HAL
open science

Condensation and extremes for a fluctuating number of independent random variables

Claude Godrèche

▶ To cite this version:

Claude Godrèche. Condensation and extremes for a fluctuating number of independent random variables. 2020. cea-03043371

HAL Id: cea-03043371
https://cea.hal.science/cea-03043371v1

Preprint submitted on 7 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Journal of Statistical Physics manuscript No.
(will be inserted by the editor)

Condensation and extremes for a fluctuating number of independent random variables

Claude Godrèche

arXiv:2006.04076v2 [cond-mat.stat-mech] 22 Oct 2020

Received: date / Accepted: date

Abstract We address the question of condensation and extremes for three classes of intimately related stochastic processes: (a) random allocation models and zero-range processes, (b) tied-down renewal processes, (c) free renewal processes. While for the former class the number of components of the system is kept fixed, for the two other classes it is a fluctuating quantity. Studies of these topics are scattered in the literature and usually dressed up in other clothing. We give a stripped-down account of the subject in the language of sums of independent random variables in order to free ourselves of the consideration of particular models and highlight the essentials. Besides giving a unified presentation of the theory, this work investigates facets so far unexplored in previous studies. Specifically, we show how the study of the class of random allocation models and zero-range processes can serve as a backdrop for the study of the two other classes of processes central to the present work—tied-down and free renewal processes. We then present new insights on the extreme value statistics of these three classes of processes which allow a deeper understanding of the mechanism of condensation and the quantitative analysis of the fluctuations of the condensate.

Keywords Condensation · extremes · renewal processes · zero-range processes · subexponentiality

1 Introduction

It is well known that n independent and identically distributed (iid) positive random variables conditioned by an atypical value of their sum exhibit the phenomenon of condensation, whereby one of the summands dominates upon the others, when their common distribution is subexponential (decaying more slowly than an exponential at large values of its argument).

Let $X_1, X_2, \ldots, X_n$ be these n random variables, henceforth taken discrete with positive integer values, whose common distribution is denoted by $f(k) = \mathrm{Prob}(X = k)$. Hereafter we consider the particular case of a subexponential distribution with asymptotic power-law decay[1]

$$f(k) \approx \frac{c}{k^{1+\theta}}, \tag{1.1}$$

Claude Godrèche
Université Paris-Saclay, CNRS, CEA, Institut de Physique Théorique, 91191 Gif-sur-Yvette, France
E-mail: claude.godreche@ipht.fr

[1] In the further course of this work, the symbol $\approx$ stands for asymptotic equivalence; the symbol $\sim$ means either 'of the order of', or 'with exponential accuracy', depending on the context.

where the index θ and the tail parameter c are both positive. Assume—for the time being—that the first moment $\langle X \rangle$ is finite (hence $\theta > 1$) and that the sum of these random variables,

$$S_n = \sum_{i=1}^{n} X_i, \tag{1.2}$$

is conditioned to take the atypically large value $L > \langle S_n \rangle = n\langle X \rangle$. In the present context the phenomenon of condensation has a simple pictorial representation. Let us consider the partial sums $S_1, S_2, \ldots$ as the successive positions of a random walk whose steps are the summands X_i. Such a representation is used in figure 1, which depicts six different paths of this walk, conditioned by a large, atypical value of its final position S_n after n steps. The steps have distribution (1.1) with $\theta = 3/2$ (see the caption for details). As can be seen on this figure, for most of the paths there is a single big step bearing the excess difference $\Delta = L - n\langle X \rangle$. In the thermodynamic limit, $L \to \infty$, $n \to \infty$ with $\rho = L/n$ fixed, the distribution of the size of this big step becomes narrow around Δ. This picture gives the gist of the phenomenon of condensation.

The two ingredients responsible for such a phenomenon are: (i) subexponentiality of the distribution $f(k)$, and: (ii) conditioning by an atypically large value of the sum S_n. In contrast, keeping the same distribution (1.1), but conditioning the sum S_n to be less than or equal to $n\langle X \rangle$, yields a 'democratic' situation where all the steps X_i are on the same footing, sharing the—now negative—excess difference Δ. Otherwise stated, in this situation, the path responds 'elastically' to the conditioning.

When the distribution $f(k)$ is exponential (e.g., a geometric distribution) the paths responds 'elastically' in all cases, i.e., regardless of whether the sum S_n is conditioned to take an atypical value larger or smaller than $n\langle X \rangle$. In both cases all the steps X_i are on the same footing, sharing the excess difference Δ (which is now either positive or negative).

This scenario of condensation has been investigated in great detail and is basically understood. It is for example encountered in random allocation models where n sites (or boxes) contain altogether L particles, the X_i representing the occupations of these sites [1,2]. This situation in turn accounts for the stationary state of dynamical urn models such as zero range processes (ZRP) or variants [3–16]. For short we shall refer to this class of models as the class of random allocation models and ZRP. Condensation means that one of the sites contains a macroscopic fraction of all the particles.

Here we shall be concerned by a different situation where the number of random variables $X_1, X_2, \ldots$, is itself a random variable, henceforth denoted by N_L and defined by conditioning the sum,

$$S_{N_L} = \sum_{i=1}^{N_L} X_i, \tag{1.3}$$

to satisfy either the inequality

$$S_{N_L} < L < S_{N_L+1}, \tag{1.4}$$

or the equality

$$S_{N_L} = L, \tag{1.5}$$

which both unambiguously determine N_L, and where L is a given positive integer number, see figure 2. These conditions are imposed irrespectively of whether the mean $\langle X \rangle$ is finite or not.

The process conditioned by the inequality (1.4) defines a *free renewal process* [17–20], the process conditioned by the equality (1.5) defines a *tied-down renewal process* [21–23]. In the former case the process is pinned at the origin, in the latter case it is also pinned at the end point. For both, the random variables X_i are the sizes of the iid (spatial or temporal) intervals between two renewals. Using the temporal language, the sum S_{N_L} is the time of occurrence

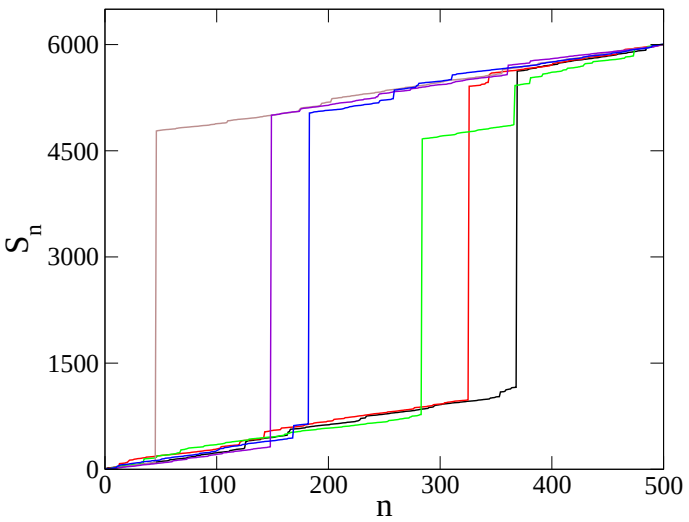

Fig. 1 Pictorial illustration of the phenomenon of condensation for the random allocation models and ZRP class. This figure depicts six paths of a random walk whose positions are given by the partial sums $S_1, S_2, \ldots$. The distribution of the steps X_i is given by $f(k) = 3k^{-5/2}/2$ ($k > 1$ is taken continuous), for which $\langle X \rangle = 3$. The random walk is conditioned to end at position $L = 6000$ at time $n = 500$. For each trajectory one can observe the occurrence of a 'big step' whose magnitude fluctuates around $\Delta = L - \langle S_n \rangle = 4500$.

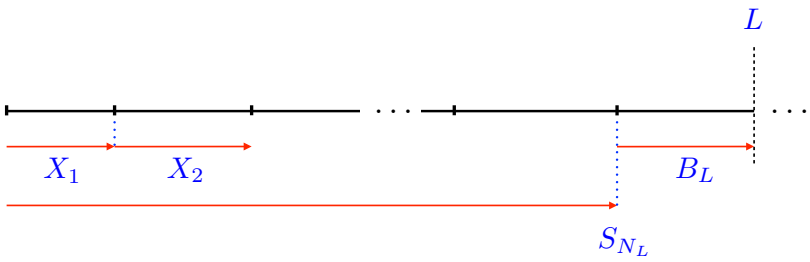

Fig. 2 In a free renewal process, the number N_L of intervals before L is a random variable such that $S_{N_L} < L < S_{N_{L+1}}$, where S_{N_L} is the sum of the N_L intervals $X_1, X_2, \ldots, X_{N_L}$. The last interval $B_L = L - S_{N_L}$ is unfinished. For a tied-down renewal process $B_L = 0$, $S_{N_L} = L$.

of the last renewal before or at time L. The last, unfinished, interval $B_L = L - S_{N_L}$ is known as the backward recurrence time in renewal theory. Tied-down renewal processes (TDRP) are special because the pinning condition (1.5) imposes $B_L = 0$.

A simple implementation of a TDRP is provided by the Bernoulli bridge, or tied-down random walk, made of ± 1 steps, starting from the origin and ending at the origin at time $2L$ [21,22,24]. The sizes of the intervals between the successive passages by the origin of the walk (where each tick mark on the $x-$axis represents two units of time) represent the random variables X_i, as depicted in figure 3. The continuum limit of the tied-down random walk is

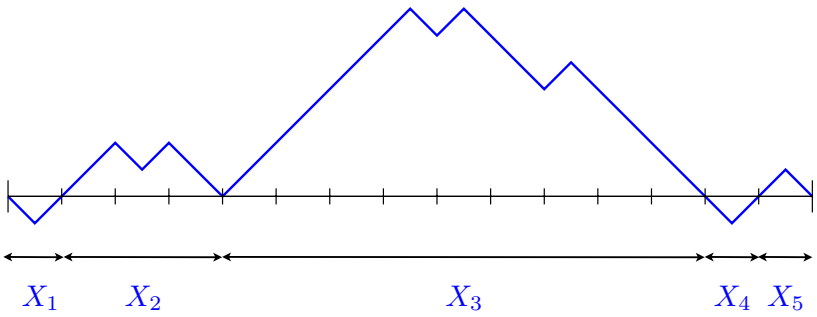

Fig. 3 A tied-down random walk, or Bernoulli bridge, is a simple random walk, with steps ± 1, starting and ending at the origin. Time is along the $x-$axis, space along the $y-$axis. The tick marks on the $x-$axis correspond to two time-steps. In this example the walk is made of $L = 15$ tick marks, with $N_{15} = 5$ intervals between zeros, $X_1, \ldots, X_5$, taking the values $1, 3, 9, 1, 1$ tick marks, respectively. The distribution of the sizes of the intervals, $f(k) = \mathrm{Prob}(X = k)$, is given by (4.1).

the *Brownian bridge*, also known as *tied-down Brownian motion* or else *pinned Brownian motion* [25].

