




Extreme value theory for constrained physical systemsMarc Höll , Wanli Wang , and Eli Barkai*Department of Physics, Institute of Nanotechnology and Advanced Materials, Bar-Ilan University, Ramat-Gan 52900, Israel* (Received 29 June 2020; revised 15 July 2020; accepted 12 October 2020; published 29 October 2020)

We investigate extreme value theory for physical systems with a global conservation law which describes renewal processes, mass transport models, and long-range interacting spin models. As shown previously, a special feature is that the distribution of the extreme value exhibits a nonanalytical point in the middle of the support. We expose exact relationships between constrained extreme value theory and well-known quantities of the underlying stochastic dynamics, all valid beyond the midpoint in general, i.e., even far from the thermodynamic limit. For example, for renewal processes the distribution of the maximum time between two renewal events is exactly related to the mean number of these events. In the thermodynamic limit, we show how our theory is suitable to describe typical and rare events which deviate from classical extreme value theory. For example, for the renewal process we unravel dual scaling of the extreme value distribution, pointing out two types of limiting laws: a normalizable scaling function for the typical statistics and a non-normalized state describing the rare events.

DOI: [10.1103/PhysRevE.102.042141](https://doi.org/10.1103/PhysRevE.102.042141)**I. INTRODUCTION**

Extreme events are a large class of phenomena in natural and man-made systems which are uncommon compared to the usual dynamics [1–5]. Despite their rare occurrence they still can have influential consequences, e.g., the fastest sperm in fertilization [6,7], the longest trapping time in transport [8], and first passage problems in Markov processes [9]. The original problem considers a set of $N \in \mathbb{N}$ independent and identically distributed (IID) random variables (x_1, \dots, x_N) and describes the statistics of its maximum $x_{\max} = \max(x_1, \dots, x_N)$. Let $\psi(x)$ be the probability density function (PDF) of the random variables and $\Psi(x)$ the cumulative distribution function (CDF). When the maximum x_{\max} has the value m then all other random variables are less than or equal to m . So the CDF of the maximum is $\text{Prob}(x_{\max} \leq m) = \Psi^N(m)$ and hence the PDF of the maximum is obviously

$$f(m) = N\psi(m)\Psi^{N-1}(m). \quad (1)$$

A central result of classical extreme value theory (EVT) is that the limiting maximum PDF for large N converges to one of three classes of distributions called Weibull, Gumbel, or Fréchet depending on the large x behavior of $\psi(x)$ when m is shifted and rescaled appropriately [3,10–12]. However, for most systems the assumption of IID random variables has to be abandoned.

Recently EVT was studied for a wide range of different models whose common property is the global confinement of their dynamics; see [5] for a review. This global conservation induces correlations among the random variables. It is a common trait shared in many models including *renewal processes* (RP) [13–17], mass transport models such as *zero range processes* (ZRP) [18–22], and long-range interacting spin models such as the *truncated inverse distance squared Ising model*

(TIDSI) [23,24]. These three models describe numerous physical systems, including zero crossing of Brownian motion, arrival times at a detector, and interacting systems, to name only a few.

Particular attention was devoted to systems which loosely speaking are scale free, such as fractal renewal theory with diverging mean waiting time and diverging variance of the waiting time (see below). These systems exhibit large fluctuations and dominance of the extreme. It was shown previously how the global constraint may modify completely the classical EVT in the sense of strong deviations from Fréchet's law. Somewhat similarly to the classical ensembles of statistical physics—e.g., microcanonical ensembles with fixed energy, volume, and number of particles and canonical ensembles where the temperature of the bath is the constraint—the different constraints discussed below also give rich physical behaviors specific to the ensemble or model. For each model there are several classes of limiting laws in the thermodynamic limit when the global constraint diverges. These classes depend on the model parameters, and were studied for RP [8,25–29], ZRP [19,30], and TIDSI [31]. For example, for RP with fat-tailed waiting times, typical fluctuations of the maximum go through a dynamical phase transition depending on the existence or nonexistence of the mean waiting time. When the mean exists Fréchet's law holds typically; when it does not exist the behavior is completely different [27]. A similar situation exists for TIDSI in the critical phase between ferromagnetic and paramagnetic phases [31,32]. However, a particular limiting law might reflect only part of the truth. For example, Fréchet's law predicts a diverging second moment of the largest waiting time in a renewal process. However, that is impossible since all waiting times are shorter than the observation time. This does not imply that Fréchet's law is incorrect, only that it must be modified in its tail. To put it

differently, the constraint yields a natural cutoff and this modifies the description of classical EVT [8,28]. In the main text below, we discuss the thermodynamic limit for each model and how our results help to classify limiting behaviours.

Our work addresses two main themes. First, without restoring to a thermodynamic limit, we provide a complete set of relations between constrained EVT and much simpler quantifiers of the underlying stochastic dynamics. These exact relations are valid for any value of the global constraint, i.e., close to and far from the thermodynamic limit. For example, for RP we find an exact and simple relation between EVT and the mean number of renewals $\langle N \rangle$. The relations are found beyond the critical point $m > C/2$, where $C > 0$ is the global constraint. It has been recognized in earlier studies [32,33] that the analysis beyond this midpoint may be simplified for the Brownian bridge and the tied-down renewal process (which is essentially TIDSI). The importance of the midpoint is easy to understand: Once we observe a value larger than half of the global constraint it is already the maximum. No following value can be larger.

Our second goal is to exploit the exact relations and consider the thermodynamic limit. We recap known limiting laws and describe additional ones. For example, for renewal processes we find dual scaling, i.e., our theory describes both types of limiting behavior. When no moment of the waiting times exists, our theory describes typical events and rare events. When only the first moment of the waiting times exists, our theory describes the correction to Fréchet’s law (considered as rare events) and its large deviations. In this sense we go beyond previous studies of the thermodynamic limit [5,8,27,28,30,31]. We further confirm that rare events can be often described within the framework of infinite densities [8,28,34–39].

The article is constructed as follows. We consider the RP in Sec. II, the ZRP in Sec. III, and the TIDSI in Sec. IV. For all three models we derive the maximum PDF in the second half of the support and relate it to well-studied stochastic quantifiers of the underlying dynamics. There we present the analysis on the RP elaborately. Furthermore, for the RP and TIDSI we derive limiting laws of the second half maximum distribution in the thermodynamic limit for fat-tailed random variables. Section V gives a summary.

II. RENEWAL PROCESS

A. Basics

RP are widely used in physics [13–17,27], for example in describing the random arrival times of radioactive debris to a Geiger counter. Mathematically these processes are described with a PDF $\psi(\tau)$ of interarrival times, sometimes called waiting times. The process starts at time $t_1 = 0$ considered as the first event. To construct the process, first, sample τ_1 from the PDF $\psi(\tau)$ (this describes the timing of the second event), then renew the process by sampling τ_2 from the same PDF so that the timing of the third event is given by $\tau_1 + \tau_2$. The process is continued this way for N events, i.e., the i th event happens at time $t_i = \tau_1 + \tau_2 + \dots + \tau_{i-1}$ with $i \in \{1, 2, \dots\}$. The waiting times τ_1, τ_2, \dots between events are IID random variables all sampled from $\psi(\tau)$. The PDF

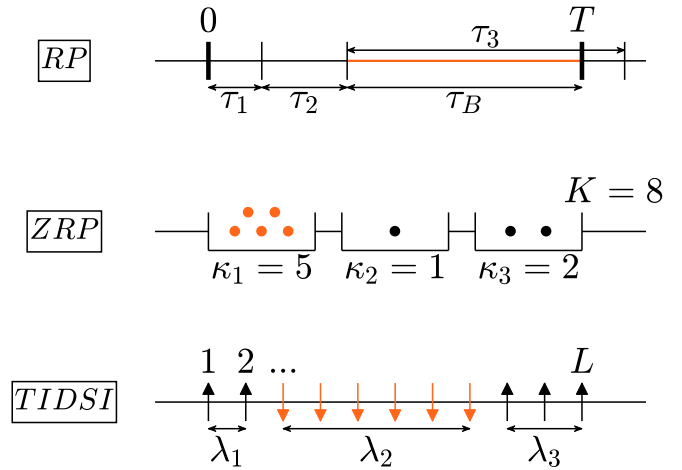


FIG. 1. Schematic figure of the three models, *renewal processes*, *zero range processes*, and the *truncated inverse distance squared Ising model*, presented in the main text. The maximum in each model is colored orange. The ZRP is described in Sec. IV and TIDSI in Sec. IV.

of $\psi(\tau)$ can be either thin tailed or fat tailed, and this has major consequences on the behavior of the extreme events. For example, an exponential (thin-tailed) PDF $\psi(\tau)$ describes arrival times of independent photons to a detector. An example of a fat-tailed process is the zero crossing of Brownian motion where $\psi(\tau) \sim \tau^{-3/2}$; similarly for blinking quantum dots [40,41] or times between jumps in the anomalous continuous time random walks [42,43].

The renewal process is observed at the observation time $t = T$. The fixed observation time T is the sum of all waiting times before the last event $i = N$ added with the backward recurrence time

$$T = \sum_{i=1}^{N-1} \tau_i + \tau_B; \tag{2}$$

see Fig. 1. The backward recurrence time τ_B is the time interval between the last event $i = N$ and the observation time T . It is distributed differently than the waiting times [44]. The constraint of a fixed observation time implies that the amount of events N is a random number. This and the cutoff of the last time interval τ_N to τ_B make the set of all waiting times $\{\tau_1, \tau_2, \dots, \tau_{N-1}, \tau_B\}$ non-IID.

B. Overview of constrained models

Before we continue with the maximum statistics of the waiting times $\{\tau_1, \tau_2, \dots, \tau_{N-1}, \tau_B\}$, we compare model details of RP with the two later studied models, ZRP and TIDSI. The common trait of these models is that the sum of the random variables

$$C = \sum_{i=1}^N x_i \tag{3}$$

is fixed to the global constraint $C > 0$. For example for the RP, the constraint is the observation time $C = T$ and the random variables are $x_i = \tau_i$ when $i \in [1, N - 1]$ and $x_N = \tau_B$. In Table I important characteristics are presented with the

TABLE I. Overview of details of the three models, *renewal processes*, *zero range processes*, and the *truncated inverse distance squared Ising model*, presented in the main text. Their relevant random variables, the constraint, and the randomness of N are shown. The ZRP is described in Sec. IV and TIDSI in Sec. IV.

	RP	ZRP	TIDSI
Random variables x_i	Waiting times τ_i	Number of particles κ_i	Domain lengths λ_i
Values of x_i	Continuous	Discrete	Discrete
Constraint C	Observation time $T = \sum_{i=1}^{N-1} \tau_i + \tau_B$	Total number $K = \sum_{i=1}^N \kappa_i$	Total length $L = \sum_{i=1}^N \lambda_i$
Number N	Random	Fixed	Random

appropriate x_i and C for each model. Figure 1 shows a schematic figure of the three models. We do not define precisely ZRP and TIDSI at this stage; we will do so later in Secs. III and IV. For now we just want to define their global constraint: For ZRP, C is the total number of particles in a system where particles are distributed in boxes, while for TIDSI, describing an interacting spin system, C is the size of the system. In both models one can say that interactions are local, i.e., only particles within a box interact and only spins within a given domain.

C. Extreme value statistics

We investigate the statistics of the maximum waiting time [8,27,28]

$$\tau_{\max} = \max(\tau_1, \tau_2, \dots, \tau_{N-1}, \tau_B). \quad (4)$$

The maximum τ_{\max} is also called the extreme event of the waiting times. The maximum PDF is defined by $f(m; T) = dF(m; T)/dm$. The maximum CDF $F(m; T) = \text{Prob}(\tau_{\max} \leq m)$ is the probability of the random variable τ_{\max} being less than or equal to m . Clearly, the maximum is constrained by $0 < m \leq T$. Since the number of events N is random, it is instructive to consider

$$f(m; T) = \sum_{N=1}^{\infty} f_N(m; T) \quad (5)$$

with $f_N(m; T) = dF_N(m; T)/dm$ being the maximum PDF with exactly N renewal events. In this context the value of N is a sampled value. The maximum CDF with exactly N events is given by [27]

$$F_N(m; T) = \int_0^m \dots \int_0^m \int_0^m \psi(\tau_1) \dots \psi(\tau_{N-1}) \varphi(\tau_B) \times \delta \left[T - \left(\sum_{i=1}^{N-1} \tau_i + \tau_B \right) \right] d\tau_1 \dots d\tau_{N-1} d\tau_B. \quad (6)$$

This is the probability of τ_{\max} being less than or equal to m when exactly N events have happened. Here, the survival probability

$$\varphi(\tau_B) = \int_{\tau_B}^{\infty} \psi(\tau) d\tau \quad (7)$$

is the probability that no other event than the first one at $t_1 = 0$ occurs until time τ_B . Equation (6) is easy to interpret: the set of waiting times $\{\tau_1, \dots, \tau_B\}$ are all less than or equal to m , and the delta function is the constraint. Since we will use this formula below we write the N -multiple integral shorter as

$$F_N(m; T) = \int_0^m \psi(\tau_1) \dots \psi(\tau_{N-1}) \varphi(\tau_N) \delta(T - \|\boldsymbol{\tau}\|_1) d\boldsymbol{\tau} \quad (8)$$

with the N -vector $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N)^T$ and the taxicab norm $\|\boldsymbol{\tau}\|_1 = \sum_{i=1}^N \tau_i$.

Before we continue our analytical investigation, let us take a look at simulation results with which we construct the PDF of τ_{\max} . In Fig. 2 we simulate the process and obtain the histograms for $f(m; T)$, where we used the exponential waiting time PDF $\psi(\tau) = \exp(-\tau)$, the Pareto waiting time PDF $\psi(\tau) = \alpha \tau^{-1-\alpha}$ with $\tau \geq 1$, and the one-sided Lévy waiting time PDF $\psi(\tau) = 1/(2\sqrt{\pi}) \tau^{-3/2} \exp[-1/(4\tau)]$. All examples show a discontinuity at the midpoint of the support,

$$m = \frac{T}{2}. \quad (9)$$

The Pareto waiting time PDF also yields an intrinsic discontinuity at $m = 9$ because $\tau \geq 1$. The importance of $T/2$ can be intuitively understood: Once a waiting time is larger than $T/2$ it is then also the maximum waiting time. No previous and following waiting time can be larger. Since the PDF of τ_{\max} is nonanalytical we cannot expect to find a global solution in the whole range $0 < m < T$. Importantly, all the results in Fig. 2 are presented for finite simulation time and far from the thermodynamic limit. For example, for the exponential process the mean waiting time is unity and the measurement time is just twice as large. Usually one does not expect general statistical laws to emerge at this limit. However, in the Fig. 2 we show a theory (derived below) that works perfectly beyond the midpoint. Our goal is then to present this theory and only later consider the thermodynamic limit.

D. Extreme value statistics in the second half

We now present the main result and its derivation afterwards. In the second half $T/2 < m < T$, the maximum PDF $f(m; T)$ is exactly related to the mean number of renewals $\langle N(t) \rangle$, namely

$$f(m; T) = \varphi(m) R(T - m) + \psi(m) \langle N(T - m) \rangle. \quad (10)$$

The function $R(T - m)$ is the rate of producing these events, namely the derivative of $\langle N \rangle$. Both $\langle N \rangle$ and R are thoroughly investigated in the physical and mathematical literature [13–16]. We find also an elegant formula of the maximum CDF,

$$F(m; T) = 1 - \varphi(m) \langle N(T - m) \rangle, \quad (11)$$

again in the second half $T/2 < m < T$. The formulas (10) and (11) present exact results for any $m > T/2$ and are very useful, as they allow us to derive both finite-time expressions and also the long-time limit (see below).

Both relationships yield insight into the maximum m when it is roughly of order T . Then we need to have information on $\langle N(T - m) \rangle$, which includes $R(T - m)$, only for very short time. Namely, once we have $\langle N(T - m) \rangle$ for time $T - m$ we

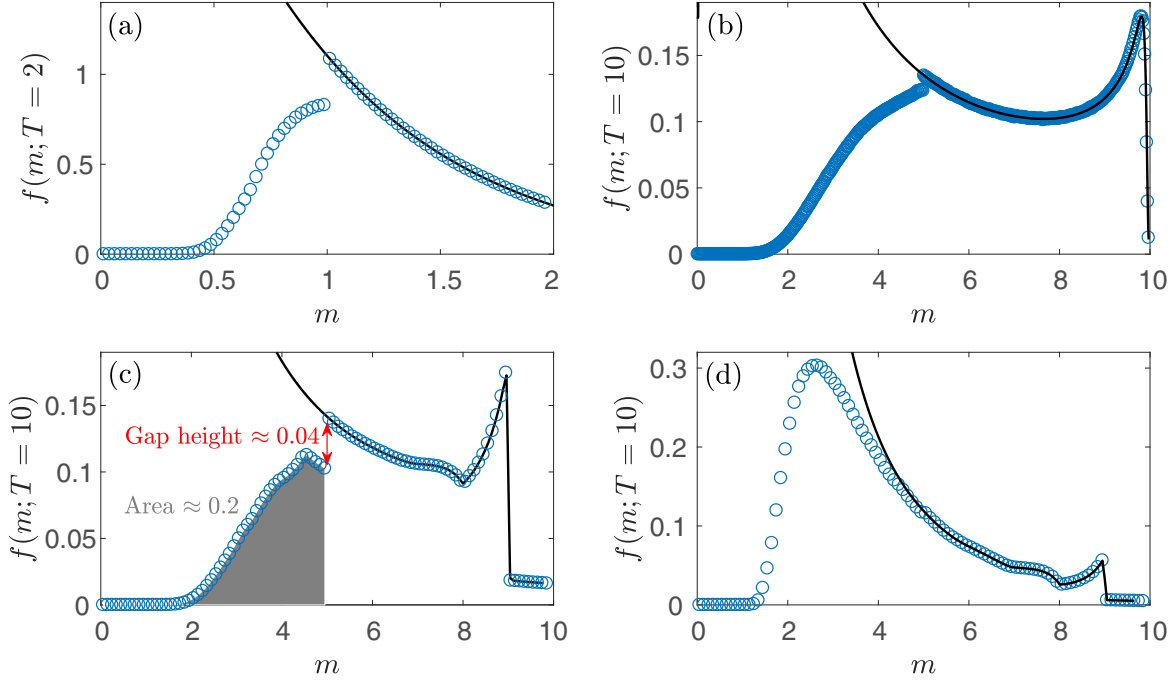


FIG. 2. Histogram of the maximum PDF $f(m; T)$ of RP from Monte Carlo simulations (blue circles) compared with the theory of Eq. (20) (black line) for (a) exponential $\psi(\tau) = \exp(-\tau)$ with $T = 2$, (b) one-sided Lévy $\psi(\tau) = 1/(\sqrt{2\pi})\tau^{-3/2} \exp[-1/(4\tau)]$ with $T = 10$ and (c), (d) Pareto $\psi(\tau) = \alpha\tau^{-\alpha}$, $\tau > 1$, with $T = 10$. The simulations were performed with 10^7 realizations. The analytical expressions of $\langle N \rangle$ and R were obtained via numerical inverse Laplace transform; see Eq. (25). The nonanalytical point $T/2$ is very visible in (a), (b), and (c). It is barely visible for $\alpha = 3/2$ in (d) because T is relatively large. In addition, we show the theoretical gap height, Eq. (30), and the theoretical area of the first half, Eq. (28), exemplarily in (c); both match with their numerical estimates.

can predict EVT for m . Intuitively, to observe a large m of order T , we need the maximum to be produced close to the start of the process.

We will now present the derivation of this main result. Taking the derivative of $F_N(m; T)$ from Eq. (6) or (8) yields two terms with each a $(N - 1)$ -multiple integral:

$$f_N(m; T) = \underbrace{\varphi(m) \int_0^m \prod_{i=1}^{N-1} \psi(\tau_i) \delta(T - m - \|\tau\|_1) d\tau}_{\mathcal{B}} + \underbrace{\psi(m)(N-1) \int_0^m \prod_{i=1}^{N-2} \psi(\tau_i) \varphi(\tau_{N-1}) \delta(T - m - \|\tau\|_1) d\tau}_{\mathcal{NB}}. \quad (12)$$

Here $\tau = (\tau_1, \dots, \tau_{N-1})^T$. The first term describes backward and the second term nonbackward processes:

$$\begin{aligned} \mathcal{B} &= \text{renewal processes with } \tau_{\max} = \tau_B, \\ \mathcal{NB} &= \text{renewal processes with } \tau_{\max} \neq \tau_B. \end{aligned} \quad (13)$$

The two integrals in Eq. (12) are special cases of this general integral:

$$I_{N-1}(m, T') = \int_0^m \prod_{i=1}^{N-1} g_i(\tau_i) \delta(T' - \|\tau\|_1) d\tau. \quad (14)$$

We assume general positive functions g_i and an arbitrary constant $T' > 0$. Compared with Eq. (12) it is $T' = T - m$ and the functions g_i are either ψ or φ . When we restrict the regime $T' < m$ [for Eq. (12) it means $T/2 < m$] then this integral is

identical to the $(N - 1)$ -fold convolution

$$I_{N-1}(m, T') = (g_1 * \dots * g_{N-1})^{(N-1)}(T'), \quad (15)$$

which we prove rigorously in Appendix A. The twofold convolution is defined as $(g_1 * g_2)(T') = \int_0^{T'} g_1(\tau_1) g_2(T' - \tau_1) d\tau_1$, and higher-order convolutions are obtained successively. Equation (15) means that the upper limit of the integration m is reduced to $T' < m$. That is because the constraint $T = \|\tau\|_1$ forces all individual τ_i to be less than T' . Therefore the integration from T' to m yields zero. Hence we remain with the convolution. In Appendix A we show Eq. (15) in detail. Importantly, we realize that this decoupling trick, valid whenever $T' < m$, is a very general theme. We use this trick also below for the two other models, i.e., the ZRP and the TIDSI.

We see now why we consider the maximum PDF instead of the maximum CDF. The delta function of the maximum

PDF depends on $T - m$. We set now $T' = T - m$ and apply Eq. (15) to Eq. (12) under the assumption of the second half $T/2 < m < T$, we obtain exactly

$$f_N(m; T) = \varphi(m)Q_{N-1}(T - m) + \psi(m)(N - 1)P_{N-1}(T - m). \quad (16)$$

Here we introduced two quantities well known from renewal theory [35]. The first quantity is the distribution $Q_{N-1}(t) = \langle \delta(t - \|\tau\|_1) \rangle$ of having the N th renewal event exactly at time t . It can be written as the iteration equation

$$Q_{N-1}(t) = (Q_{N-2} * \psi)(t) \quad (17)$$

with $Q_0(t) = \delta(t)$. The second quantity $P_{N-1}(t)$ is the probability of finding $N - 1$ renewal events up to time t . Both are connected with the survival probability via

$$P_{N-1}(t) = (Q_{N-1} * \varphi)(t). \quad (18)$$

In case of a single event process $N = 1$, we get from Eq. (16) that

$$f_1(m; T) = \varphi(m)\delta(T - m), \quad (19)$$

which describes the delta peak of the maximum PDF $f(m; T)$ at $m = T$.

The first term of Eq. (16) means that the last waiting time τ_B is maximum and the second term describes the $N - 1$ other cases where the maximum ended before T . When the last waiting time is maximum then at time $T - m$ exactly N events have happened, which gives $Q_{N-1}(T - m)$. This is multiplied with the probability of not having an event during m , namely $\varphi(m)$. Hence we have the first term. Now the second term consists of $\psi(m)$, i.e., the maximum ended before T , and $(N - 1)P_{N-1}(T - m)$. It simply means that we had $N - 1$ events in the remaining time $T - m$. Note that any of the $N - 1$ waiting times excluding the backward recurrence time might be the largest, so the second term is multiplied by $N - 1$. Hence we have the second term.

Summing up all number of events in Eq. (16) yields the main result of this section, namely the maximum PDF when $T/2 < m < T$ exactly given by

$$f(m; T) = \varphi(m)R(T - m) + \psi(m)\langle N(T - m) \rangle. \quad (20)$$

Thus, we derived Eq. (10). The first term contains the rate function

$$R(T - m) = \sum_{N=1}^{\infty} Q_{N-1}(T - m), \quad (21)$$

which is the probability of finding some event exactly at time $T - M$, [35]. The delta function $\delta(T - M)$ from $N = 1$ does not contribute since $m < T$, and further we have already pointed out the behavior of the solution when the maximum is equal to the observation time, Eq. (19). The second term in Eq. (20) contains the mean number of renewal events,

$$\langle N(T - m) \rangle = \sum_{N=1}^{\infty} NP_N(T - M). \quad (22)$$

It is related to the rate function via the definite integral

$$\langle N(T - M) \rangle = \int_0^{T-M} R(t)dt. \quad (23)$$

Note that if T is large but we limit ourselves to rare events when also m is large, such that $T - m$ is small, Eq. (20) states that all we need to evaluate is the short-time behavior R and $\langle N \rangle$.