For renewal processes (both free or tied-down) with a subexponential distribution $f(k)$, we shall show that, by weighting the configurations according to the number N_L of summands, a phase transition occurs, as $L \to \infty$, when the positive weight parameter w conjugate to N_L, varies from larger values, favouring configurations with a large number of summands, to smaller ones, favouring atypical configurations with a smaller number of summands. In the former case the weight parameter w is to be interpreted as a reward, in the latter case as a penalty.

Characterising this transition is the aim of the present work, with main focus on the quantitative analysis of the fluctuations of the condensate. Again, the occurrence of the phenomenon of condensation is due to: (i) subexponentiality of the distribution $f(k)$, and: (ii) atypicality of the configurations.

Tied-down renewal processes fall into the class of *linear systems* considered by Fisher [26]. The latter are defined as one-dimensional chains of total length L, made up, e.g., of alternating intervals of two kinds, A and B. This class encompasses the Poland-Scheraga model [27,28], consisting of an alternating sequence of straight paths A and loops B, wetting models, where A and B represent two phases, etc. If the direction of the chain is taken as a time axis, the loops in the perpendicular direction can be seen as random walks. The Bernoulli bridge or tied-down random walk of figure 3 is a natural implementation of this situation, where there is only one kind of intervals, say the loops B, representing the intervals between two passages at the origin of the walk. In the same vein, a variant of the random allocation model defined in [1] considers the case where the number of sites is varying [29], with an occupation variable X starting at $k = 1$. This model, as well as the spin domain model considered in [30–32] are examples of linear systems with only one kind of intervals. Both models are actually equivalent and are just particular instances of the TDRP considered in the present work. Let us finally mention the random walk (or polymer) models considered in [33, 34] which are free or tied-down renewal processes with a penalty (or reward) at each renewal events, as in the present work. In these models the condensation transition is interpreted as the transition between a localised phase and a delocalised one [33,34]. For the example of figure 3, a localised configuration corresponds to many contacts of the walk with the origin, while a delocalised one corresponds to the presence of a macroscopic excursion.

Renewal theory is a classic in probability studies. It is also ubiquitous in statistical physics and has a wide range of applicability (see examples in [35–38]). Yet, besides [33,34], which give rigorous mathematical results, studies of weighted free renewal processes are scarce. In particular, there are no available detailed characterisation of the condensation phenomenon for these processes in the existing literature, nor considerations on the statistics of extremes in the condensed phase.

We now describe the organisation of the paper, with further details on the literature on the subject.

The text is composed of three parts, which are respectively section 2, dealing with the class of random allocation models and ZRP, sections 3 to 7, dealing with TDRP, and finally, sections 8 to 10, dealing with free renewal processes. These three parts are both conceptually and analytically related.

Section 2 is a short presentation of how condensation arises for random allocation models and ZRP. The first two subsections, §2.1.1 and §2.1.2, give the basic formalism. Equation (2.20) introduces a simple remark on the distribution of the maximum in the second half ($k > L/2$), which will turn out to be instrumental in §2.3, and which will be generalized in the later sections on renewal processes. Subsections §2.2 and §2.2.2 summarise the main features of the phenomenon of condensation in the thermodynamic limit. This classical topic has been much investigated in the past, both in statistical physics and in mathematics. The summary given in these subsections rely on the recent review [16] to which we refer the reader for further bibliographical references. A complete mathematical review of the subject can be found in [2].

Subsection §2.3, which will serve as a backdrop for the study contained in the two other parts, contains novel aspects of the phenomenon of condensation for the class of models at hand. Namely, instead of considering the thermodynamic limit $L \to \infty$, $n \to \infty$ with fixed ratio $\rho = L/n$, we investigate the situation where the number of summands n is kept fixed and the value of the sum $S_n = L$ increases to infinity. Such a framework is precisely that considered in [12]. In this reference it is shown that for n fixed, $L \to \infty$, assuming that $\rho_c = \langle X \rangle$ exists, i.e., $\theta > 1$, if the largest summand is removed, the measure on the remaining summands converges to the product measure with density ρ_c, a feature which is apparent on figure 1. Therefore the largest summand is the unique condensate, with size $L - (n-1)\langle X \rangle$. As emphasised in [12] '*this phenomenon is a combinatorial fact that can be observed without making the number of sites grow to infinity*'. Simply stated, there is total condensation in this case, in the sense that the condensate essentially bears the totality of the L particles. We shall prove this result by elementary means in subsection §2.3 (see (2.48)) and extend it to the case where $\theta < 1$, thus proving that even though the first moment $\rho_c = \langle X \rangle$ does not exist, there is however (total) condensation[2]. If $\theta < 1$, the correction of the mean largest summand to L—in other words the fluctuations of the condensate—scales as $L^{1-\theta}$, with a known amplitude, given in (2.46).

It turns out that this scenario of (total) condensation is precisely that prevailing for the two other processes studied in the present work, namely, tied-down and free renewal processes. This scenario will be the red thread for the rest of the paper, with all the complications introduced by a now fluctuating number of summands. This red thread in particular links the three figures 5, 10 and 12, and equations (2.46), (2.48), (7.31), (7.32), (10.27) and (10.28), which are the central results of the present study.

Sections 3 to 7 are devoted to TDRP with a power-law distribution of summands (1.1). Section 3 gives a systematic presentation of the formalism, valid for any arbitrary distribution $f(k)$, followed, in sections 5 to 7 by the analysis of the behaviour of the system in the different phases, when $f(k)$ is a subexponential distribution of the form (1.1). As said in the abstract,

[2] The fact that condensation also occurs when $\theta < 1$ in the present context has been previously mentioned in [39]. I am indebted to S Grosskinsky for pointing this to me.

these topics are scattered in the literature [29–32]. The analyses presented in sections 5 to 7 are comprehensive and go deeper than previous studies, especially in the description of the condensation phenomenon and in the analysis of the statistics of extremes, as detailed in section 7.

Likewise, sections 8 to 10, devoted to free renewal processes with power-law distribution of summands (1.1), give a thorough analysis both of the formalism and of the phenomenon of condensation. It turns out that this case, which is more generic than that of TDRP since the process is not pinned at the end point L, is yet more complicated to analyse, the reason being that besides the intervals X_i, the last interval B_L, depicted in figure 2, enters the analysis.

Section 11, together with tables 1 and 2, gives a summary of the present study.

2 Condensation for random allocation models and ZRP

The key quantity for the study of the processes described above (random allocation models and ZRP, free and tied-down renewal processes) is the statistical weight of a configuration, or in the language of the random walk of figure 1, the statistical weight of a path.

The choice of conventions on the initial values of n, L, k is a matter of convenience which depends on the kind of reality that we wish to describe, as will appear shortly.

We start by giving, in §2.1.1 and §2.1.2, some elements of the formalism, where $f(k) = \mathrm{Prob}(X = k)$ is any arbitrary distribution of the positive random variable X. In the rest of section 2, $f(k)$ has a power-law tail (1.1).

2.1 General formalism

2.1.1 Statistical weight of a configuration

Let $X_1, X_2, \ldots, X_n$ be n positive iid integer random variables with sum S_n conditioned to be equal to L. The joint conditional probability associated to a configuration $\{X_1 = k_1, \ldots, X_n = k_n\}$, with $S_n = L$ given, reads[3]

$$\begin{aligned} p(k_1, \ldots, k_n|L) &= \mathrm{Prob}(X_1 = k_1, \ldots, X_n = k_n|S_n = L) \\ &= \frac{1}{Z_n(L)} f(k_1) \ldots f(k_n)\, \delta\Big(\sum_{i=1}^{n} k_i, L\Big), \end{aligned} \tag{2.1}$$

where $\delta(.,.)$ is the Kronecker delta, and where the denominator, whose presence stems from the constraint $S_n = L$, is the partition function

$$\begin{aligned} Z_n(L) &= \sum_{\{k_i\}} f(k_1) \ldots f(k_n) \delta\Big(\sum_{i=1}^{n} k_i, L\Big) = (f\star)^n(L) \\ &= \langle \delta(S_n, L)\rangle = \mathrm{Prob}(S_n = L). \end{aligned} \tag{2.2}$$

So $Z_n(L)$ is another notation for the distribution of the sum S_n.

The physical picture associated to these definitions correspond to a system of n sites (or boxes), L particles in total, and where the summands X_i are the occupation numbers of these sites, i.e., the number of particles on each of them. Since these sites can be empty, the occupation probability $f(0)$ is non zero in general. It is therefore natural to initialise L to 0.

[3] For the sake of simplicity we restrict the study to the case where $f(k)$ is a normalisable probability distribution.

In particular, the probability that the sum of occupations be zero is that all sites are empty, i.e.,

$$Z_n(0) = f(0)^n. \tag{2.3}$$

It also turns out to be convenient to start n at 0, and set

$$Z_0(L) = \delta(L, 0), \tag{2.4}$$

which serves as an initial condition for the recursion

$$Z_n(L) = \sum_{k=0}^{L} f(k) Z_{n-1}(L-k). \tag{2.5}$$

As can be seen either from (2.2) or (2.5), the generating function of $Z_n(L)$ with respect to L yields

$$\tilde{Z}_n(z) = \sum_{L\geq 0} z^L Z_n(L) = \tilde{f}(z)^n, \tag{2.6}$$

where the generating function of $f(k)$ with respect to k is

$$\tilde{f}(z) = \sum_{k\geq 0} z^k f(k). \tag{2.7}$$

The marginal distribution of the occupation of a generic site, say site 1, is

$$\pi_n(k|L) = \mathrm{Prob}(X_1 = k|S_n = L) = \langle \delta(X_1, k)\rangle = \frac{f(k) Z_{n-1}(L-k)}{Z_n(L)}, \tag{2.8}$$

where $\langle\cdot\rangle$ is the average with respect to (2.1). The mean conditional occupation is

$$\langle X|L\rangle = \sum_{k\geq 0} k\pi_n(k|L) = \frac{L}{n} = \rho. \tag{2.9}$$

2.1.2 Distribution of the largest occupation

Condensation corresponds to the presence of a site with a macroscopic occupation. We are therefore led to investigate the statistics of the largest occupation. This topic has been discussed in [2,7,8,13–16]. We briefly revisit this topic and supplement it with equation (2.20) at the end of this subsection which sheds some new light on the subject and lay the ground for renewal processes.