Although R and $\langle N \rangle$ are well investigated observables within renewal theory, still an exact and explicit analysis of Eq. (20) is difficult due to the convolutions. As is well known, it is beneficial to analyze such problems in Laplace space. The Laplace transform of some function $h(t)$ is defined by

$$\widehat{h}(s) = \mathcal{L}_{t \rightarrow s}\{h(t)\} = \int_0^{\infty} h(t)e^{-st} dt. \quad (24)$$

The Laplace transform of Eq. (20) with respect to the observation time T is

$$\widehat{f}(m; s) = \varphi(m) \frac{e^{-sm}}{1 - \widehat{\psi}(s)} + \psi(m) \frac{e^{-sm}}{s[1 - \widehat{\psi}(s)]}, \quad (25)$$

which is easy to prove with the convolution theorem of Laplace transforms. In detail, we used $\widehat{Q}_N(s) = \widehat{\psi}^N(s)$ and $\widehat{P}_N(s) = \widehat{\psi}^{N-1}(s)[1 - \widehat{\psi}(s)]/s$ and the geometric series. Note that Eq. (25) is only valid for inverse Laplace transforms $\mathcal{L}_{s \rightarrow T}^{-1}$ when $T/2 < m < T$.

E. Maximum CDF in the second half

From normalization, the maximum CDF at $m = T$ is clearly $F(T; T) = 1$. The CDF is discontinuous due to samples with the only renewal at $t_1 = 0$. We separate the contribution from these realizations with single renewal events, described by Eq. (19), and the remaining processes where we had at least two renewals. Therefore we have

$$\lim_{m \rightarrow T} F(m; T) + \varphi(T) = 1, \quad (26)$$

where $F(T; T) = \varphi(T)$ is the probability of $m = T$. With this boundary condition, we may integrate the PDF $f(m; T)$ and then get the maximum CDF for $T/2 < m < T$ as

$$F(m; T) = 1 - \varphi(m)\langle N(T - m) \rangle. \quad (27)$$

Thus, we derived Eq. (11). In Fig. 3 we simulate both sides of this formula and find perfect matching in the second half. Especially the estimation of the right hand side, i.e., the estimation of $\langle N \rangle$ and putting this into $1 - \varphi(m)\langle N(T - m) \rangle$, demonstrates how our theory can be used empirically to find the extreme value statistics in an indirect way.

In principle, the same can be done for f .

Specifically the probability of finding the maximum τ_{\max} in the first half time $m < T/2$ is

$$F(T/2; T) = 1 - \varphi(T/2)\langle N(T/2) \rangle, \quad (28)$$

which is valid for all waiting time PDFs $\psi(\tau)$.

F. Gap height of the maximum PDF at the midpoint

The nonanalytical behavior at $T/2$ arises from double event processes with $N = 2$ renewals. The set of waiting times is (τ_1, τ_B) . The maximum of this set is always larger than $T/2$, i.e., the probability of $m < T/2$ is zero. We can quantify this

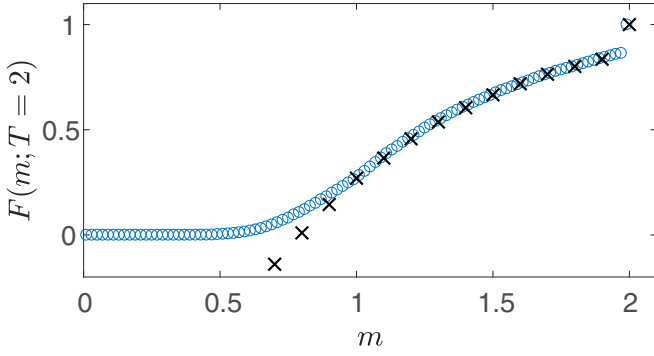


FIG. 3. Comparison of two approaches to estimate the CDF $F(m; T)$ with $T = 2$ for exponential $\psi(\tau) = \exp(-\tau)$: first by direct estimation (blue circles) and second by estimation of $\langle N \rangle$ and using Eq. (27) (black crosses). The number of simulations is 10^4 . As predicted by Eq. (27), both curves match for the second half $T/2 < m < T$. At $m = T$ the discontinuity is also captured by both curves because $\langle N(0) \rangle = 0$. Equation (27) can also be calculated analytically as $1 - \exp(T - m + 1)$.

with

$$f_2(M; T) = \begin{cases} 0 & \text{if } m < T/2, \\ \varphi(m)\psi(T - m) & \text{if } m < T/2, \\ +\psi(m)\phi(T - m) & \text{if } m > T/2, \end{cases} \quad (29)$$

which is derived from Eq. (12). Thus the height of the gap between the first and second half time expressions of $f(M; T)$ at $T/2$ is

$$f((T/2)^+; T) - f((T/2)^-; T) = 2\psi(T/2)\varphi(T/2). \quad (30)$$

Here $(T/2)^\pm$ means we approach $T/2$ from left/right. When $T \rightarrow \infty$ the gap closes, i.e., tends to zero. This prediction is later verified for the simulations presented in Fig. 5. Similarly, for $N \geq 3$ the maximum has to be larger than T/N . So the PDF $f(m; T)$ is nonanalytical at points T/N which all are in the first half time $m < T/2$; see [27].

G. Long-time limits for fractal renewal processes

We calculate the long-time limit $T \rightarrow \infty$ of $f(m; T)$ in the second half $T/2 < m < T$, i.e., Eq. (20), for power law waiting time PDFs

$$\psi(\tau) \sim b_\alpha \tau^{-1-\alpha} \quad (31)$$

with $\alpha \in (0, 1)$ or $\alpha \in (1, 2)$. The second half implies that we are dealing with large values of m . Therefore we consider the linear order $m = O(T)$ when the maximum is of the order of the observation time. Then the second half shows rich behavior for $f(m; T)$ for the power law waiting time PDF (31) as described now.

Since $R(T - m)$ and $\langle N(T - m) \rangle$ depend on the remaining time $T - m$, we have to specify how the remaining time $T - m$ behaves. We first consider linear order $T - m = O(T)$; see Fig. 4. Hence we have to calculate the long-time limit of R and $\langle N \rangle$. This is equivalent to calculating the small s behavior of Eq. (25). The small s behavior of the waiting time PDF is

$$\widehat{\psi}(s) \sim \begin{cases} 1 - b_\alpha |\Gamma(-\alpha)| s^\alpha & \text{for } \alpha \in (0, 1), \\ 1 - \langle \tau \rangle s & \text{for } \alpha \in (1, 2) \end{cases} \quad (32)$$

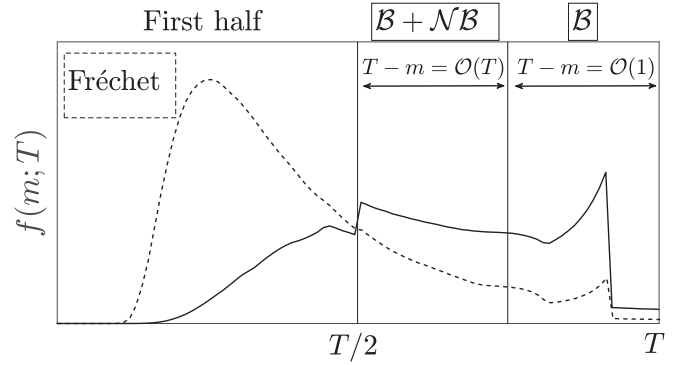


FIG. 4. Presentation of the two scaling regimes in the second half $T/2 < m < T$, which we apply for power law waiting time PDFs with $\alpha \in (0, 1)$ (black line) and $\alpha \in (1, 2)$ (dashed line). When the remaining time scales linearly, $T - m = O(T)$, both processes \mathcal{B} and \mathcal{NB} contribute. We obtain the scaling function $\mathcal{G}(m/T)$ in Eq. (34) for $\alpha \in (0, 1)$ and $\mathcal{I}(m/T)$ in Eq. (36) for $\alpha \in (1, 2)$. When $T - m = O(1)$ then only \mathcal{B} contributes, which is related to the rate function; see Eq. (38). For $\alpha \in (0, 1)$, there is a matching between \mathcal{G} and the $T - m = O(1)$ regime but not between the first half time and \mathcal{G} because $f(m; T/2)$ is not differentiable. For $\alpha \in (1, 2)$, there is a matching between Fréchet's law and \mathcal{I} and also between \mathcal{I} and the $T - m = O(1)$ regime. For a detailed analysis of small m we refer to [27].

with the mean waiting time $\langle \tau \rangle = \int_0^\infty \tau \psi(\tau) d\tau$.

In case of $\alpha \in (0, 1)$ we obtain from the small s behavior of Eq. (25) the scaling law of the second half maximum PDF as

$$f(m; T) \sim \frac{1}{T} \mathcal{G}\left(\frac{m}{T}\right) \quad (33)$$

with

$$\mathcal{G}(\xi) = \frac{\sin(\pi\alpha)}{\pi} \xi^{-\alpha} (1 - \xi)^{\alpha-1} + \frac{\sin(\pi\alpha)}{\pi} \xi^{-\alpha-1} (1 - \xi)^\alpha. \quad (34)$$

The rescaled variable is $\xi = m/T$. The first term represents processes \mathcal{B} and the second term \mathcal{NB} . Of course one can sum both terms and get the right-hand side as $\sin(\pi\alpha)/\pi \xi^{-1-\alpha} (1 - \xi)^{\alpha-1}$ which was already found in [27,45]. This function is valid for $1/2 < \xi < 1$ due to the restriction on the second half. The midpoint $\xi = 1/2$ is nonanalytical, as can be seen by the kink in Fig. 5 where we compare the theory with Monte Carlo simulations. Note that for $\xi \rightarrow 1$ the function blows up to infinity. In reality, for any finite observation time the maximum PDF does not diverge. Hence later we cure this problem by considering constant remaining time $T - m = O(1)$; see also Fig. 4. That analysis will show how a second scaling law describes rare events.

In the case of $\alpha \in (1, 2)$ we obtain from the small s behavior of Eq. (25) the scaling law of the second half maximum PDF as

$$f(m; T) \sim \frac{1}{T^\alpha} \mathcal{I}\left(\frac{m}{T}\right) \quad (35)$$

with

$$\mathcal{I}(\xi) = \frac{b_\alpha}{\langle \tau \rangle} \xi^{-1-\alpha} (1 - \xi) + \frac{b_\alpha}{\alpha \langle \tau \rangle} \xi^{-\alpha}. \quad (36)$$

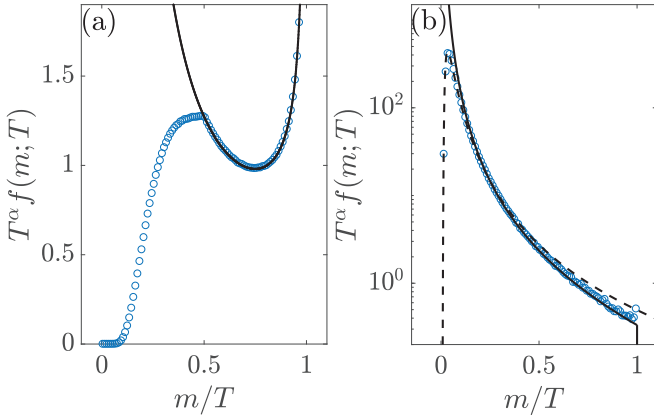


FIG. 5. Rescaled histogram from simulations (blue circles) for the Pareto waiting time PDF with (a) $\alpha = 1/2$ and (b) $\alpha = 3/2$ compared with the theory (black line), i.e., (a) \mathcal{G} of Eq. (34) and (b) \mathcal{I} of Eq. (36). In the latter, we also plotted Fréchet's law, Eq. (37) (dashed line). The number of realizations is 5×10^7 . The observation times are (a) $T = 5000$ and (b) $T = 1000$. Note that when $T \rightarrow \infty$ the discontinuity for $\alpha \in (0, 1)$ is with respect to the derivation of $f_1(m; T)$ at $T/2$, while for finite time T we observe a gap discontinuity of the maximum PDF itself; see Fig. 2.

The rescaled variable is $\xi = m/T$. The first term represents processes \mathcal{B} and the second term $\mathcal{N}\mathcal{B}$. This formula was already found in the context of the big jump principle in physical modeling for the two-state Lévy walk [8,28,46,47]. The important point of this scaling law is that it cures the nonphysical diverging second moment of Fréchet's law as pointed out by [28]. Simply put, most values of the random variable τ_{\max} are found for values below $m < T/2$ when T is large. These typical events follow Fréchet's law,

$$f(m; T) \sim \frac{1}{(T/\langle\tau\rangle)^{1/\alpha}} b_\alpha \xi^{-1-\alpha} \exp\left(-\frac{b_\alpha \xi^{-\alpha}}{\alpha}\right) \quad (37)$$

with $\xi = m(T/\langle\tau\rangle)^{-1/\alpha}$, i.e., $m = O(T^{1/\alpha})$. See [27] for a rigorous derivation. But Fréchet's law predicts the divergence of the variance of τ_{\max} , which is nonphysical since $m \leq T$. The scaling law (37) matches with Fréchet's law: The small m behavior of Eq. (35) equals the large m behavior of Fréchet's law (37), namely $b_\alpha T m^{-1-\alpha}/\langle\tau\rangle$. Hence both scaling regimes, i.e., Fréchet and the far tail, are complementary. In Fig. 5 we compare Eq. (35) with numerical simulations and Fréchet's law (37). We see that the gap at the midpoint vanishes. Furthermore, as explained in [8,28,34], the function \mathcal{I} of Eq. (36) is an infinite covariant density because it is non-normalizable. However, it describes the second moment of $f(m; T)$.

In summary, although the limit laws (33) and (35) are known, we showed that they arise from the second half distribution, Eq. (20). It is not surprising for $\alpha \in (0, 1)$ but it is for $\alpha \in (1, 2)$ because of the different behavior of the nonanalytical midpoint in the thermodynamic limit. We now present a result unraveled by our approach.

H. Long-time limits for fractal renewal processes with constant remaining time

Here we calculate the long-time limit of $f(m; T)$ in the second half time $T/2 < m < T$, i.e., Eq. (20), also with

$m = O(T)$ but now we consider constant remaining time $T - m = O(1)$; see Fig. 4. Hence the rate function $R(T - m)$ and the mean number of events $\langle N(T - m) \rangle$ stay constant in Eq. (20). So we only have to compare their prefactors $\varphi(m) \sim b_\alpha m^{-\alpha}/\alpha$ and $\psi(m) \sim b_\alpha m^{-1-\alpha}$. The first one is dominant and scales as $b_\alpha T^{-\alpha}/\alpha$. The second term can be neglected. Therefore we find the scaling law

$$f(m; T) \sim \frac{1}{T^\alpha} \frac{b_\alpha}{\alpha} R(T - m), \quad (38)$$

which is valid for both cases $\alpha \in (0, 1)$ and $\alpha \in (1, 2)$. This formula means that the maximum waiting time is always the last one; i.e., this long-time limit comes solely from the process \mathcal{B} . If a waiting time is the maximum but not the last one then it has to end exactly in such a way that the remaining time $T - m$ is of order 1. But as T increases this probability becomes zero so that only \mathcal{B} contributes. Furthermore, the remaining time can be a small value and therefore the full form of the waiting time PDF $\psi(\tau)$ [and consequently the full form of $R(T - m)$] is required. This is in contrast to the previous study of $T - m = O(T)$ where the asymptotic behavior of $\psi(\tau)$ fully describes the scaling of the maximum PDF in the second half.

The scaling function $b_\alpha R(T - m)/\alpha$ is obviously non-normalizable because

$$T^\alpha \int_0^T f(m; T) dm \sim \frac{b_\alpha}{\alpha} \int_0^\infty R(\epsilon) d\epsilon \rightarrow \infty \quad (39)$$

with $\epsilon = T - m$. Technically, the integral over m shown here is only correct for $m > T/2$, but this does not change the divergence. However, $b_\alpha R(T - m)/\alpha$ matches with integrable scaling PDFs: For $\alpha \in (0, 1)$ it matches with \mathcal{G} and for $\alpha \in (1, 2)$ it matches with \mathcal{I} which matches itself with Fréchet's law. The small $T - m$ limit of $f(m; T)$ with $T - m = O(T)$ is the same as the large $T - m$ limit of $f(m; T)$ with $T - m = O(1)$. For $\alpha \in (0, 1)$ this is

$$f(m; T) \sim \frac{\sin(\pi\alpha)}{\pi} T^{-\alpha} (T - m)^{\alpha-1} \quad (40)$$

and for $\alpha \in (1, 2)$ this is

$$f(m; T) \sim \frac{b_\alpha}{\alpha\langle\tau\rangle} T^{-\alpha}. \quad (41)$$

In Fig. 6 we simulate $f(m; T)$ for two different values of α and compare $T^\alpha f(m; T)$ plotted over $T - m$ with $b_\alpha/\alpha R(T - m)$ of Eq. (38) and also with the matching functions of Eqs. (40) and (41). We find that simulation and theory match. The difference between Eq. (38) and the matching functions of Eq. (40) and Eq. (41) is easy to see only when $T - m$ is relatively small. So the analysis of the $T - m = O(1)$ regime is suitable to describe rare events very close to T .

III. ZERO RANGE PROCESS

A. Basics

Zero range processes in equilibrium describe a system with a fixed number K of interacting particles. These particles are located in well separated traps where transition times between the traps are very fast. We have N such traps, and in each trap

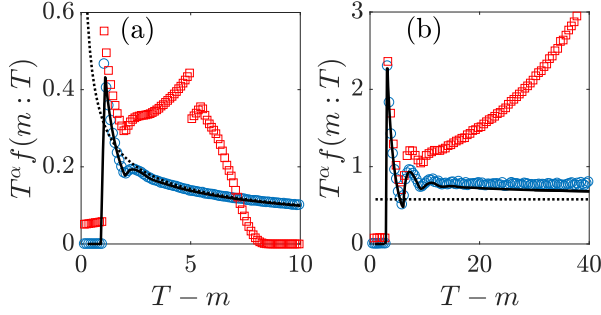


FIG. 6. Rescaled histogram from simulations (red squares and blue circles) compared with the theory $b_\alpha/\alpha R(T - m)$ of Eq. (38) (black line) and the matching functions of (a) Eq. (40) and (b) Eq. (41) (dotted lines). The simulations were performed for two different times in each figure: (a) $T = 10$ (red squares) and $T = 1000$ (blue circles), and (b) $T = 100$ (red squares) and $T = 1000$ (blue circles). The number of realizations is (a) 10^8 and (b) 5×10^7 . Note that we used $\tau_0 = 3$ and not $\tau_0 = 1$ in (b) to reduce computation time.

$i \in [1, N]$ we have $\kappa_i \geq 0$ particles. Clearly the constraint is

$$K = \sum_{i=1}^N \kappa_i; \quad (42)$$

see Fig. 1 and Table I. Here $\psi(\kappa)$ is the probability of finding κ_i particles in the trap i . In thermal equilibrium, $\psi(\kappa)$ is the Boltzmann factor, though more generally it depends on the microscopical description of the transitions [23,24]. In this model the number of traps N is fixed, unlike the random number of renewals in the previous model. A well-studied phenomenon in this model is condensation [18–22]. When the density of the system K/N crosses a critical value, a macroscopic number of particles may occupy one trap. It is then natural to wonder what is the distribution $f(m; K)$ of the maximum $\kappa_{\max} = \max(\kappa_1, \dots, \kappa_N)$ with value m since that describes the statistical properties of the condensation [5,19,30].

B. Extreme value statistics in the second half

We investigate the statistics of the maximum particle number [5,19,30]

$$\kappa_{\max} = \max(\kappa_1, \kappa_2, \dots, \kappa_N). \quad (43)$$

The maximum probability mass function (PMF) is defined by $f_N(m; K) = F_N(m; K) - F_N(m - 1; K)$. The maximum CDF was derived in [30] and is given by

$$F_N(m; K) = \frac{1}{Z_N(K)} \sum_{\kappa=0}^m \prod_{i=1}^N \psi(\kappa_i) \delta_{K, \|\kappa\|_1}. \quad (44)$$

Similar to Eq. (8) (but there for integrals) we just wrote the N -multiple sums in short form as $\sum_{\kappa=0}^m = \sum_{\kappa_1=0}^m \cdots \sum_{\kappa_N=0}^m$ with the N -vector $\tau = (\tau_1, \dots, \tau_N)^T$. The partition function is the N -fold convolution

$$Z_N(K) = (\psi * \cdots * \psi)^{(N)}(K); \quad (45)$$

see [30]. The twofold convolution for discrete functions is $(\psi * \psi)^{(2)}(K) = \sum_{\kappa_1=0}^K \psi(\kappa_1) \psi(K - \kappa_1)$ and higher orders

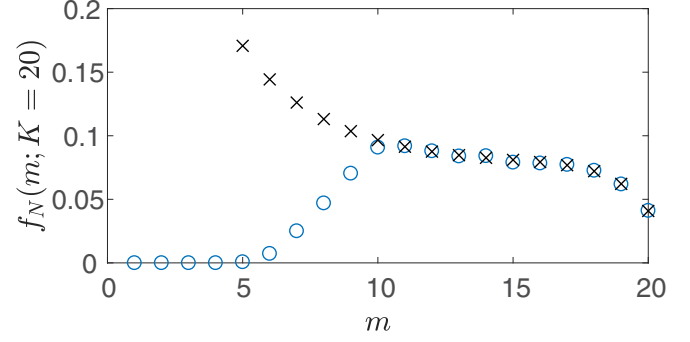


FIG. 7. Histogram of the maximum PMF $f_N(m; K)$ of ZRP from Monte Carlo simulations (blue circles) compared with the theory of Eq. (47) (black crosses) for zeta $\psi(\kappa) = 1/\zeta(1 + \alpha)(\kappa + 1)^{-1-\alpha}$ with $\alpha = 1/2$, $K = 20$, and $N = 5$. The simulations were performed with 10^7 realizations. The analytical expression of the PMF of the sum of IID random variables $\Phi_{N-1}(K - m)$ [as well as $Z_N(K)$] is obtained via inverse z transform. The kink at the midpoint $K/2$ is visible although the random variables are discrete.

are defined successively. Equation (44) is easy to interpret, the set of particle numbers $\{\kappa_1, \dots, \kappa_N\}$ are all less than or equal to m , and the Kronecker delta is the constraint.

Similarly to the RP we can calculate the maximum PMF for the second half $K/2 < m < K$ with the almost identical analysis. The maximum PMF is generally derived from Eq. (44) as

$$f_N(m; K) = \frac{N\psi(m)}{Z_N(K)} \sum_{\kappa=0}^m \prod_{i=1}^{N-1} \psi(\kappa_i) \delta_{K-m, \|\kappa\|_1}. \quad (46)$$

Here $\kappa = (\kappa_1, \dots, \kappa_{N-1})^T$. This formula is similar Eq. (12), i.e., the maximum PDF for RP in the second half, but the integrals are replaced by sums and the delta function is replaced by the Kronecker delta. Also the ZRP formula has only one term because there exists no \mathcal{B} . In Appendix A we show that the sums in Eq. (46) are identical to the N -fold convolution in the range $K/2 < m < K$. This means that the maximum PMF in the second half is

$$f_N(m; K) = \frac{1}{Z_N(K)} N\psi(m) \Phi_{N-1}(K - m). \quad (47)$$

Here, we introduced

$$\Phi_{N-1}(K - m) = (\psi * \cdots * \psi)^{(N-1)}(K - m), \quad (48)$$

which is the PMF with value $K - m$ of the sum of $N - 1$ discrete IID random variables whose common PMF is $\psi(\kappa)$. So Eq. (47) relates extreme statistics with one of the most well-studied problems in stochastic theory: the sum of IID random variables; in physics this is simply the problem of an $N - 1$ step random walk. In addition, we see here a useful modification of the classical EVT case Eq. (1). The CDF $\Psi^{N-1}(m)$ is replaced by $\Phi_{N-1}(K - m)$ which is also divided by $Z_N(K)$. When the maximum particle number κ_{\max} is m , all other particle numbers add up to the remaining number $K - m$ due to the constraint. In Fig. 7 we compare theory and simulation.

C. Relationship to condensation

From the joint PMF of the particle numbers,

$$p_N(\kappa; K) = \frac{1}{Z_N(K)} \psi(\kappa_1) \cdots \psi(\kappa_N) \delta_{K, \|\kappa\|_1}, \quad (49)$$

we obtain the well-studied single trap distribution of the particle number,

$$\rho_N(m; K) = \frac{1}{Z_N(K)} \psi(m) \Phi_{N-1}(K - m), \quad (50)$$

by summing over $N - 1$ random variables; see [30]. Comparing with Eq. (47) yields

$$f_N(m; K) = N \rho_N(m; K) \quad (51)$$

for the second half $K/2 < m < K$. This is a modification of Eq. (1) when we set $\Psi(m) = 1$ due to the constraint. This result was obtained in [19] as a limiting law in the condensation phase of the model. Our result shows that it is exactly valid close to and far from the thermodynamic limit, regardless of the occurrence of condensation. It is independent of the structure of $\psi(\kappa)$. Hence, our result provides a general connection between EVT and the single trap distribution of the particle number. We refer to [5, 19, 30] where the thermodynamic limit of $\rho_N(m; K)$ was studied.

Finally, the main interest of our result on the ZRP is to demonstrate how the decoupling trick of Eq. (15) used for the RP can easily be applied also here. The model differences summarized in Table I do not alter the general theme. The same is true for the last studied model.