Let $X_{\max}$ be the largest summand (or occupation) under the conditioning $S_n = L$,

$$X_{\max} = \max(X_1, \ldots, X_n). \tag{2.10}$$

The distribution function of this variable is

$$F_n(k|L) = \mathrm{Prob}(X_{\max} \leq k|S_n = L) = \frac{\mathrm{Prob}(X_{\max} \leq k, S_n = L)}{Z_n(L)}, \tag{2.11}$$

whose numerator is

$$F_n(k|L)_{|\mathrm{num}} = \mathrm{Prob}(X_{\max} \leq k, S_n = L) = \sum_{k_1=0}^{k} f(k_1) \cdots \sum_{k_n=0}^{k} f(k_n) \delta\Big(\sum_{i=1}^{n} k_i, L\Big). \tag{2.12}$$

The generating function of the latter reads

$$\sum_{L\geq 0} z^L F_n(k|L)_{|\mathrm{num}} = \prod_{i=1}^{n} \Big(\sum_{k_i=0}^{k} f(k_i) z^{k_i}\Big) = \tilde{f}(z, k)^n, \tag{2.13}$$

where

$$\tilde{f}(z,k) = \sum_{j=0}^{k} z^j f(j). \tag{2.14}$$

The distribution of the largest occupation is thus given by the difference

$$p_n(k|L) = \mathrm{Prob}(X_{\max} = k|S_n = L) = F_n(k|L) - F_n(k-1|L), \tag{2.15}$$

where

$$F_n(0|L)_{|\mathrm{num}} = f(0)^n \delta(L,0). \tag{2.16}$$

Its generating function is

$$\sum_{L\geq 0} z^L p_n(k|L)_{|\mathrm{num}} = \tilde{f}(z,k)^n - \tilde{f}(z,k-1)^n. \tag{2.17}$$

The numerator (2.12) obeys the recursion

$$F_n(k|L)_{|\mathrm{num}} = \sum_{j=0}^{\min(k,L)} f(j)\, F_{n-1}(k|L-j)_{|\mathrm{num}}, \tag{2.18}$$

with initial condition

$$F_0(k|L)_{|\mathrm{num}} = \delta(L,0). \tag{2.19}$$

Let us note that if the occupation number X_1 is larger than $L/2$, then it is necessarily the largest one, $X_{\max}$. If so, the probability distribution of the latter, $p_n(k|L)$, is identical to $n\pi_n(k|L)$, since there are n possible choices of the generic summand X_1. Denoting the restriction of $p_n(k|L)$ to the range $k > L/2$ by $q_n(k|L)$, we thus have

$$q_n(k|L) = n\pi_n(k|L) = \frac{nf(k)Z_{n-1}(L-k)}{Z_n(L)}. \tag{2.20}$$

We shall see later that this relation, as simple as it may seem, is instrumental for the analysis of the fluctuations of the condensate and extends naturally to the case of TDRP or free renewal processes. When $k > L/2$ we have also

$$F_n(k|L-k)_{|\mathrm{num}} = Z_n(L-k), \tag{2.21}$$

since $X_{\max}$ is necessarily less than k (with $k > L/2$) when $S_n = L-k$. Thus

$$q_n(k|L) = nf(k)F_{n-1}(k|L-k). \tag{2.22}$$

2.2 Phenomenology of condensation in the thermodynamic limit

This subsection is a reminder of well-known facts on the phenomenon of condensation for a thermodynamic system with a large number n of sites and large total occupation L, at fixed density $\rho = L/n$. More detailed accounts or complements on this topic can be found, e.g., in [1,2,8–11,13–16], with further bibliographical references contained in the last reference.

This reminder will help emphasising the differences between the scenario of condensation *in the thermodynamic limit*, described in the present subsection, with another scenario of condensation, to be described in §2.3, where L is still large, but n (the number of summands or sites) is kept fixed. In the latter regime, condensation will turn out to be *total*: the condensed fraction is asymptotically equal to unity.

For the time being, we consider the situation where n and L are both large, with $\rho = L/n$ kept fixed, assuming that $f(k)$ has a power-law tail (1.1) and that $\langle X\rangle = \rho_c$ is finite ($\theta > 1$).

2.2.1 Regimes for the single occupation distribution

Evidence for the existence of a condensate, i.e., a site with a macroscopic occupation, is demonstrated by the behaviour of the single occupation distribution (2.8). There are three regimes to consider, according to the respective values of ρ and ρ_c.

1. Subcritical regime ($\rho < \rho_c$)
The asymptotic estimate of the partition function $Z_n(L)$ is given by the saddle-point method

$$Z_n(L) = \oint \frac{\mathrm{d}z}{2\pi \mathrm{i} z^{L+1}}\, \tilde{f}(z)^n \sim \frac{\tilde{f}(z_0)^n}{z_0^L}, \tag{2.23}$$

where z_0 obeys the saddle-point (sp) equation

$$\frac{z_0 \tilde{f}'(z_0)}{\tilde{f}(z_0)} = \rho. \tag{2.24}$$

This equation has a solution $z_0(\rho)$ for any $\rho < \rho_c$. It follows that

$$\pi_n(k|L)_{\mathrm{sp}} \approx \frac{z_0^k f(k)}{\tilde{f}(z_0)}, \tag{2.25}$$

which is no longer dependent on L and n separately and only depends on their ratio ρ. Note that (2.24) and (2.25) entail that

$$\langle X|L\rangle_{\mathrm{sp}} = \sum_{k\geq 0} k\pi_n(k|L)_{\mathrm{sp}} \approx \frac{z_0 \tilde{f}'(z_0)}{\tilde{f}(z_0)} = \rho, \tag{2.26}$$

consistently with (2.9). In this regime, the system is made of a fluid of independent particles with common distribution (2.25).

2. Critical regime ($\rho = \rho_c$)
A phase transition occurs when the saddle-point value z_0 reaches the maximum value of z, equal to one, where $\tilde{f}(z)$ is singular, with a branch cut chosen to be on the negative axis. The bulk of the partition function is given by the generalised central limit theorem and

$$\pi_n(k|L) \approx f(k), \tag{2.27}$$

up to finite-size corrections. At criticality the equality $\langle X|L\rangle = \langle X\rangle$ holds identically. In this regime, the system is made of a critical fluid of independent particles with common distribution (2.27).

3. Supercritical regime ($\rho > \rho_c$)
In this regime the saddle-point equation (2.24) can no longer be satisfied because z_0 sticks to the head of the cut of $\tilde{f}(z)$. The excess difference,

$$\Delta = L - n\langle X\rangle = n(\rho - \rho_c), \tag{2.28}$$

instead of being equally shared by all the sites, is, with high probability, accommodated by a single site, the *condensate*. The partition function $Z_n(L)$ is asymptotically given by its right tail (see [16] for more details),

$$Z_n(L) \underset{n\to\infty}{\overset{L=n\rho}{\approx}} \frac{nc}{\Delta^{1+\theta}}. \tag{2.29}$$

In the supercritical regime, the marginal distribution $\pi_n(k|L)$ has different behaviours in the three regions of values of the occupation variable.

(a) The *critical background* corresponds to values of k finite, for which (2.27) holds again. The main contribution to the total weight comes from this region.

(b) The *condensate* is located in the region $k \approx \Delta$ (i.e., the difference $\Delta - k$ is subextensive). The ratio of $f(k) \approx c/\Delta^{1+\theta}$ to $Z_n(L)$, given by (2.29), is asymptotically equal to

$$\frac{f(k)}{Z_n(L)} \approx \frac{c/\Delta^{1/\theta}}{nc/\Delta^{1/\theta}} = \frac{1}{n}. \tag{2.30}$$

On the other hand, $Z_{n-1}(L-k)$, is given by its bulk since $L - k \approx n\rho_c$. Hence, if $1 < \theta < 2$,

$$\pi_n(k|L)_{|\mathrm{cond}} \approx \frac{1}{n} Z_{n-1}(L-k) \approx \frac{1}{n} \frac{1}{n^{1/\theta}} \mathcal{L}_{\theta,c} \left(\frac{\Delta - k}{n^{1/\theta}} \right), \tag{2.31}$$

where $\mathcal{L}_{\theta,c}$ is the stable Lévy distribution of index θ, asymmetry parameter $\beta = 1$, and tail parameter c [40], while, if $\theta > 2$,

$$\pi_n(k|L)_{|\mathrm{cond}} \approx \frac{1}{n} Z_{n-1}(L-k) \approx \frac{1}{n} \frac{1}{n^{1/2}} \mathcal{G} \left(\frac{\Delta - k}{n^{1/2}} \right), \tag{2.32}$$

where $\mathcal{G}$ is the Gaussian distribution [40]. These expressions describe the bulk of the fluctuating condensate which manifests itself by a hump in the marginal distribution $\pi_n(k|L)$, in the neighbourhood of $k \approx \Delta$, visible on figure 4. The weight of this region is obtained from (2.31) or (2.32), according to the value of θ, as

$$\mathrm{Prob}(X_1 \in \mathrm{cond}) = \sum_{k \in \mathrm{hump}} \pi_n(k|L)_{|\mathrm{cond}} \approx \frac{1}{n}, \tag{2.33}$$

which demonstrates that the excess difference Δ is typically borne by only one summand.

This hump becomes peaked in the thermodynamic limit. For a finite system, most often there is a single condensate, i.e., a site with a macroscopic occupation, while more rarely there are two sites with macroscopic occupations, both of order L. This situation corresponds to the *dip region*, described next.

(c) The range of values of k such that k and $\Delta - k$ are large and comparable, interpolates between the critical part of $\pi_n(k|L)$, for k or order 1, and the condensate, for k close to Δ. It corresponds to the *dip region* on figure 4. In this region, $Z_{n-1}(L-k)$ is given by its right tail (2.29). So, for any $\theta > 1$,

$$\pi_n(k|L)_{|\mathrm{dip}} \approx c \left[\frac{\Delta}{k(\Delta - k)} \right]^{1+\theta} \approx \frac{f(k) f(\Delta - k)}{f(\Delta)}. \tag{2.34}$$

The interpretation of this result is that in the dip region typical configurations where one summand takes the value k are such that the remaining $\Delta - k$ excess difference is borne by a single other summand. The dip region is therefore dominated by rare configurations where the excess difference is shared by *two* summands [41]. An example of such a configuration is the green trajectory in figure 1.

Setting $k = \lambda\Delta$ in (2.34) and introducing a cutoff $\Lambda = \epsilon\Delta$, the weight of these configurations can be estimated as

$$\mathrm{Prob}(X_1 \in \mathrm{dip}) = \sum_{k=\epsilon\Delta}^{(1-\epsilon)\Delta} \pi_n(k|L)_{|\mathrm{dip}} \sim \Delta^{-\theta} \sim n^{-\theta}. \tag{2.35}$$

The relative weights of the dip (2.35) and condensate (2.33) regions is therefore of order $n^{-(\theta-1)}$, i.e., the weight of events where the condensate is broken into two pieces of order n is subleading with respect to events with a single big summand.