IV. TRUNCATED INVERSE DISTANCE SQUARED ISING MODEL

A. Basics

The TIDSI describes a one-dimensional system of spin domains with each domain having spins $+1$ or -1 ; see Fig. 1. There is an inverse squared long-range interaction between spins within the same domain. Let N be the random number of domains $i \in [1, N]$ with each of domain length $\lambda_i \geq 1$. The constraint is the fixed total length of the system

$$L = \sum_{i=1}^N \lambda_i; \quad (52)$$

see Fig. 1 and Table I. The domain i of length λ_i is associated with the weight $\psi(\lambda) \propto \lambda^{-\gamma}$, where the domain length decays with the parameter $\gamma \geq 1$ which is the product of the inverse temperature $1/(k_B T)$ and the long-range interaction [31]. The relevance of TIDSI is that it exhibits a mixed order phase transition, i.e., it shows features of phase transitions of first and of second kind. Depending on the temperature, there is either a ferromagnetic phase with a large number of domains or a paramagnetic phase with one domain of order L . Thus the analysis of the extreme domain size $\lambda_{\max} = \max(\lambda_1, \dots, \lambda_N)$ is important [5, 31].

B. Extreme value statistics in the second half

We investigate the statistics of the maximum domain length [31]

$$\lambda_{\max} = \max(\lambda_1, \lambda_2, \dots, \lambda_N). \quad (53)$$

Since the number of events N is random it is instructive to consider

$$f(m; L) = \sum_{N=1}^{\infty} f_N(m; L) \quad (54)$$

with $f_N(m; L) = F_N(m; L) - F_N(m - 1; L)$ being the maximum PDF with exactly N renewal events. In this context the value of N is a sampled value. The maximum CDF with given N were derived in [31] and is given by

$$F_N(m; L) = \frac{1}{Z(L)} \sum_{\lambda=0}^m \prod_{i=1}^N \psi(\lambda_i) \delta_{L, \|\lambda\|_1}. \quad (55)$$

This formula is almost identical to Eq. (44) for ZRP but the partition function is here

$$Z(L) = \sum_{N=1}^{\infty} (\psi * \cdots * \psi)^{(N)}(L). \quad (56)$$

The maximum PMF with given N is

$$f_N(m; L) = \frac{N \psi(m)}{Z(L)} \sum_{\lambda=0}^m \prod_{i=1}^{N-1} \psi(\lambda_i) \delta_{L-m, \|\lambda\|_1}. \quad (57)$$

Again, we use that in the second half $L/2 < m < L$ this formula is identical to the convolution. The second half maximum PMF with given N is

$$f_N(m; L) = N \psi(m) P_{N-1}(L - m). \quad (58)$$

Here, the probability of having $N - 1$ spin domains is

$$\begin{aligned} P_{N-1}(L - m) &= \frac{1}{Z(L)} \langle \delta_{L-m, \|\lambda\|_1} \rangle \\ &= \frac{1}{Z(L)} (\psi * \cdots * \psi)^{(N-1)}(L - m) \end{aligned} \quad (59)$$

where the average $\langle \circ \rangle$ is performed over all possible domain lengths. Finally, averaging over all N yields the second half maximum PMF

$$f(m; L) = \psi(m) \frac{Z(L - m)}{Z(L)} [\langle N(L - m) \rangle + 1] \quad (60)$$

with $L/2 < m < L$. The mean number of domains is

$$\langle N(L - m) \rangle = \sum_{N=1}^{\infty} N P_N(L - m). \quad (61)$$

One could write Eq. (60) also as

$$f_N(m; L) = \frac{1}{Z(L)} \psi(m) \sum_{N=1}^{\infty} N \Phi_{N-1}(L - m). \quad (62)$$

in order to emphasize the relationship to the random walk picture, similarly to what we did for the ZRP. When the maximum spin domain length λ_{\max} is m , all other lengths add

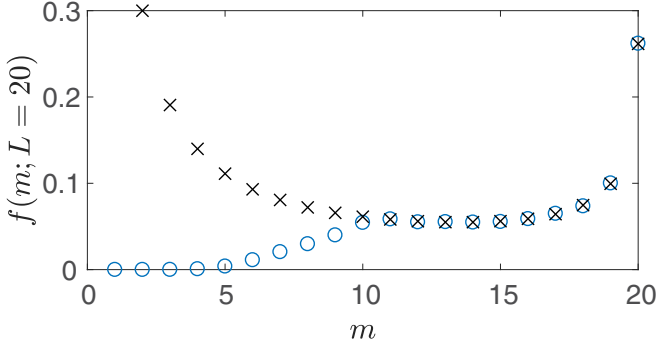


FIG. 8. Histogram of the maximum PMF $f(m; L)$ of TIDSI from Monte Carlo simulations (blue circles) compared with the theory of Eq. (60) (black line) for zeta $\psi(\lambda) = 1/\zeta(1 + \alpha)\lambda^{-1-\alpha}$ with $\alpha = 0.2$ and $K = 20$. The simulations were performed with 10^6 realizations. The analytical expression of the mean number of domains $\langle N(L - m) \rangle$ [as well as $Z(L - m)$ and $Z(L)$] is obtained via inverse z transform. The kink at the midpoint $L/2$ is visible although the random variables are discrete.

up to the remaining length $L - m$ due to the constraint. In Fig. 8 we compare theory and simulation.

C. Limiting law in the critical phase

We consider the large total length limit $L \rightarrow \infty$ of $f(m; L)$ in the second half $L/2 < m < L$, i.e., Eq. (60) or (62), for the critical phase between ferromagnetic and paramagnetic phases [31]. Then the calculations can be transferred almost effortlessly from the above RP techniques. For the limiting laws in the ferromagnetic and paramagnetic phases we refer to [31]. The weight is generally

$$\psi(\lambda) = \frac{e^{-\beta\Delta}}{\lambda^{1+\alpha}} \quad (63)$$

with the inverse temperature $\beta = 1/(k_B T)$, the chemical potential Δ , and $1 + \alpha = \beta J \geq 1$, where J is the strength of the inverse squared long-range interaction within a single spin domain; see [31]. In the critical phase the marginal domain size decays algebraically. Then the weight is

$$\psi(\lambda) = \frac{1}{\zeta(1 + \alpha)\lambda^{1+\alpha}} \quad (64)$$

with the Riemann zeta function $\zeta(1 + \alpha) = \sum_{N=1}^{\infty} N^{-1-\alpha}$, i.e.; the fugacity is $e^{-\beta\Delta} = 1/\zeta(1 + \alpha)$. It was shown in [31] that there are two regimes in the critical phase for $\alpha \in (0, 1)$ and $\alpha > 1$. We restrict the latter to $\alpha \in (1, 2)$ in order to compare it to RP. As explained in [31] the analysis using z transform can be replaced by Laplace transforms in the critical phase, which we use now.

The z transform of the weight is

$$\psi(\lambda) \circ \bullet \hat{\psi}(z) = \sum_{L=1}^{\infty} \psi(L) z^L, \quad (65)$$

the z -transform of the denominator of Eq. (60) is

$$\sum_{N=1}^{\infty} N \Phi_{N-1}(L - m) \circ \bullet \sum_{N=1}^{\infty} N z^m \hat{\psi}^{N-1}(z) = \frac{z^m}{[1 - \hat{\psi}(z)]^2} \quad (66)$$

and the z transform of the numerator of Eq. (60) is

$$\sum_{N=1}^{\infty} \Phi_N(L) \circ \bullet = \sum_{N=1}^{\infty} \hat{\psi}^N(z) = \frac{\hat{\psi}(z)}{1 - \hat{\psi}(z)}. \quad (67)$$

The symbol $\circ \bullet$ means we perform the z transform as defined in Eq. (65).

We study here the scaling $m = O(L)$ and $L - m = O(L)$. Hence, we need the large L limit of both the denominator and numerator. We set $z = \exp(-s)$ and consider the small s behavior of the weights,

$$\hat{\psi}(s) \sim \begin{cases} 1 - \frac{|\Gamma(-\alpha)|}{\zeta(1+\alpha)} s^\alpha & \text{for } 0 < \alpha < 1, \\ 1 - \langle \lambda \rangle s & \text{for } 1 < \alpha < 2. \end{cases} \quad (68)$$

This is equivalent to the asymptotic behavior of $\hat{\psi}(z) \sim 1 - |\Gamma(-\alpha)|/\zeta(1 + \alpha)(1 - z)^\alpha - \zeta(\alpha)/\zeta(1 + \alpha)(1 - z)$ at the branch point $z = 1$; see [31].

For $\alpha \in (0, 1)$ we get from the inverse Laplace transform the scaling law

$$f(m; L) \sim \frac{1}{L} \mathcal{G}\left(\frac{m}{L}\right) \quad (69)$$

with

$$\mathcal{G}(\xi) = \frac{\Gamma(\alpha)}{|\Gamma(-\alpha)|\Gamma(2\alpha)} \xi^{-1-\alpha} (1 - \xi)^{2\alpha-1} \quad (70)$$

with the rescaled variable $\xi = m/L$. The same limiting law has been derived in [32]. It has also been derived in [31] but with a different expression depending on hypergeometric functions. The results are identical; see Appendix B. Equation (69) is valid for $1/2 < \xi < 1$ due to the restriction on the second half. The midpoint $\xi = 1/2$ is nonanalytical as reported in [31,32]. Note that at $\xi \rightarrow 1$ the function blows up to infinity. In reality, for any finite observation time, the maximum PMF does not diverge. Below we cure this problem again by considering constant remaining length $L - m = O(1)$. This describes the rare events where the scaling law (69) is not valid anymore.

For $\alpha \in (1, 2)$ we get

$$f(m; L) \sim \frac{1}{L^\alpha} \mathcal{I}\left(\frac{m}{L}\right) \quad (71)$$

with

$$\mathcal{I}(\xi) = \frac{1}{\zeta(1 + \alpha)\langle \lambda \rangle} \xi^{-1-\alpha} (1 - \xi). \quad (72)$$

The rescaled variable is $\xi = m/L$. The important point of this scaling law is that it cures the unphysical diverging second moment of Fréchet's law describing typical events,

$$f(m; L) \sim \frac{1}{(L/\langle \lambda \rangle)^{1/\alpha}} b_\alpha \xi^{-1-\alpha} \exp\left(-\frac{b_\alpha \xi^{-\alpha}}{\alpha}\right) \quad (73)$$

with $b_\alpha = 1/\zeta(1 + \alpha)$ and the rescaled variable $\xi = m(L/\langle \lambda \rangle)^{-1/\alpha}$, i.e., $m = O(L^{1/\alpha})$. The mean length is $\langle \lambda \rangle = \zeta(\alpha)/\zeta(1 + \alpha)$. See [31] for a rigorous derivation. But Fréchet's law predicts the divergence of the variance of λ_{\max} which is unphysical since $m \leq T$. The scaling law Eq. (71) matches with Fréchet's law: The small m behavior of Eq. (71) equals the large m behavior of Fréchet's law, namely $b_\alpha L m^{-1-\alpha}/\langle \lambda \rangle$. Hence the two scaling regimes are

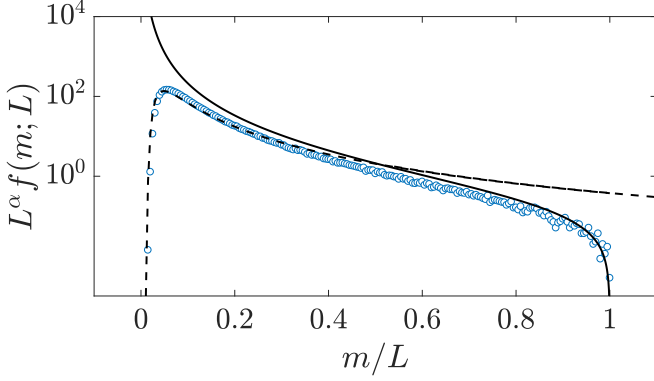


FIG. 9. Rescaled histogram of the maximum PMF $f(m; L)$ for TIDS I (blue circles) for $L = 200$ compared with the limiting law of Eq. (71) (solid line) and Fréchet's law of Eq. (73) (dashed line). The simulations were performed with 10^6 realizations and $\alpha = 3/2$. Clearly, Eq. (71) works relatively well already for not too large L provided that $m > L/2$.

complementary. In Fig. 9 we compare Eq. (71) with numerical simulations and Fréchet's law (73).

The function of Eq. (36) is non-normalizable:

$$L^\alpha \int_0^L f(m; L) dm \sim \frac{1}{\zeta(1+\alpha)\langle\lambda\rangle} \int_0^1 \xi^{-1-\alpha} (1-\xi) d\xi \rightarrow \infty. \quad (74)$$

Similar to the RP, this limiting function describing rare events cures the infinite variance problem of Fréchet's law.

D. Limiting law in the critical phase with constant remaining length

Here, we calculate the long-time limit of $f(m; L)$ in the second half time $L/2 < m < L$, i.e., Eq. (60), also with $m = O(L)$, but now we consider constant remaining total length $L - m = O(1)$. Hence the $Z(L - m)$ and the mean number of domains $\langle N(L - m) \rangle$ stay constant in Eq. (60). So we only have to consider $\psi(m) \sim \psi(L)$ and the large L behavior of $Z(L)$. Therefore we find the scaling law

$$\begin{aligned} f(m; L) &\sim \frac{\psi(L)}{Z(L)} Z(L - m) [\langle N(L - m) \rangle + 1] \\ &= Z(L - m) [\langle N(L - m) \rangle + 1] \\ &\times \begin{cases} \frac{|\Gamma(-\alpha)\Gamma(\alpha)|}{\zeta^2(1+\alpha)} L^{-2\alpha} & \text{for } \alpha \in (0, 1), \\ \frac{\langle\lambda\rangle}{\zeta(1+\alpha)} L^{-1-\alpha} & \text{for } \alpha \in (1, 2). \end{cases} \quad (75) \end{aligned}$$

In particular, for $Z(L)$ we used the small z behavior of Eq. (67) and calculated the inverse Laplace transform with $z = \exp(-s)$.

The meaning of this scaling law is similar to the RP limiting law in Sec. IIH: it describes the rare events of m very close to the constraint L . In Fig. 10(a) we show Monte Carlo simulations for a system size $L = 200$. The figure illustrates that the exact expression for $f(m; L = 200)$ in the second half equation (60) works well as expected. The region near L is well described by the asymptotic theory, Eq. (75), while the law, Eq. (69), is not performing well. The latter observation is to be expected as we are dealing with rare events. Then in Fig. 10(b) we consider a larger system, $L = 5000$. Here, Monte Carlo simulations do not converge in a reasonable time. We can, however, explore this regime with our exact solution, Eq. (60), again a solution valid in the domain $m > L/2$. This points out to the fact that the exact solution can be exploited to

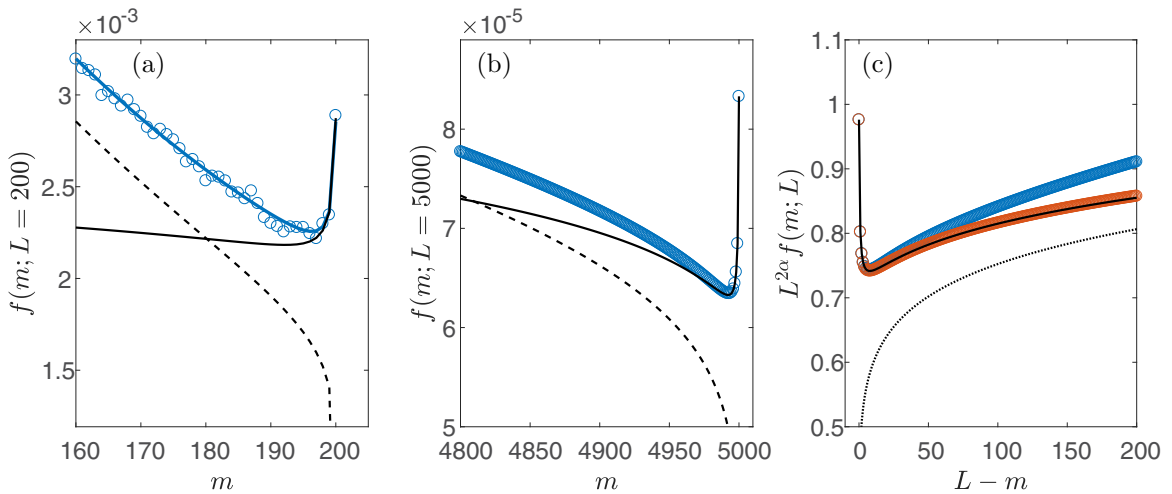


FIG. 10. (a) Comparison of the maximum PMF $f(m; L)$ for TIDS I with $L = 200$ and $\alpha = 0.55$ [see Eq. (64)] of Monte Carlo simulations (red circles), the exact half time distribution (60) (red solid line), the $m/L \rightarrow \text{const.}$ scaling law (69) for typical fluctuations (dashed line), and the $L - m = O(1)$ scaling law of Eq. (75) describing rare events (black solid line). The simulation is performed for 10^8 realizations. (b) The maximum PMF with $L = 5000$ of the exact half time distribution (60) (blue circles), the $m/L \rightarrow \text{const.}$ scaling law (69) (dashed line), and the $L - m = O(1)$ scaling law of Eq. (75) (solid line). (c) Rescaled maximum PMF plotted over $L - m$ of the exact maximum distribution (60) with $L = 5000$ and $L = 10^5$ compared with the $L - m = O(1)$ scaling law (75) and the matching function with the $L - m = O(L)$ scaling law (76).

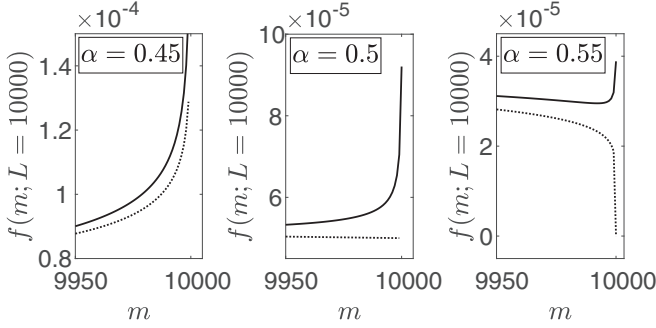


FIG. 11. Right tails of $f(m; L)$ for TIDSI with $L = 10^4$ described by the $L - m = O(1)$ limiting law of Eq. (75) (solid line) and the matching function to the $L - m = O(L)$ limiting law (76) (dotted line). Three different values of $\alpha \in (0, 1)$ have been chosen. There is a significant change of the behavior of the matching function when α passes $1/2$ while the exact behavior always increases for $m \rightarrow L$.

investigate rare fluctuations where sampling of rare events, at least with straightforward simulations, is difficult or impossible. Further, the exact theory also matches the asymptotic theory where it should, namely on the far right-hand side of the figure. We are able to plot the exact behavior of $f(m; L)$ near L for any large value of L . The detailed procedure is explained as follows. First, we replace $Z(L)$ by its large L behavior $\zeta(1 + \alpha)/[|\Gamma(-\alpha)|\Gamma(\alpha)]L^{-1+\alpha}$ in Eq. (60). Second, the denominator $Z(L - m)[\langle N(L - m) \rangle + 1]$ is exactly obtained via Taylor series of its z transform. And here it is important that the expression $Z(L - m)[\langle N(L - m) \rangle + 1]$ only depends on $L - m$. Since we are only interested in small $L - m \leq 200$ we are able to derive the Taylor series for any value of L with *Mathematica*. Thus, we obtain the exact expression of $f(m; L)$ near L . Finally, we compare this replacement of the data with the scaling laws in Fig. 10.

The matching between the two scaling laws with $L - m = O(L)$ of Eq. (69) and $L - m = O(1)$ of Eq. (75) can be analytically calculated with an argumentation identical to the previous RP comparison between the two regimes with $T - m = O(T)$ and $T - m = O(1)$ in Sec. IIH. The small $L - m$ limit of $f(m; L)$ with $L - m = O(L)$ is equal to the large $L - m$ limit of $f(m; L)$ with $T - m = O(1)$. For $\alpha \in (0, 1)$ this is

$$f(m; L) \sim \frac{\Gamma(\alpha)}{|\Gamma(-\alpha)|\Gamma(2\alpha)} L^{-2\alpha} (L - m)^{2\alpha - 1}; \quad (76)$$

see Fig. 10. For $\alpha \in (1, 2)$ the matching function is

$$f(m; L) \sim \frac{1}{\zeta(1 + \alpha)\langle \lambda \rangle} L^{-1-\alpha} (L - m). \quad (77)$$

An interesting observation is that this matching function for different values of $\alpha \in (0, 1)$ behaves totally differently than the exact solution when $m \rightarrow L$. In Fig. 11, we compare for $\alpha = 0.45, 0.5$, and 0.55 the matching function (76) with Eq. (75). Although both solutions match for small m , the $L - m = O(1)$ law (75) diverges at $m \rightarrow L$ while the matching solution (76) [and therefore also the $L - m = O(L)$ scaling law] changes its behavior at $\alpha = 1/2$. This behavior shows

that the rare event behavior is correctly described by assuming $L - m = O(1)$.

V. SUMMARY

We have analyzed EVT of the longest waiting time τ_{\max} of the RP, the largest particle number per site κ_{\max} of the ZRP, and the largest spin domain size λ_{\max} of the TIDSI. These three models share the global constraint for the sum of the random variables, i.e., the waiting times, the particle numbers per site, and the spin domain lengths. The exact details of the models differ from each other. While the number of sites in the ZRP is fixed, the number of waiting times (spin domains) is random in the RP (TIDSI). Furthermore, the last waiting time for the RP is cut off to the backward recurrence time. However, we found that despite these differences the common trait of the global constraint enabled us to decouple the problem when the extreme value is larger than half of the constraint. One of our main results is the revelation of the deep connection between two different fields: constrained EVT and well-known quantifiers of stochastic dynamics. The latter are the mean number of renewal events, Eq. (10), the sum of independent and identically distributed random variables, Eq. (51), and the mean number of spin domains, Eq. (60). Our results are in perfect accordance in the second half of the support as presented in Fig. 2 for the RP, in Fig. 7 for the ZRP, and in Fig. 8 for the TIDSI. In these figures the practical calculation of the theory relies on Laplace transforms (or z transforms). Since our theory relates two fields, namely EVT and underlying stochastic dynamics, we demonstrated exemplarily for RP in Fig. 3 that one can also obtain the EVT indirectly: The estimation of the mean number of renewals $\langle N \rangle$ is sufficient to obtain the maximum CDF by using Eq. (11). Another advantage of our theory is that we can plot the extreme value statistics for cases when Monte Carlo sampling demands huge computational resources. This was demonstrated for the TIDSI in Fig. 10.

After this general result of the second half maximum distribution, we considered different asymptotic limits for RP with power law waiting times with exponent $\alpha \in (0, 1)$ and $\alpha \in (1, 2)$. For TIDSI we have chosen to study the asymptotic limit in the critical phase between ferromagnetic and paramagnetic phases because the behavior is comparable to the RP behavior. We recapped known results and also found new limiting laws when the global constraint diverges $C \rightarrow \infty$, i.e., we have diverging observation time $C = T$ for the RP and diverging total domain length $C = L$ for the TIDSI. The limiting behavior of the maximum distribution $f(m; C)$ is summarized as

(a) When $C \rightarrow \infty$ and m/C is fixed, the second half maximum distribution for $\alpha \in (0, 1)$ describes typical events. For the RP we found Eq. (34) and explained previous results [27,45] by identifying contributions from \mathcal{B} and $\mathcal{N}\mathcal{B}$, i.e., both processes with the maximum being the last waiting time or not. For the TIDSI we found Eq. (69), which was derived in [31,32].

(b) When $C \rightarrow \infty$ and m/C is fixed, the second half maximum distribution for $\alpha \in (1, 2)$ complements the typical events described by Fréchet's law. For the RP we found

TABLE II. Collection of limiting laws of $f(m; C)$ for RP with $C = T$ and TIDSI with $C = L$ in the critical phase. The random variables (waiting times and spin domains lengths) are fat-tail distributed with exponent α . For the study on the first half $m \in (0, C/2)$ we refer for RP to [27] and for TIDSI to [31], in particular for $\alpha \in (0, 1)$ the first half shows a different scaling law than Fréchet’s law. In the second half, i.e., $C/2 < m < C$, the scaling $m = O(C)$ is applied. The rescaled variable is $\xi = m/C$. We find again the laws of (a) [27,45], (b) [28], and (c) [32]. Another expression of (c) has been derived in [31] using other methods; they express the law as a sum of two hypergeometric functions while the expression in the table is simpler; see Appendix B. Note that we present the expressions in this table without prefactors.