2.2.2 Statistics of the largest summand in the condensed phase

In view of (2.33) and the following equations (2.34) and (2.35), we infer that, for k larger than $\Delta/2$, which is the centre of the dip, the excess difference Δ is typically borne by only one summand—namely the condensate $X_{\max}$—thus

$$p_n(k|L) \underset{k>\Delta/2}{\approx} n\pi_n(k|L). \tag{2.36}$$

According to (2.20) we know that the two sides of this equation are actually identical for $k > L/2$, for any finite values of n and L. Note that $L/2$ is always larger than $\Delta/2$, but it can be smaller or larger than Δ depending on whether ρ is larger or smaller than $2\rho_c$. The significance of (2.36) is that the equality (2.20), valid for $k > L/2$, extends asymptotically to the entire region $k > \Delta/2$.

More precise statements have been given on the asymptotic distribution of the largest summand $X_{\max}$ in [13–15,2]. The result is that, if $L = n\rho$, $\rho > \rho_c$, $n \to \infty$, the rescaled variable $n^{-1/\alpha}(\Delta - X_{\max})$ converges to a stable law of index α, with $\alpha = \theta$ if $\theta < 2$, or $\alpha = 2$ if $\theta > 2$. This means that, asymptotically, the probability distribution of $X_{\max}$ coincides, up to a factor n, with the estimates of the marginal density in the condensate region ($\Delta - k \sim n^{1/\alpha}$), that is with (2.31) or (2.32) according to the value of θ, which is precisely the content of (2.36).

On the other hand, denoting the $r-$th largest summand by $X^{(r)}$ ($r = 1, \ldots, n$), with $X^{(1)} \equiv X_{\max}$, the distributions of these ranked summands, denoted by $p_n^{(r)}(k|L)$ with $p_n^{(1)}(k|L) \equiv p_n(k|L)$, sum up exactly to

$$p_n(k|L) + \sum_{r=2}^{n} p_n^{(r)}(k|L) = n\pi_n(k|L). \tag{2.37}$$

Thus according to (2.36) the sum upon $r \geq 2$ in the left side of (2.37) is negligible for $k > \Delta/2$.

As shown in [2] the distribution of the second largest summand, $X^{(2)}$, is asymptotically Fréchet, and the subsequent ones, $X^{(r)}$ ($r \geq 2$), are the order statistics of $n-1$ iid random variables X_i with distribution $f(k)$, which amounts to saying that, in the supercritical regime, the dependency between the summands X_i introduced by the conditioning goes asymptotically in the big jump $X_{\max}$.

Since $X_{\max}$ typically scales as n, while $X^{(2)}, X^{(3)}, \ldots$ typically scale as $n^{1/\theta}$, the condensate is increasingly separated from the background as n increases, leaving space to the dip region (k and $\Delta - k$ large and comparable). We know from the analysis made in §2.2 (see discussion following (2.34)) that this region is dominated by configurations where the excess difference is shared by two summands, namely $X_{\max}$ and $X^{(2)}$, so

$$p_n(k|L) + p_n^{(2)}(k|L) \underset{k\in\text{dip}}{\approx} n\pi_n(k|L)|_{\text{dip}}, \tag{2.38}$$

and that the contributions of these events to $n\pi_n(k|L)$ are of relative order $n^{-(\theta-1)}$. To the right of $\Delta/2$ the predominant contribution to the sum on the left side of (2.38) comes from $p_n(k|L)$, to the left of $\Delta/2$ it comes from $p_n^{(2)}(k|L)$.

An illustration

Figure 4 depicts a comparison between $n\pi_n(k|L)$, obtained from (2.8), and $p_n(k|L)$ obtained from (2.17), on the following example, defined by the normalised distribution

$$f(k) = \frac{1}{\zeta(1+\theta)} \frac{1}{(k+1)^{1+\theta}}, \quad (k \geq 0), \tag{2.39}$$

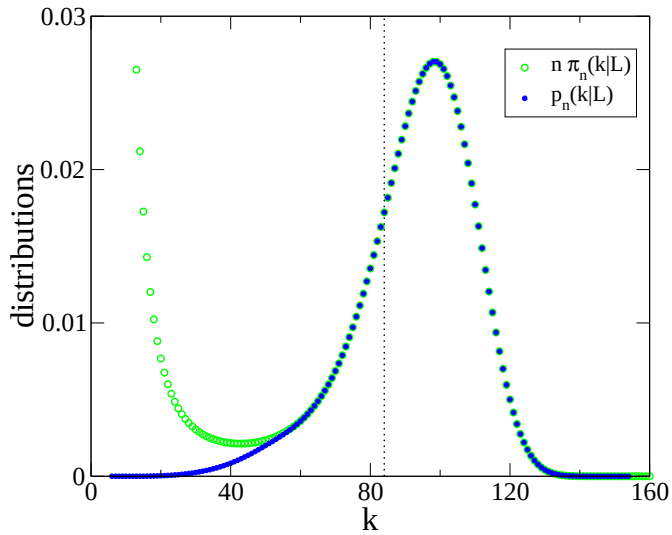

Fig. 4 Random allocations models and ZRP: comparison of $n\pi_n(k|L)$ (where $\pi_n(k|L)$ is the single occupation distribution) with $p_n(k|L)$ (distribution of the maximum) for the example (2.39), with $\theta = 3$, $\rho_c \approx 0.1106$, $n = 600$, $L = 168$, $\Delta \approx 102$. The vertical dotted line is at $L/2$. There is an exact identity between $n\pi_n(k|L)$ and $p_n(k|L)$, for $k > L/2$, as explained in §2.1.2. Moreover, there is already excellent numerical coincidence between the two curves as soon as $k > 58 \gtrsim \Delta/2$.

where $\zeta(s) = \sum_{s\geq 1} 1/n^s$ is the Riemann zeta function. This model has been introduced in [1], then further investigated in [5,6,42].

In this figure $\theta = 3$, $\rho_c \approx 0.1106$, $n = 600$, $L = 168$. This choice of parameters corresponds to a density $\rho = 0.28$ slightly larger than $2\rho_c$, where Δ and $L/2$ coincide. The top of the hump is approximately located at $\Delta \approx 102$ and the minimum of the dip is a bit less than $\Delta/2$. These curves are practically indiscernible as soon as $k \approx 58$, which is less than $L/2 = 84$, indicated by the vertical dotted line on the figure, from which the identity becomes exact.

Finally, let us compute the mean condensed fraction $X_{\max}/L$. For $\theta > 2$, using (2.32), it can be estimated as

$$\begin{aligned}\frac{\langle X_{\max}\rangle}{L} &= \frac{1}{L}\sum_{k=0}^{L} k\, p_n(k|L) \approx \frac{1}{L}\sum_{k=\Delta/2}^{L} k\, n\, \pi_n(k|L) \\ &\approx \frac{1}{L}\frac{1}{n^{1/2}\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}k\, k\, \mathrm{e}^{-\frac{1}{2}\left(\frac{\Delta-k}{n^{1/2}}\right)^2} = 1 - \frac{\rho_c}{\rho} = \frac{\Delta}{L}. \end{aligned} \tag{2.40}$$

The same result holds if $1 < \theta < 2$, using (2.31).

As ρ increases, the peak of the condensate moves towards the right end L, hence if $\rho \gg \rho_c$ the condensed fraction tends to unity, corresponding to total condensation. As detailed in §2.3, this scenario still holds when n is kept fixed.

2.3 Phenomenon of total condensation when n is kept fixed and $L \to \infty$

As seen above, if $\rho \gg \rho_c$, condensation becomes total, and the peak of the condensate is asymptotically located at L. As we now show, this still holds true if the number of summands

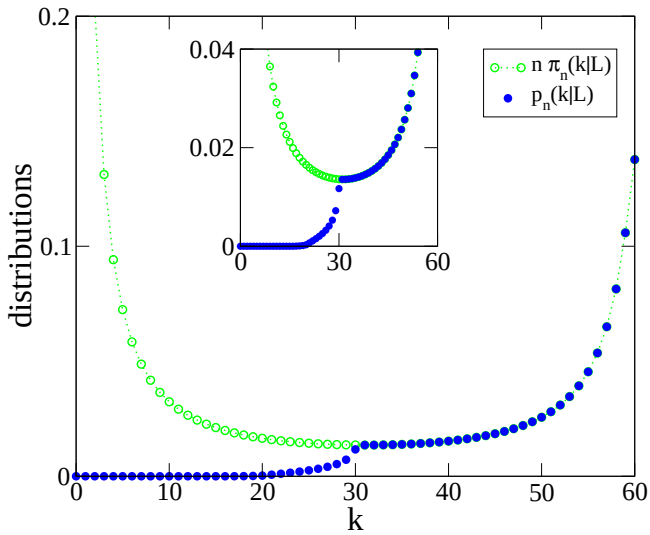

Fig. 5 Random allocations models and ZRP: comparison of $n\pi_n(k|L)$ with $p_n(k|L)$ for the example (2.41) corresponding to a tail index $\theta = 1/2$, with $L = 60$, $n = 4$. The inset highlights the cusp at $L/2$.

n is kept fixed, and L is large, irrespective of the existence of a first moment $\rho_c = \langle X \rangle$, or in other words, irrespective of whether θ is smaller or larger than one. Existence of condensation in such a situation has been pointed out in [12,39]. The quantitative characterisation of this phenomenon is the aim of this subsection.

2.3.1 An illustration

We start by giving an illustration of the phenomenon on the following example, corresponding to a tail index $\theta = 1/2$ for the decay of $f(k)$,

$$\tilde{f}(z) = \frac{1 - \sqrt{1-z}}{z}, \tag{2.41}$$

which entails that

$$f(k) = \frac{1}{2^{2k+1}} \frac{(2k)!}{k!(k+1)!} \underset{k\to\infty}{\approx} \frac{c}{k^{3/2}}, \tag{2.42}$$

with $c = 1/(2\sqrt{\pi})$, and

$$Z_n(L) = \frac{n}{2^{2L+n}(2L+n)} \binom{2L+n}{L} \overset{n \text{ fixed}}{\underset{L\to\infty}{\approx}} \frac{nc}{L^{3/2}}, \tag{2.43}$$

for n kept fixed, L large, which is a different regime from that leading to (2.29). In the general case, whenever $f(k)$ obeys (1.1), the asymptotic estimate (2.43) becomes

$$Z_n(L) \overset{n \text{ fixed}}{\underset{L\to\infty}{\approx}} \frac{nc}{L^{1+\theta}}, \tag{2.44}$$

as can be deduced from (2.6).

Figure 5 depicts a comparison of $n\pi_n(k|L)$ with $p_n(k|L)$ for this example with $L = 60$, $n = 4$. These two quantities are identical for $k > L/2$, The inset highlights the existence of a cusp at $L/2$. The distribution of the maximum $p_n(k|L)$ is significantly depressed for $k < L/2$. It vanishes identically for $k < L/4$. These features are analysed in the two subsections below.

2.3.2 Fluctuations of the condensate

Let us turn to the general case. A measure of the fluctuations of the condensate is provided by the width of the peak of the maximum, i.e., the mass outside the condensate. It can be estimated by the sum (see §2.3.3 below for details)

$$L - \langle X_{\max}\rangle = \sum_{\ell=0}^{L-1} \ell\, p_n(L-\ell|L) \approx \sum_{\ell=0}^{L/2-1} \ell\, q_n(L-\ell|L) \approx \sum_{\ell=0}^{L/2-1} \ell\, n\pi_n(L-\ell|L), \tag{2.45}$$

where $\ell = L - k$. The dominant contribution to this sum depends on whether θ is smaller or larger than one.

• If $\theta < 1$, the dominant contribution comes from values of $\ell = L - k$ comparable to L. Setting $\ell = \lambda L$ in (2.45), we have (see §2.3.3 below for details)

$$L - \langle X_{\max}\rangle \approx (n-1)cL^{1-\theta} \int_0^{1/2} \mathrm{d}\lambda \frac{\lambda}{[(\lambda(1-\lambda)]^{1+\theta}} \approx (n-1)c\,\mathrm{B}_{\frac{1}{2}}\Big(1-\theta, -\theta\Big) L^{1-\theta}, \tag{2.46}$$

where the incomplete beta function is defined as

$$\mathrm{B}_x(a,b) = \int_0^x \mathrm{d}t\, t^{a-1}(1-t)^{b-1}. \tag{2.47}$$

For example, if $\theta = 1/2$, we have $\mathrm{B}_{\frac{1}{2}}\Big(1-\theta, -\theta\Big) = 2$.

• If $\theta > 1$, the main contribution comes from finite values of ℓ,

$$\begin{aligned} L - \langle X_{\max}\rangle &\approx n \sum_{\ell=0}^{L/2} \ell\, \frac{f(L-\ell) Z_{n-1}(\ell)}{Z_n(L)} \approx \sum_{\ell=0}^{L/2} \ell\, Z_{n-1}(\ell) \\ &\approx (n-1) \sum_{\ell=0}^{L/2} \ell\, f(\ell) \to (n-1) \sum_{\ell=0}^{\infty} \ell f(\ell) = (n-1)\langle X\rangle. \end{aligned} \tag{2.48}$$

This last result (2.48) has a simple interpretation. It says that the correction $L - \langle X_{\max}\rangle$ comes from the $n-1$ sites of the fluid, each with mean occupation $\langle X\rangle = \rho_c$, in accordance with the prediction made in [12] and recalled in the introduction.

2.3.3 A finer analysis

Let us now add some more details on the derivations made above. The aim of this subsection is to give a detailed analysis of the distributions in the various regimes, in order to eventually compute the corrections to the scaling expressions predicted in (2.46) and (2.48) above.

We start with the discussion of the regimes for the single occupation distribution $\pi_n(k|L)$. There are three regimes to consider:

1. *Downhill region.* For k finite, using (2.44), we have

$$\pi_n(k|L) \underset{L\to\infty}{\approx} \frac{n-1}{n} f(k), \tag{2.49}$$

reflecting the fact that, with probability $(n-1)/n$, a randomly chosen site belongs to the fluid.
Introducing a cutoff Λ, such that $1 \ll \Lambda \ll L$, the weight of this region can be estimated by the sum

$$\sum_{k=0}^{\Lambda} \pi_n(k|L) \approx \sum_{k=0}^{\Lambda} \frac{n-1}{n} f(k) \underset{L\to\infty}{\to} \sum_{k=0}^{\infty} \frac{n-1}{n} f(k) = 1 - \frac{1}{n}. \tag{2.50}$$

2. *Dip region.* In the dip region, where k and $L-k$ are simultaneously large, setting $k = \lambda L$ in (2.8) where $0 < \lambda < 1$, and using (2.44), yields the estimate

$$\pi_n(k|L) \underset{L\to\infty}{\approx} \frac{n-1}{n} \frac{f(k)f(L-k)}{f(L)} \approx \frac{n-1}{n} \frac{c}{L^{1+\theta}} \frac{1}{[\lambda(1-\lambda)]^{1+\theta}}, \tag{2.51}$$

In this region the distribution is therefore U-shaped: the most probable configurations are those where almost all the particles are located on one of two sites. The dip centred around $k = L/2$ becomes deeper and deeper with L.
The weight of this region reads, choosing $\Lambda = \epsilon L$,

$$\sum_{k=\Lambda}^{L-\Lambda} \pi_n(k|L) \approx \frac{n-1}{n} c L^{-\theta} \int_{\epsilon}^{1-\epsilon} \frac{\mathrm{d}\lambda}{[\lambda(1-\lambda)]^{1+\theta}}. \tag{2.52}$$

3. *Uphill region.* The condensate region corresponds to $\ell = L-k$ finite, where (2.8) simplifies into

$$\pi_n(k|L) \underset{L\to\infty}{\approx} \frac{1}{n} Z_{n-1}(\ell), \tag{2.53}$$

as in (2.31) or (2.32). The weight of this uphill region can be estimated as

$$\sum_{\ell=0}^{\Lambda} \pi_n(L-\ell|L) \approx \frac{1}{n} \sum_{\ell=0}^{\Lambda} Z_{n-1}(\ell) \to \frac{1}{n} \sum_{\ell=0}^{\infty} Z_{n-1}(\ell) = \frac{1}{n}, \tag{2.54}$$

where Λ is yet another cutoff, and where the last step is obtained by setting $z = 1$ in the expression of the generating function (2.6). Thus, as seen in (2.50) and (2.54) the weights of the downhill and uphill regions add up to one, in line with the fact that the contribution of the dip region is subdominant, as shown in (2.52) above.

We now proceed to the discussion of the regimes for the distribution of the maximum, $p_n(k|L)$. There are again three regimes to consider, that we describe in turn, starting from the right. In the two first regions, the uphill and dip regions, such that $k > L/2$, we have $q_n(k|L) = n\pi_n(k|L)$, whose estimates follow from those of $\pi_n(k|L)$ seen above.

1. *Uphill region*, for ℓ finite. Using (2.53), we have

$$q_n(L-\ell|L) \approx Z_{n-1}(\ell), \tag{2.55}$$

with weight equal to 1 up to the subleading corrections detailed in (2.57).
2. *Dip region*, for $L/2 < k \sim L-k$. We have, according to (2.51),

$$q_n(k|L) = n\pi_n(k|L) \approx \frac{(n-1)\,c}{L^{1+\theta}} \frac{1}{[\lambda(1-\lambda)]^{1+\theta}}. \tag{2.56}$$

The weight of this region is proportional to $L^{-\theta}$, as seen in (2.52).

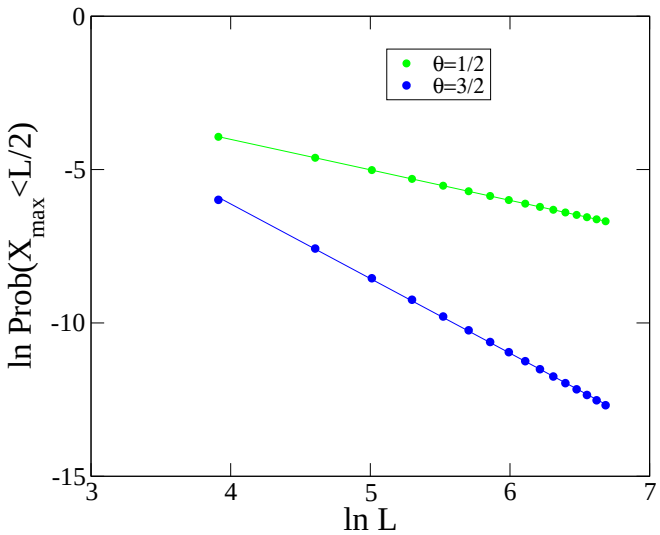

Fig. 6 Random allocations models and ZRP: comparison of the prediction $\mathrm{Prob}(X_{\max} < L/2) \sim L^{-\beta}$ (lines) for $n = 3$ summands and $\theta = 1/2$ ($\beta = 1$) and $\theta = 3/2$ ($\beta = 5/2$) with exact numerical computations (dots).

3. *Left region, $k < L/2$.* The weight of this region is subdominant with respect to that of the two previous ones. A simple heuristic argument shows that

$$\mathrm{Prob}(X_{\max} < L/2) = F_n(L/2|L) \sim L^{-\beta}, \tag{2.57}$$

(where F_n is defined in §2.1.2) irrespective of the value of n, with

$$\beta = \begin{cases} 2\theta & \text{if } \theta < 1 \\ 1+\theta & \text{if } \theta > 1. \end{cases} \tag{2.58}$$

This is illustrated in figure 6 where this probability is depicted for $\theta = 1/2$ and $\theta = 3/2$, with $n = 3$.

Equation (2.57) eventually justifies the approximation made in (2.45), where the contribution of the left region to the sum $\sum_{\ell=0}^{L-1} \ell\, p_n(L-\ell|L)$ was neglected. In view of (2.57), this contribution is $O(L^{1-\beta})$, thus subdominant by a factor $L^{-\theta}$ with respect to the first correction—respectively $O(L^{1-\theta})$ if $\theta < 1$ (see (2.46)), or $O(1)$ for $\theta > 1$ (see (2.48))—whether θ is smaller of larger than unity.

Finally, in appendix A, we give an exact calculation of the weight of the left region when $f(k)$ is the continuous Lévy $\frac{1}{2}$ stable density.

All the discussion given in the present subsection is a preparation for subsections §7.3 and §10.4, where figures 10 and 12 are to be compared to figure 5, and equations (7.31), (7.32), (10.27) and (10.28) are to be compared to equations (2.46) and (2.48).

3 General statements on tied-down renewal processes

The random variables X_i now represent the sizes of (spatial or temporal) intervals, that we take strictly positive, hence

$$f(0) = 0. \tag{3.1}$$

In the temporal language L is the total duration of the process, in the spatial language it is the length of the system. To each interval (equivalently, to each renewal event) is associated a positive weight w, to be interpreted as a reward if $w > 1$ or a penalty if $w < 1$. In the models considered in [29–32], w has the interpretation of the ratio y/y_c, where y is a fugacity, and y_c its value at criticality.