		Second half: $m \in (C/2, C)$		
Model	$\alpha \in$	First half: $m \in (0, C/2)$	Remaining constraint $C - m = O(C)$	Remaining constraint $C - m = O(1)$
RP	(0,1)	Beyond Fréchet’s law	(a) $T f(m; T) \sim \xi^{-1-\alpha}(1 - \xi)^{\alpha-1}$	(e) $T^\alpha f(m; T) \sim R(T - m)$
	(1,2)	Fréchet’s law	(b) $T^\alpha f(m; T) \sim \xi^{-1-\alpha}[1 - (1/\alpha - 1)\xi]$	(f) $T^\alpha f(m; T) \sim R(T - m)$
TIDSI	(0,1)	Beyond Fréchet’s law	(c) $L f(m; L) \sim \xi^{-1-\alpha}(1 - \xi)^{2\alpha-1}$	(g) $L^{2\alpha} f(m; L) \sim Z(L - m)[(N(L - m) + 1)]$
	(1,2)	Fréchet’s law	(d) $L^\alpha f(m; L) \sim \xi^{-1-\alpha}(1 - \xi)$	(h) $L^{1+\alpha} f(m; L) \sim Z(L - m)[(N(L - m) + 1)]$

Eq. (36) which was derived in [28]. For the TIDSI we found Eq. (71). Both limiting laws are infinite densities.

(c) We find the rare events of the statistics of the maximum for $\alpha \in (0, 1)$. Especially for the RP this is relevant because it cures the divergent behavior of the typical events near the observation time. The scaling of the rare events assumes $T - m$ is fixed while $T \rightarrow \infty$. Equation (38) shows that only the process \mathcal{B} is important. Here the rate function, of the mean number of renewals, is a useful tool in the analysis of the large deviations. Of course while this rate function describes rare events, it is very different from the rate function of standard large deviation theory [48]. Finally, the presented results are used in [49] where we established the so-called big jump principle [28] for the ballistic Lévy walk model. In summary, there we show the usefulness of the approach, in the sense that the statistics of τ_{\max} might be used to predict the large deviations of a widely applicable model of anomalous transport. In addition, we found the same scaling behavior to describe the rare event near L for the TIDSI in the critical phase in Eq. (75). However, there is obviously no distinction between \mathcal{B} and $\mathcal{N}\mathcal{B}$.

We collect the just described limiting laws in Table II together with the behavior in the first half $0 < m < T/2$. While for classical EVT the limiting behavior is described by Fréchet’s law, the global constraint yields rich limiting behavior with different scaling laws for which our theory provides a helpful tool to derive them as presented in the main text.

Note added. Recently C. Godrèche published related results [50]. We also thank him for pointing out Refs. [32,33].

ACKNOWLEDGMENTS

M.H. is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), 436344834. E.B. acknowledges the Israel Science Foundations Grant No. 1898/17. W.W. was supported by Bar-Ilan University together with the Planning and Budgeting Committee fellowship program.

APPENDIX A: INTEGRALS IDENTICAL TO THE CONVOLUTION

We first consider the RP and later summarize the results also for the ZRP and the TIDSI. In Eq. (14) we have integrals

of the form

$$I_N(m, T') = \int_0^m d\tau_1 \cdots \int_0^m d\tau_N \prod_{i=1}^N g_i(\tau_i) \delta\left(T' - \sum_{j=1}^N \tau_j\right). \tag{A1}$$

Note that in Eq. (14) there are $(N - 1)$ -multiple integrals but we consider now N -multiple integrals. The functions $g_i(\tau_i)$ in Eq. (14) are the waiting time PDFs $\psi(\tau_i)$ or the survival probability $\varphi(\tau_i)$. Furthermore the parameter T' in Eq. (14) is the remaining time $T - m$. Here we discuss general functions which must be positive $g_i(\tau_i) \geq 0$ with positive arguments $\tau_i \geq 0$. And we consider an arbitrary constraint $T' > 0$. The main result of this section is that the integral $I_N(m, T')$ is identical to the convolution

$$I_N(m, T') = (g_1 * \cdots * g_N)^{(N)}(T') \tag{A2}$$

when the condition $m > T'$ is fulfilled. This condition will lead to the range of the second half $T/2 < m < T$ when $T' = T - m$. The twofold convolution is $(g_1 * g_2)^{(2)}(T') = \int_0^{T'} d\tau_1 g_1(\tau_1) g_2(T' - \tau_1)$ and higher orders are defined successively.

We derive Eq. (A2) with a proof by induction. Let us start with $N = 2$, i.e., we show now that

$$I_2(m, T') = (g_1 * g_2)^{(2)}(T') \tag{A3}$$

when $m > T'$. By definition we have

$$I_2(m, T') = \int_0^m d\tau_1 \int_0^m d\tau_2 g_1(\tau_1) g_2(\tau_2) \delta(T' - \tau_1 - \tau_2). \tag{A4}$$

For the inner integral we take both limits to infinity while putting two Heaviside functions into the integrand

$$\begin{aligned} & \int_0^m d\tau_2 g_2(\tau_2) \delta(T' - \tau_1 - \tau_2) \\ &= \int_{-\infty}^{+\infty} d\tau_2 g_2(\tau_2) \Theta(\tau_2) \Theta(m - \tau_2) \delta(T' - \tau_1 - \tau_2) \\ &= g_2(T' - \tau_1) \Theta(T' - \tau_1) \Theta(m - [T' - \tau_1]). \end{aligned} \tag{A5}$$

Hence this inner integral is only nonzero under the condition

$$T' - m < \tau_1 < T'. \tag{A6}$$

The further analysis of the outer integral of Eq. (A4) depends on this condition (A6) and the relationship between T' and m .

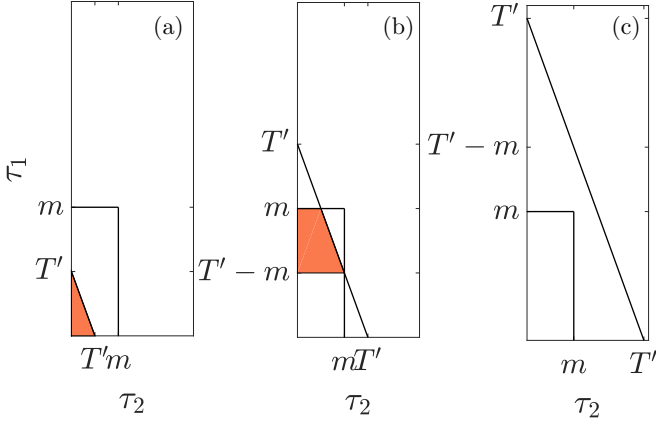


FIG. 12. Areas of integration of $I_2(m; T')$ for three different regimes depending on the relationship between the maximum m to some parameter T' ; see Eq. (A7). The most relevant integration is (a). Our claim is that in this case we may restrict the integration in Eq. (A1) to $m = T'$, since the constraint limits the relevant domain of the integration variables.

We may consider the three regimes

$$(a) \ 0 < T' < m,$$

$$I_{N+1}(m, T') = \int_0^m d\tau_{N+1} g_{N+1}(\tau_{N+1}) \left[\int_0^m d\tau_1 \cdots \int_0^m d\tau_N \prod_{i=1}^N g_i(\tau_i) \delta\left(T' - \tau_{N+1} - \sum_{j=1}^N \tau_j\right) \right]. \quad (A10)$$

Now we consider $T' < m$, which corresponds to regime (a) from Eq. (A7). From this obviously $T' - \tau_{N+1} < m$ because $T' - \tau_{N+1} < T'$. This inequality $T' - \tau_{N+1} < m$ is exactly the condition for which the N -multiple integral inside the square brackets of Eq. (A10) is the N -fold convolution

$$I_{N+1}(m, T') = \int_0^m d\tau_{N+1} g_{N+1}(\tau_{N+1}) \times [(g_1 * \cdots * g_N)^{(N)}(T' - \tau_{N+1})] \quad (A11)$$

according to the assumption of the induction proof. The remaining integral over τ_{N+1} is zero from T' to m . The difference $T' - \tau_{N+1} = \sum_{i=1}^N \tau_i$ is positive because all τ_i are positive. So when $\tau_{N+1} > T'$ we cannot fulfill the constrain. This property is controlled by the convolution in the integrand of Eq. (A11) which is zero for negative arguments. So we get

$$I_{N+1}(m, T') = \int_0^{T'} g_{N+1}(\tau_{N+1}) \times d\tau_{N+1} [(g_1 * \cdots * g_N)^{(N)}(T' - \tau_{N+1})] \quad (A12)$$

and this is the convolution. Remember that we assumed $T' < m$ in Eq. (A11). Therefore we showed Eq. (A2).

With the same arguments, Eq. (A2) can also be stated for discrete random variables with some arbitrary constraint

$$(b) \ m < T' < 2m, \\ (c) \ 2m < T'. \quad (A7)$$

Both conditions of Eq. (A6) and Eq. (A7) lead to

$$I_2(m, T') = \begin{cases} \int_0^{T'} d\tau_1 g_1(\tau_1) g_2(T' - \tau_1) & \text{for (a),} \\ \int_{T'-m}^m d\tau_1 g_1(\tau_1) g_2(T' - \tau_1) & \text{for (b),} \\ 0 & \text{for (c).} \end{cases} \quad (A8)$$

See also Fig. 12 for three different areas of integration. We are only interested in the first regime when $0 < T' < m$. Then the double integral is the convolution and hence Eq. (A3) is shown for $N = 2$.

In order to finish the proof of Eq. (A2) we show it for $N + 1$ while assuming that the statement is true for N . We write again the definition of the integral

$$I_{N+1}(m, T') = \int_0^m d\tau_1 \cdots \int_0^m d\tau_{N+1} \prod_{i=1}^{N+1} g_i(\tau_i) \delta\left(T' - \sum_{j=1}^{N+1} \tau_j\right). \quad (A9)$$

We rearrange the order of integration and separate $-\tau_N + 1$ in the delta function:

$C' > 0$. It is equivalently

$$\sum_{y_1=0}^m \cdots \sum_{y_N=0}^m \prod_{i=1}^N g_i(y_i) \delta_{C', \sum_{j=1}^N y_j} = (g_1 * \cdots * g_N)^{(N)}(C') \quad (A13)$$

for $m > C'$. For ZRP it is $y_i = \kappa_i$ and $C' = K - m$ and for TIDSI it is $y_i = \lambda_i$ and $C' = L - m$. For ZRP and TIDSI the functions are $g_i = \psi$ for all i .

APPENDIX B: TYPICAL FLUCTUATIONS OF TIDSI FOR THE PARAMETER $\alpha \in (0, 1)$

In [31] the typical fluctuations of $f(m; L)$ in the second half $L/2 < m < L$ were calculated as

$$Lf(m; L) \sim \frac{1}{\xi^2} \frac{d}{du} H(u)|_{u=1/\xi} \quad (B1)$$

with the function

$$H(u) = \frac{\Gamma(\alpha)}{\Gamma(2\alpha + 1)|\Gamma(-\alpha)|} u^{1-\alpha} (u-1)^{2\alpha} \times {}_2F_1(1, 1 + \alpha, 1 + 2\alpha, 1 - u). \quad (B2)$$

The hypergeometric function defined as

$${}_2F_1(a, b, c, z) = \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \frac{z^j}{j!} \tag{B3}$$

with the Pochhammer symbol $(a)_j = \Gamma(a + j)/\Gamma(a)$.

We show now that Eq. (B1) is identical to our result from Eq. (69). For that let us first take the derivative of the right-hand side of Eq. (B1) while $u = 1/\xi$:

$$\begin{aligned} u^2 \frac{d}{du} H(u) &= u^2 \frac{\Gamma(\alpha)}{\Gamma(2\alpha + 1) |\Gamma(-\alpha)|} \\ &\times \left[[(1 - \alpha)u^{-\alpha}(u - 1)^{2\alpha} + 2\alpha u^{1-\alpha}(u - 1)^{2\alpha-1}] \right. \\ &\times {}_2F_1(1, 1 + \alpha, 2\alpha + 1, 1 - u) \\ &- \frac{1 + \alpha}{1 + 2\alpha} u^{1-\alpha}(u - 1)^{2\alpha} \\ &\left. \times {}_2F_1(2, 2 + \alpha, 2\alpha + 2, 1 - u) \right], \tag{B4} \end{aligned}$$

where we used $d/dz {}_2F_1(a, b, c, z) = ab/c {}_2F_1(1 + a, 1 + b, 1 + c, z)$. Now we take out the term $u^{-\alpha}(u - 1)^{2\alpha-1}$ so that

$$\begin{aligned} u^2 \frac{d}{du} H(u) &= \frac{\Gamma(\alpha)}{\Gamma(2\alpha + 1) |\Gamma(-\alpha)|} u^{2-\alpha}(u - 1)^{2\alpha-1} \\ &\times \left[[(1 - \alpha)(u - 1) + 2\alpha u] \right. \\ &\times {}_2F_1(1, 1 + \alpha, 2\alpha + 1, 1 - u) \\ &- \frac{1 + \alpha}{1 + 2\alpha} u(u - 1) \\ &\left. \times {}_2F_1(2, 2 + \alpha, 2\alpha + 2, 1 - u) \right]. \tag{B5} \end{aligned}$$

To show the identity to Eq. (69) we have to show that the expression inside the big square brackets of Eq. (B5) is identical to 2α . Let us write this question in shorter form as

$$f(u)F(1, 1 - u) + g(u)F(2, 1 - u) = 2\alpha; \tag{B6}$$

i.e., is this statement true? Here $f(u) = (1 - \alpha)(u - 1) + 2\alpha u$, $g(u) = -(1 + \alpha)/(1 + 2\alpha)u(u - 1)$, and $F(i, 1 - u) = {}_2F_1(i + 1, i + 1 + \alpha, i + 1 + 2\alpha, 1 - u)$.