In the present section, $f(k) = \mathrm{Prob}(X = k)$ is any arbitrary distribution of the positive random variable X. Later on, in sections 4, 5, 6 and 7, $f(k)$ will obey the form (1.1).

3.1 Joint distribution

The probability of the configuration $\{X_1 = k_1, \dots, X_{N_L} = k_n, N_L = n\}$, given that $S_{N_L} = L$, reads

$$\begin{aligned} p(k_1, \dots, k_n, n|L) &= \mathrm{Prob}(X_1 = k_1, \dots, X_{N_L} = k_n, N_L = n|S_{N_L} = L) \\ &= \frac{1}{Z^{\mathrm{td}}(w, L)} w^n f(k_1) \dots f(k_n) \delta\Big(\sum_{i=1}^n k_i, L\Big), \end{aligned} \tag{3.2}$$

where the denominator is the tied-down partition function

$$\begin{aligned} Z^{\mathrm{td}}(w, L) &= \sum_{n\geq 0} w^n \sum_{\{k_i\}} f(k_1) \dots f(k_n) \delta\Big(\sum_{i=1}^n k_i, L\Big) \\ &= \sum_{n\geq 0} w^n Z_n(L) = \delta(L, 0) + \sum_{n\geq 1} w^n (f\star)^n(L). \end{aligned} \tag{3.3}$$

The probability $Z_n(L)$ is still defined as in (2.2), except for the change of the initial value (3.1) of $f(k)$, which entails that $Z_n(L)$ is only defined for $n \leq L$. The first term $\delta(L, 0)$ follows from (2.4). The first values of $Z^{\mathrm{td}}(w, L)$ are

$$\begin{aligned} Z^{\mathrm{td}}(w, 0) = 1, \quad Z^{\mathrm{td}}(w, 1) = wf(1), \quad Z^{\mathrm{td}}(w, 2) = wf(2) + w^2 f(1)^2, \\ Z^{\mathrm{td}}(w, 3) = wf(3) + 2w^2 f(1)f(2) + w^3 f(1)^3, \end{aligned} \tag{3.4}$$

and so on. The generating function of $Z^{\mathrm{td}}(w, L)$ with respect to L is

$$\tilde{Z}^{\mathrm{td}}(w, z) = \sum_{L\geq 0} z^L Z^{\mathrm{td}}(w, L) = \sum_{n\geq 0} w^n \tilde{f}(z)^n = \frac{1}{1 - w\tilde{f}(z)}. \tag{3.5}$$

Note that $Z^{\mathrm{td}}(w, L)$ can be seen as the grand canonical partition function of the system with respect to N_L.

For $w = 1$, the tied-down partition function,

$$Z^{\mathrm{td}}(1, L) = \sum_{n\geq 0} \mathrm{Prob}(S_n = L) = \mathrm{Prob}(S_{N_L} = L) = \langle \delta(S_{N_L}, L)\rangle, \tag{3.6}$$

is the probability that a renewal occurs at L.

Finally, we note that

$$p(k_1, \dots, k_n, n|L)_{|\mathrm{num}} = w^n p_n(k_1, \dots, k_n|L)_{|\mathrm{num}}, \tag{3.7}$$

hence

$$p(k_1, \dots, k_n, n|L) = p_n(k_1, \dots, k_n|L)\mathrm{p}_n(L), \tag{3.8}$$

where $p_n(\{k_i\}|L)$ and $\mathrm{p}_n(L)$ are respectively defined in (2.1) and (3.9).

3.2 Distribution of the number of intervals

The distribution of the number of intervals is obtained by summing the distribution (3.2) upon all variables except n,

$$\mathrm{p}_n(L) = \mathrm{Prob}(N_L = n) = \frac{w^n Z_n(L)}{Z^{\mathrm{td}}(w,L)}. \tag{3.9}$$

For instance, taking the successive terms of (3.3) divided by $Z^{\mathrm{td}}(w,L)$ yields

$$\mathrm{p}_0(L) = \frac{\delta(L,0)}{Z^{\mathrm{td}}(w,L)}, \quad \mathrm{p}_1(L) = \frac{wf(L)}{Z^{\mathrm{td}}(w,L)}, \quad \mathrm{p}_2(L) = \frac{w^2 \sum_{k_1} f(k_1) f(L-k_1)}{Z^{\mathrm{td}}(w,L)}, \ldots \tag{3.10}$$

and more generally, for $n \geq 1$,

$$\mathrm{p}_n(L) = \frac{w^n (f\star)^n(L)}{Z^{\mathrm{td}}(w,L)} = \frac{w^n \left[\tilde{f}(z)^n\right]_L}{Z^{\mathrm{td}}(w,L)}, \tag{3.11}$$

where the notation $[\cdot]_L$ stands for the $L-$th coefficient of the series inside the brackets. Hence the generating function with respect to L of the numerator of (3.9) reads

$$\sum_{L\geq 0} z^L \mathrm{p}_n(L)_{|\mathrm{num}} = w^n \tilde{f}(z)^n. \tag{3.12}$$

Taking the sum of the right side upon $n \geq 0$ yields back $\tilde{Z}^{\mathrm{td}}(w,z)$ given in (3.5).

The first moment of this distribution is by definition

$$\langle N_L \rangle = \sum_{n\geq 1} n\, \mathrm{p}_n(L). \tag{3.13}$$

The generating function of its numerator reads, using (3.12)

$$\sum_{L\geq 0} z^L \langle N_L \rangle_{|\mathrm{num}} = \sum_{n\geq 1} n (w\tilde{f}(z))^n = \frac{w\tilde{f}(z)}{\left(1 - w\tilde{f}(z)\right)^2} = w \frac{\mathrm{d}\tilde{Z}^{\mathrm{td}}(w,z)}{\mathrm{d}w}, \tag{3.14}$$

hence

$$\langle N_L \rangle = w\, \frac{\mathrm{d} \ln Z^{\mathrm{td}}(w,L)}{\mathrm{d}w}, \tag{3.15}$$

as expected in the grand canonical ensemble with respect to N_L. Alternatively, since

$$\sum_{L\geq 0} z^L \langle N_L \rangle_{|\mathrm{num}} = \tilde{Z}^{\mathrm{td}}(w,z)^2 - \tilde{Z}^{\mathrm{td}}(w,z), \tag{3.16}$$

we have

$$\langle N_L \rangle_{|\mathrm{num}} = (Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w,L) - Z^{\mathrm{td}}(w,L). \tag{3.17}$$

More generally, the generating function of the moments of N_L is given by

$$\langle v^{N_L} \rangle = \sum_{n\geq 0} v^n \mathrm{p}_n(L). \tag{3.18}$$

Taking the generating function with respect to L of the numerator of this expression, using (3.12),

$$\sum_{L\geq 0} z^L \langle v^{N_L} \rangle_{|\mathrm{num}} = \sum_{n\geq 0} v^n (w\tilde{f}(z))^n = \frac{1}{1 - vw\tilde{f}(z)} = \tilde{Z}^{\mathrm{td}}(vw,z), \tag{3.19}$$

yields

$$\langle v^{N_L}\rangle = \frac{Z^{\rm td}(vw,L)}{Z^{\rm td}(w,L)}. \tag{3.20}$$

Likewise, the inverse moment $\langle 1/N_L\rangle$ is

$$\Big\langle \frac{1}{N_L}\Big\rangle = \sum_{n\geq 1}\frac{{\rm p}_n(L)}{n} = \frac{1}{Z^{\rm td}(w,L)}\sum_{n\geq 1}\frac{\Big[[w\tilde f(z)]^n\Big]_L}{n} = \frac{1}{Z^{\rm td}(w,L)}\Big[-\ln(1-w\tilde f(z))\Big]_L. \tag{3.21}$$

3.3 Single interval distribution

The marginal distribution of one of the summands, say X_1, is by definition

$$\pi(k|L) = {\rm Prob}(X_1 = k|S_{N_L} = L) = \langle \delta(X_1,k)\rangle, \tag{3.22}$$

where $\langle\cdot\rangle$ is the average with respect to (3.2), with a summation upon the variables $k_1,\ldots,k_n$ (with $1\leq k\leq L$) and $n\geq 1$, resulting in

$$\begin{aligned}\pi(k|L)_{|\rm num} &= \sum_{n\geq 1}\sum_{k_1}\delta(k_1,k)wf(k_1)\sum_{k_2,\ldots}w^{n-1}f(k_2)\ldots f(k_n)\delta\Big(k_1+\sum_{i=2}^n k_i, L\Big)\\ &= \sum_{k_1}\delta(k_1,k)wf(k_1)\delta(k_1,L) + \sum_{k_1,k_2}\delta(k_1,k)w^2f(k_1)f(k_2)\delta(k_1+k_2,L)+\cdots\\ &= wf(k)\delta(k,L) + wf(k)Z^{\rm td}(w,L-k)\big(1-\delta(k,L)\big).\end{aligned} \tag{3.23}$$

Finally

$$\pi(k|L) = \underbrace{\frac{wf(k)}{Z^{\rm td}(w,L)}\delta(k,L)}_{p(k,1|L)} + wf(k)\frac{Z^{\rm td}(w,L-k)}{Z^{\rm td}(w,L)}\big(1-\delta(k,L)\big), \tag{3.24}$$

where the first term corresponds to $n=1$, i.e.,

$$\pi(L|L) = \frac{wf(L)}{Z^{\rm td}(w,L)} = {\rm Prob}(N_L = 1). \tag{3.25}$$

Also, since $Z^{\rm td}(w,0)=1$, (3.24) can be more compactly written as

$$\pi(k|L) = \frac{wf(k)Z^{\rm td}(w,L-k)}{Z^{\rm td}(w,L)}. \tag{3.26}$$

The generating function of the numerator of (3.26) with respect to L yields

$$\sum_{L\geq k} z^L\pi(k|L)_{|\rm num} = wz^kf(k)\tilde Z^{\rm td}(w,z) = \frac{wz^kf(k)}{1-w\tilde f(z)}. \tag{3.27}$$

Summing (3.26) upon k we obtain

$$Z^{\rm td}(w,L) = \sum_{k=1}^L wf(k)Z^{\rm td}(w,L-k), \quad L\geq 1, \tag{3.28}$$

which can also be obtained by multiplying the recursion (2.5) for $Z_n(L)$ by w^n and summing on n.