Since the hypergeometric function depends on $1 - u$ we consider the series expansion at $u = 1$ of the inner brackets. In principle any other point could be considered but the problem becomes simpler at $u = 1$. The Taylor series of Eq. (B6) is

$$\begin{aligned} f(u)F(1, 1 - u) + g(u)F(2, 1 - u) &= \sum_{j=0}^{\infty} (f(u)F(1, 1 - u) + g(u)F(2, 1 - u))^{(j)}|_{u=1} \\ &\times \frac{(u - 1)^j}{j!}. \tag{B7} \end{aligned}$$

We apply the general Leibniz rule of derivation,

$$\begin{aligned} f(u)F(1, 1 - u) + g(u)F(2, 1 - u) &= \sum_{j=0}^{\infty} \left(\sum_{k_1=0}^j \binom{j}{k_1} F^{(j-k_1)}(1, 1 - u) f^{(k_1)}(u)|_{u=1} \right. \\ &\left. + \sum_{k_2=0}^j \binom{j}{k_2} F^{(j-k_2)}(2, 1 - u) g^{(k_2)}(u)|_{u=1} \right) \frac{(u - 1)^j}{j!}. \tag{B8} \end{aligned}$$

The derivatives of f and g are

$$\begin{aligned} f^{(k_1)}(u)|_{u=1} &= \begin{cases} 2\alpha & \text{for } k_1 = 0, \\ 1 + \alpha & \text{for } k_1 = 1, \\ 0 & \text{for } k_1 \geq 2, \end{cases} \\ g^{(k_2)}(u)|_{u=1} &= \begin{cases} 0 & \text{for } k_2 = 0, \\ -\frac{1+\alpha}{1+2\alpha} & \text{for } k_2 = 1, \\ -2\frac{1+\alpha}{1+2\alpha} & \text{for } k_2 = 2, \\ 0 & \text{for } k_2 \geq 3. \end{cases} \tag{B9} \end{aligned}$$

The two sums in Eq. (B8) are only nonzero for $k_1 = 0, 1$ and $k_2 = 2, 3$. Thus we can write

$$\begin{aligned} f(u)F(1, 1 - u) + g(u)F(2, 1 - u) &= \sum_{j=0}^{\infty} \left(\sum_{k_1=0}^1 \binom{j}{k_1} F^{(j-k_1)}(1, 1 - u) f^{(k_1)}(u)|_{u=1} \right. \\ &\left. + \sum_{k_2=2}^3 \binom{j}{k_2} F^{(j-k_2)}(2, 1 - u) g^{(k_2)}(u)|_{u=1} \right) \frac{(u - 1)^j}{j!}. \tag{B10} \end{aligned}$$

The binomial is zero when $k_1 > j$ and $k_2 > j$ so this expression is valid for all j . Now we express the hypergeometric function $F(2, 1 - u)$ by $F(1, 1 - u)$ via the relationship of their derivatives. The j th derivative of the hypergeometric function at $u = 1$ is

$$F^{(j)}(1, 1 - u)|_{u=1} = (-1)^j \frac{(1)_j (1 + \alpha)_j}{(1 + 2\alpha)_j}, \tag{B11}$$

thus

$$F^{(j)}(2, 1 - u)|_{u=1} = -\frac{1 + 2\alpha}{1 + \alpha} F^{(j+1)}(1, 1 - u). \tag{B12}$$

So we can write

$$\begin{aligned} f(u)F(1, 1 - u) + g(u)F(2, 1 - u) &= \sum_{j=0}^{\infty} \left(\sum_{k_1=0}^1 \binom{j}{k_1} F^{(j-k_1)}(1, 1 - u) f^{(k_1)}(u)|_{u=1} \right. \\ &- \frac{1 + 2\alpha}{1 + \alpha} \\ &\left. \times \sum_{k_2=1}^2 \binom{j}{k_2} F^{(j-k_2+1)}(1, 1 - u) g^{(k_2)}(u)|_{u=1} \right) \frac{(u - 1)^j}{j!}. \tag{B13} \end{aligned}$$

We order according to the hypergeometric functions:

$$\begin{aligned}
 & f(u)F(1, 1 - u) + g(u)F(2, 1 - u) \\
 &= \sum_{j=0}^{\infty} \left(F^{(j)}(1, 1 - u) \right. \\
 &\quad \times \left[\binom{j}{0} f^{(0)}(u) - \frac{1 + 2\alpha}{1 + \alpha} \binom{j}{1} g^{(1)}(u) \right] \Big|_{u=1} \\
 &\quad + F^{(j-1)}(1, 1 - u) \\
 &\quad \times \left. \left[\binom{j}{1} f^{(1)}(u) - \frac{1 + 2\alpha}{1 + \alpha} \binom{j}{2} g^{(2)}(u) \right] \Big|_{u=1} \right) \frac{(u - 1)^j}{j!}.
 \end{aligned} \tag{B14}$$

With Eq. (B9) we get

$$f(u)F(1, 1 - u) + g(u)F(2, 1 - u)$$

$$\begin{aligned}
 &= \sum_{j=0}^{\infty} ((2\alpha + j)F^{(j)}(1, 1 - u)|_{u=1} \\
 &\quad + j(\alpha + j)F^{(j-1)}(1, 1 - u)|_{u=1}) \frac{(u - 1)^j}{j!}.
 \end{aligned} \tag{B15}$$

Now we split the summation over j for $j = 0$ and all other $j \geq 1$. For the latter we use the relationship between successive orders of the derivative for the hypergeometric function,

$$F^{(j)}(1, 1 - u)|_{u=1} = -\frac{j(\alpha + j)}{2\alpha + j} F^{(j-1)}(1, 1 - u)|_{u=1}, \tag{B16}$$

valid for $j \geq 1$. This gives zero for all terms with $j \geq 1$ in Eq. (B15) and only the term with $j = 0$ remains. With $F^{(0)}(1, 1 - u)|_{u=1} = 1$ we obtain

$$f(u)F(1, 1 - u) + g(u)F(2, 1 - u) = 2\alpha. \tag{B17}$$

Thus we finally showed that indeed

$$u^2 \frac{d}{du} H(u) = \frac{\Gamma(\alpha)}{\Gamma(2\alpha)|\Gamma(-\alpha)|} u^{2-\alpha} (u - 1)^{2\alpha-1}. \tag{B18}$$

Hence Eq. (B1) is identical to our result from Eq. (69).

[1] J.-P. Bouchaud and M. Mézard, *J. Phys. A Math. Gen.* **30**, 7997 (1997).

[2] S. Albeverio, V. Jentsch, and H. Kantz, *Extreme Events in Nature and Society* (Springer-Verlag, Berlin, 2006).

[3] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events: For Insurance and Finance*, Applications of Mathematics Vol. 33 (Springer-Verlag, Berlin, 2013).

[4] J.-Y. Fortin and M. Clusel, *J. Phys. A Math. Theor.* **48**, 183001 (2015).

[5] S. N. Majumdar, A. Pal, and G. Schehr, *Phys. Rep.* **840**, 1 (2020).

[6] B. Meerson and S. Redner, *Phys. Rev. Lett.* **114**, 198101 (2015).

[7] Z. Schuss, K. Basnayake, and D. Holcman, *Phys. Life Rev.* **28**, 52 (2019).

[8] W. Wang, A. Vezzani, R. Burioni, and E. Barkai, *Phys. Rev. Research* **1**, 033172 (2019).

[9] D. Hartich and A. Godec, *J. Phys. A Math. Theor.* **52**, 244001 (2019).

[10] L. Zarfaty, E. Barkai, and D. A. Kessler, *arXiv:2006.13677*.

[11] E. J. Gumbel, *Statistics of Extremes* (Dover, New York, 1958).

[12] R. A. Fisher and L. H. C. Tippett, *Math. Proc. Cambridge Philos. Soc.* **24**, 180 (1928).

[13] C. Godrèche and J. Luck, *J. Stat. Phys.* **104**, 489 (2001).

[14] M. Niemann, E. Barkai, and H. Kantz, *Math. Model. Nat. Phenom.* **11**, 191 (2016).

[15] W. Wang, J. H. P. Schulz, W. Deng, and E. Barkai, *Phys. Rev. E* **98**, 042139 (2018).

[16] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 2 (Wiley, New York, 1971).

[17] S. B. Lowen and M. C. Teich, *Phys. Rev. E* **47**, 992 (1993).

[18] M. Evans, S. N. Majumdar, and R. Zia, *J. Stat. Phys.* **123**, 357 (2006).

[19] S. Majumdar, *Exact Methods in Low-dimensional Statistical Physics and Quantum Computing*, Lecture Notes of the Les Houches Summer School: Vol. 89, July 2008 (Oxford University Press, Oxford, 2010), p. 407.

[20] S. N. Majumdar, M. R. Evans, and R. K. P. Zia, *Phys. Rev. Lett.* **94**, 180601 (2005).

[21] R. Zia, M. Evans, and S. N. Majumdar, *J. Stat. Mech.* (2004) L10001.

[22] M. R. Evans and T. Hanney, *J. Phys. A: Math. Gen.* **38**, R195 (2005).

[23] A. Bar and D. Mukamel, *Phys. Rev. Lett.* **112**, 015701 (2014).

[24] A. Bar and D. Mukamel, *J. Stat. Mech.* (2014) P11001.

[25] C. Godrèche, S. N. Majumdar, and G. Schehr, *Phys. Rev. Lett.* **102**, 240602 (2009).

[26] C. Godrèche, S. N. Majumdar, and G. Schehr, *J. Phys. A: Math. Theor.* **47**, 255001 (2014).

[27] C. Godrèche, S. N. Majumdar, and G. Schehr, *J. Stat. Mech.* (2015) P03014.

[28] A. Vezzani, E. Barkai, and R. Burioni, *Phys. Rev. E* **100**, 012108 (2019).

[29] C. L. Scheffer, *Stoch. Process. Their Appl.* **55**, 101 (1995).

[30] M. R. Evans and S. N. Majumdar, *J. Stat. Mech.* (2008) P05004.

[31] A. Bar, S. N. Majumdar, G. Schehr, and D. Mukamel, *Phys. Rev. E* **93**, 052130 (2016).

[32] C. Godrèche, *J. Phys. A: Math. Theor.* **50**, 195003 (2017).

[33] J. Wendel, *Math. Scand.* **14**, 21 (1964).

[34] A. Rebenshtok, S. Denisov, P. Hänggi, and E. Barkai, *Phys. Rev. Lett.* **112**, 110601 (2014).

[35] T. Akimoto, E. Barkai, and G. Radons, *Phys. Rev. E* **101**, 052112 (2020).

[36] D. A. Kessler and E. Barkai, *Phys. Rev. Lett.* **105**, 120602 (2010).

[37] X. Wang, W. Deng, and Y. Chen, *J. Chem. Phys.* **150**, 164121 (2019).

- [38] T. Akimoto, S. Shinkai, and Y. Aizawa, *J. Stat. Phys.* **158**, 476 (2015).
- [39] T. Akimoto and T. Miyaguchi, *Phys. Rev. E* **82**, 030102(R) (2010).
- [40] F. D. Stefani, J. P. Hoogenboom, and E. Barkai, *Phys. Today* **62**, 34 (2009).
- [41] G. Margolin and E. Barkai, *Phys. Rev. Lett.* **94**, 080601 (2005).
- [42] R. Metzler and J. Klafter, *Phys. Rep.* **339**, 1 (2000).
- [43] R. Kutner and J. Masoliver, *Eur. Phys. J. B* **90**, 50 (2017).
- [44] D. R. Cox, *Renewal Theory* (Methuen, London, 1962).
- [45] J. Lamperti, *Am. Math. Soc* **12**, 724 (1961).
- [46] A. Vezzani, E. Barkai, and R. Burioni, *Sci. Rep.* **10**, 2732 (2020).
- [47] R. Burioni and A. Vezzani, *J. Stat. Mech.* (2020) 034005.
- [48] H. Touchette, *Phys. Rep.* **478**, 1 (2009).
- [49] W. Wang, M. Höll, and E. Barkai, [arXiv:2007.02341](https://arxiv.org/abs/2007.02341) (2020) [Phys. Rev. E (to be published)].
- [50] C. Godrèche, [arXiv:2006.04076](https://arxiv.org/abs/2006.04076).