Remarks

1. An alternative route to (3.26) is as follows. We have

$$\begin{aligned}\pi(k|L)_{|\mathrm{num}} &= \sum_{n\geq 1} w^n \sum_{\{k_i\}} \delta(k_1,k) p_n(\{k_i\}|L)_{|\mathrm{num}} = \sum_{n\geq 1} w^n \pi_n(k|L)_{|\mathrm{num}} \\ &= \sum_{n\geq 1} w^n f(k) Z_{n-1}(L-k) = w f(k) \sum_{n\geq 0} w^n Z_n(L-k) \\ &= w f(k) Z^{\mathrm{td}}(w, L-k)_{|\mathrm{num}}. \end{aligned} \tag{3.29}$$

2. We also note that

$$\pi(k|L) = \sum_{n\geq 1} \pi_n(k|L) \mathrm{p}_n(L), \tag{3.30}$$

which is a simple consequence of (3.8).

3.4 Mean interval $\langle X|L\rangle$

This is, by definition,

$$\langle X|L\rangle = \sum_{k\geq 1} k\pi(k|L). \tag{3.31}$$

Multiplying (3.27) by k and summing upon k yields

$$\sum_{L\geq 1} z^L \langle X|L\rangle_{|\mathrm{num}} = \frac{wz\tilde{f}'(z)}{1-w\tilde{f}(z)}, \tag{3.32}$$

which can also be obtained by taking the derivative with respect to z of the expression for the inverse moment $\langle 1/N_L\rangle$. Indeed,

$$\langle X|L\rangle_{|\mathrm{num}} = \left\langle \frac{L}{N_L} \right\rangle_{|\mathrm{num}} = L\left[-\ln(1-w\tilde{f}(z))\right]_L, \tag{3.33}$$

then, taking the generating function of the right side gives

$$\begin{aligned}\sum_{L\geq 0} z^L \langle X|L\rangle_{|\mathrm{num}} &= \sum_{L\geq 1} z^L L\left[-\ln(1-w\tilde{f}(z))\right]_L \\ &= z\frac{\mathrm{d}}{\mathrm{d}z}\Big(-\ln(1-w\tilde{f}(z))\Big) = \frac{wz\tilde{f}'(z)}{1-w\tilde{f}(z)}. \end{aligned} \tag{3.34}$$

3.5 The longest interval

By definition, the longest interval is

$$X_{\max} = \max(X_1, \ldots, X_{N_L}). \tag{3.35}$$

Its distribution function is defined as

$$F(k|L) = \mathrm{Prob}(X_{\max} \leq k|L) = \sum_{n\geq 0} \sum_{k_1=1}^{k} \cdots \sum_{k_n=1}^{k} p(\{k_i\}, n|L) = \frac{F(k|L)_{|\mathrm{num}}}{Z^{\mathrm{td}}(w,L)}, \tag{3.36}$$

with initial value

$$F(k|0)_{|\mathrm{num}} = 1. \tag{3.37}$$

The numerator in (3.36) reads

$$F(k|L)_{|\mathrm{num}} = \sum_{n\geq 0} w^n \sum_{k_1=1}^{k} \cdots \sum_{k_n=1}^{k} p_n(\{k_i\}|L)_{|\mathrm{num}}$$
$$= \sum_{n\geq 0} w^n F_n(k|L)_{|\mathrm{num}}, \tag{3.38}$$

where $F_n(k|L)_{|\mathrm{num}}$ is defined in (2.12). Note that $F(L|L)_{|\mathrm{num}} = Z^{\mathrm{td}}(w, L)$, hence $F(L|L) = 1$. The generating function of the numerator is

$$\sum_{L\geq 0} z^L F(k|L)_{|\mathrm{num}} = 1 + \sum_{n\geq 1} \prod_{i=1}^{n} \Big(\sum_{k_i=1}^{k} w f(k_i) z^{k_i} \Big)$$
$$= 1 + \sum_{n\geq 1} \Big(w\tilde{f}(z,k) \Big)^n = \frac{1}{1 - w\tilde{f}(z,k)}, \tag{3.39}$$

where

$$\tilde{f}(z,k) = \sum_{j=1}^{k} z^j f(j). \tag{3.40}$$

The numerator obeys the recursion (renewal) equation, which generalises the Rosén-Wendel's result (2.4) of [22],

$$F(k|L)_{|\mathrm{num}} = \sum_{j=1}^{\min(k,L)} w f(j)\, F(k|L-j)_{|\mathrm{num}}, \tag{3.41}$$

with initial condition (3.37).

The distribution of $X_{\max}$ is given by the difference

$$p(k|L) = \mathrm{Prob}(X_{\max} = k) = F(k|L) - F(k-1|L), \tag{3.42}$$

where $F(0|L) = \delta(L, 0)$, with generating function

$$\sum_{L\geq 0} z^L p(k|L)_{|\mathrm{num}} = \frac{1}{1 - w\tilde{f}(z,k)} - \frac{1}{1 - w\tilde{f}(z,k-1)}$$
$$= \frac{w z^k f(k)}{[1 - w\tilde{f}(z,k)][1 - w\tilde{f}(z,k-1)]}. \tag{3.43}$$

Its end point value is the same as $\pi(L|L)$ (3.25), i.e.,

$$p(L|L) = \mathrm{Prob}(N_L = 1) = \frac{w f(L)}{Z^{\mathrm{td}}(w, L)}. \tag{3.44}$$

Note that (3.43) can be obtained by multiplying (2.17) by w^n and summing on n. In other words,

$$p(k|L)_{|\mathrm{num}} = \sum_{n\geq 1} w^n p_n(k|L)_{|\mathrm{num}}, \tag{3.45}$$

as can also be inferred from (3.38). And therefore (see (3.8))

$$p(k|L) = \sum_{n\geq 1} p_n(k|L) \mathrm{p}_n(L). \tag{3.46}$$

When $X_{\max} = k > L/2$, the longest interval is unique. Denoting the restriction of $p(k|L)$ to the range $k > L/2$ by $q(k|L)$, and generalising the reasoning made in [21,22] we can decompose a configuration into three contributions to obtain

$$\begin{aligned} q(k|L)_{|\mathrm{num}} &= \sum_{i=0}^{L-k} Z^{\mathrm{td}}(w,i) w f(k) Z^{\mathrm{td}}(w, L-k-i) \\ &= w f(k) (Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w, L-k), \end{aligned} \tag{3.47}$$

which, using (3.17), can be alternatively written as

$$q(k|L) = \frac{w f(k) Z^{\mathrm{td}}(w, L-k)}{Z^{\mathrm{td}}(w, L)} (1 + \langle N_{L-k} \rangle), \tag{3.48}$$

that is (see (3.26)),

$$q(k|L) = \pi(k|L)(1 + \langle N_{L-k} \rangle). \tag{3.49}$$

For $k = L$, (3.44) with (3.25) is recovered.

Remarks

1. An alternative route to (3.49) can be inferred from (3.45) as follows. We have

$$q(k|L)_{|\mathrm{num}} = \sum_{n \geq 1} w^n q_n(k|L)_{|\mathrm{num}} = \sum_{n \geq 1} w^n n \pi_n(k|L)_{|\mathrm{num}}, \tag{3.50}$$

where $\pi_n(k|L)$ is given in (2.8). So

$$\begin{aligned} q(k|L)_{|\mathrm{num}} &= w f(k) \sum_{n \geq 1} n w^{n-1} Z_{n-1}(L-k) \\ &= w f(k) \sum_{n \geq 0} (w^n Z_n(L-k) + n w^n Z_n(L-k)) \\ &= w f(k) (Z^{\mathrm{td}}(w, L-k) + \langle N_{L-k} \rangle_{|\mathrm{num}}), \end{aligned} \tag{3.51}$$

which, after division by $Z^{\mathrm{td}}(w, L)$, yields (3.49).
2. Comparison between (3.51) and (3.47) entails the equality

$$(Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w, L) = \sum_{n \geq 1} n w^{n-1} Z_{n-1}(L), \tag{3.52}$$

which can also be checked directly by taking the generating functions of both sides,

$$\tilde{Z}^{\mathrm{td}}(w, z)^2 = \sum_{n \geq 1} n w^{n-1} \tilde{f}(z)^{n-1} = \frac{1}{(1 - \tilde{w} f(z))^2}. \tag{3.53}$$

3. From (2.21) we infer that, if $k > L/2$,

$$F^{\mathrm{td}}(k|L-k)_{|\mathrm{num}} = Z^{\mathrm{td}}(w, L-k). \tag{3.54}$$

4 Phase transition for tied-down renewal processes

In this section and in the following sections 5, 6 and 7 the distribution $f(k) = \mathrm{Prob}(X = k)$ is taken subexponential with asymptotic power-law decay (1.1). Before discussing the phase diagram of the process, we give some illustrative examples of such distributions.

4.1 Illustrative examples

In the sequel, we shall illustrate the general results derived for TDRP in the current section, and for free renewal processes in section 8, on the following examples.

Example 1. This first example corresponds to the tied-down random walk of figure 3 on which we come back in more detail. The distribution of the size of intervals, $f(k)$, representing the probability of first return at the origin of the walk after $2k$ steps, or equivalently after k tick marks on figure 3, reads

$$f(k) = \frac{1}{2^{2k-1}} \frac{(2k-2)!}{(k-1)!k!} = \frac{\Gamma(k-1/2)}{2\sqrt{\pi}\Gamma(k+1)} \approx \frac{1}{2\sqrt{\pi}k^{3/2}}, \tag{4.1}$$

since the number of such walks is equal to $(2k-2)!/[(k-1)!k!]$. Its generating function reads

$$\tilde{f}(z) = 1 - \sqrt{1-z}. \tag{4.2}$$

The partition function (3.6) for $w = 1$ represents the probability that the walk returns at the origin after $2L$ steps, or equivalently after L tick marks,

$$Z^{\mathrm{td}}(1, L) = \frac{1}{2^{2L}} \binom{2L}{L} \approx \frac{1}{\sqrt{\pi L}}, \tag{4.3}$$

since the number of such walks is equal to $(2L)!/(L!)^2$. Its generating function reads

$$\tilde{Z}^{\mathrm{td}}(1, z) = \frac{1}{\sqrt{1-z}}. \tag{4.4}$$

Note that

$$f(L) = Z^{\mathrm{td}}(1, L-1) - Z^{\mathrm{td}}(1, L). \tag{4.5}$$

The partition function $Z_n(L)$ is explicit for this case,

$$Z_n(L) = \frac{n}{2^{2L-n}} \frac{(2L-n-1)!}{L!(L-n)!} \approx \frac{n}{2\sqrt{\pi}L^{3/2}}, \tag{4.6}$$

with $n \leq L$.

Example 2. This second example is defined for any $\theta > 0$ by

$$f(k) = \frac{1}{\zeta(1+\theta)} \frac{1}{k^{1+\theta}}, \quad k > 0. \tag{4.7}$$

So

$$\tilde{f}(z) = \frac{\mathrm{Li}_{1+\theta}(z)}{\zeta(1+\theta)}, \tag{4.8}$$

where $\mathrm{Li}_s(z) = \sum_{k\geq 1} z^k/k^s$ is the polylogarithm function. If $\theta > 1$ the mean $\langle X \rangle = \zeta(\theta)/\zeta(1+\theta)$. This is the distribution used, e.g., in [1,29–32,42].

4.2 Phase diagram

Demonstrating the existence of a phase transition in the model defined by (3.2), with distribution $f(k)$ given by (1.1), when w crosses the value one, is a classical subject. This model is a particular instance of a linear system, as described in [26], where the mechanism of the transition is explained in simple terms. This transition is also studied in [29–32] for *Example 2* (see (4.7)). Let us first analyse the large L behaviour of $Z^{\mathrm{td}}(w,L)$. Recalling (3.5) we have, for a contour encircling the origin,

$$Z^{\mathrm{td}}(w,L) = \oint \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{\tilde{Z}^{\mathrm{td}}(w,z)}{z^{L+1}} = \oint \frac{\mathrm{d}z}{2\pi\mathrm{i}\, z^{L+1}} \frac{1}{1-w\tilde{f}(z)}. \tag{4.9}$$

Since $\tilde{f}(z)$ is monotonically increasing for $z \in (0,1)$ the denominator of $\tilde{Z}^{\mathrm{td}}(w,z)$, $1-w\tilde{f}(z)$, is monotonically decreasing between 1 and $1-w$.

Disordered phase. If $w>1$, the denominator vanishes for $z=z_0<1$ such that $w\tilde{f}(z_0)=1$, hence $\tilde{Z}^{\mathrm{td}}(w,z)$ has a pole at z_0, and therefore $Z^{\mathrm{td}}(w,L)$ is exponentially increasing,

$$Z^{\mathrm{td}}(w,L) \approx \frac{z_0^{-L}}{w z_0 \tilde{f}'(z_0)}. \tag{4.10}$$

Critical regime. If $w=1$, then $z_0=1$. The asymptotic estimates of $Z^{\mathrm{td}}(1,L)$ are given in (6.5) and (6.23).

Condensed phase. If $w<1$, the denominator $1-w\tilde{f}(z)$ has no zero, but it is singular for $z=z_0=1$ (which is the singularity of $\tilde{f}(z)$). Hence z_0 sticks to 1. The asymptotic estimate of $Z^{\mathrm{td}}(w,L)$ is given in (7.1).

This is the *switch* mechanism of Fisher [26]: the condition determining z_0 switches from z_0 being the smallest root of the equation $1-w\tilde{f}(z)=0$ to being the closest real singularity of $\tilde{f}(z)$, which is a cut at $z=z_0=1$. This non analytical switch signals the phase transition. The free energy density [26]

$$\mathbf{f} = \lim_{L\to\infty} -\frac{1}{L} \ln Z^{\mathrm{td}}(w,L) = \ln z_0, \tag{4.11}$$

therefore vanishes when $w \le 1$. The three cases above are successively reviewed in the next sections.

5 Disordered phase ($w>1$) for tied-down renewal processes

The asymptotic expression at large L of the distribution of the size of a generic interval is obtained by carrying (4.10) in (3.26), which leads to

$$\pi(k|L) \approx w f(k) z_0^k = w f(k) \mathrm{e}^{-k/\xi}, \qquad \xi = \frac{1}{|\ln z_0|}, \tag{5.1}$$

where ξ is the correlation length, divergent at the transition. This expression is independent of L and normalised, since summing on k restores $w\tilde{f}(z_0)=1$. This exponentially decaying distribution has a finite mean,

$$\langle X|L\rangle \approx w z_0 \tilde{f}'(z_0), \tag{5.2}$$

an expression which can also be inferred from (3.32). Thus (4.10) can be recast as

$$Z^{\mathrm{td}}(w,L) \approx \frac{z_0^{-L}}{\langle X|L\rangle}. \tag{5.3}$$

The distribution of N_L is given by (3.9)

$$\mathrm{p}_n(L) = \frac{w^n Z_n(L)}{Z^{\mathrm{td}}(w,L)} \approx w^n Z_n(L)\, w\tilde{f}'(z_0) z_0^{L+1}. \tag{5.4}$$

This distribution obeys the central limit theorem, as illustrated on the example below. Using (3.15), we obtain the asymptotic expression of $\langle N_L \rangle$,

$$\langle N_L \rangle \approx -L \frac{w}{z_0} \frac{\mathrm{d}z_0}{\mathrm{d}w} \approx \frac{L}{\langle X|L \rangle}, \tag{5.5}$$

which means that

$$\frac{1}{\langle N_L \rangle} \approx \left\langle \frac{1}{N_L} \right\rangle. \tag{5.6}$$

Let us denote the density of points (or intervals) for a finite system as

$$\nu_L = \frac{\langle N_L \rangle}{L}, \tag{5.7}$$

then, asymptotically, we have

$$\nu = \lim_{L\to\infty} \nu_L = \lim_{L\to\infty} \frac{\langle N_L \rangle}{L} = \lim_{L\to\infty} \frac{1}{\langle X|L \rangle}. \tag{5.8}$$

We illustrate these general statements on *Example 1* (see (4.1)), for which z_0 is explicit,

$$z_0 = \frac{2w-1}{w^2}, \tag{5.9}$$

hence

$$\xi \approx (w-1)^{-2}, \tag{5.10}$$

and the following asymptotic expressions hold,

$$\begin{aligned}
Z^{\mathrm{td}}(w,L) &\approx \frac{2(w-1)w^{2L}}{(2w-1)^{L+1}}, \\
\langle X|L \rangle &\underset{L\to\infty}{\to} \frac{2w-1}{2(w-1)}, \\
\langle N_L \rangle &\approx \frac{L}{\langle X|L \rangle} + \frac{w}{(w-1)(2w-1)}, \\
\operatorname{Var} N_L &\approx L\frac{2w}{(2w-1)^2} - \frac{w(2w^2-1)}{(2w-1)^2(w-1)^2}, \\
\mathrm{p}_n(L) &\approx \frac{1}{\sqrt{2\pi \operatorname{Var} N_L}} \exp\Big(- \frac{(n-\langle N_L \rangle)^2}{2 \operatorname{Var} N_L} \Big), \\
\nu &= \frac{2(w-1)}{2w-1}.
\end{aligned} \tag{5.11}$$

Figure 7 depicts a comparison between the exact finite-size expression of the density ν_L obtained by means of (3.14) for $L = 1000$ as a function of w, and the asymptotic expression (5.11). It vanishes at the transition $w = 1$, where the system becomes critical.

More generally, if $\theta < 1$, close to the transition, we get

$$\nu \sim (w-1)^{1/\theta - 1}, \tag{5.12}$$

as can be easily inferred by means of the expansion (6.1), a result already present in [29], later recovered in [30]. If $\theta > 1$, the density ν tends to $1/\langle X \rangle$ when $w \to 1$, as can be seen

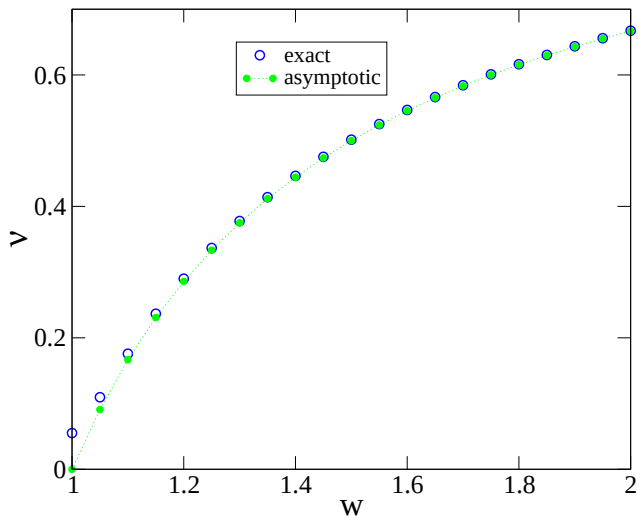

Fig. 7 TDRP: density of intervals ν against w in the disordered phase, for *Example 1* (see (4.1)). *Exact* refers to the finite-size expression (5.7) where $\langle N_L \rangle$ is extracted from (3.14) with $L = 1000$, *asymptotic* to (5.11).

on (5.2) and (5.8), using the expansion (6.2) (see also (6.24)). The density vanishes in the condensed phase since $\langle N_L \rangle$ is finite (see (7.5)), it is therefore discontinuous at the transition, as noted in [29,30]. Likewise, it is easy to see that

$$\xi^{-1} \sim \begin{cases} (w-1)^{1/\theta} & \theta < 1 \\ w - 1 & \theta > 1. \end{cases} \tag{5.13}$$

The correlation length diverges at the transition, while the order parameter ν is either continuous ($\theta < 1$) or discontinuous ($\theta > 1$), as seen above. The transition is therefore of mixed order [29,30].

Finally we note that the intervals X_i behave essentially as iid random variables, with distribution (5.1), hence the statistics of the longest interval belongs to the Gumbel class [43,44]. This is detailed on *Example 2* in [31].

6 Critical regime ($w = 1$) for tied-down renewal processes

In this regime, the behaviour of the quantities of interest strongly depends on whether the index θ is smaller or larger than unity. The discussion below is organised accordingly. Part of the material of this section can be found in more details in [21,22] and is also addressed in [31,32]. Here we summarise these former studies and complement them by a detailed analysis of the distribution of the number of intervals N_L and of the distribution $\pi(k|L)$ of the size of a generic interval. We also come back on the distribution of the longest interval.

If $w = 1$ the singularity is at $z = 1$, or, setting $z = \mathrm{e}^{-s}$, at $s = 0$. The generating function $\tilde{f}(z)$ becomes the Laplace transform $\hat{f}(s)$ which has the expansion

$$\hat{f}(s) \underset{s\to 0}{\approx} 1 - |a|s^{\theta}, \qquad \theta < 1 \tag{6.1}$$

$$\hat{f}(s) \underset{s\to 0}{\approx} 1 - s\langle X\rangle + \cdots + as^{\theta}, \qquad \theta > 1 \tag{6.2}$$

with

$$a = c\,\Gamma(-\theta) = \theta\Gamma(-\theta)k_0^{\theta}, \tag{6.3}$$

i.e., $c = \theta k_0^{\theta}$, where k_0 is a microscopic scale, defined as

$$g(k) = \sum_{j>k} f(j) \underset{k\to\infty}{\approx} \left(\frac{k_0}{k}\right)^{\theta}. \tag{6.4}$$

The parameter a is negative if $0 < \theta < 1$, positive if $1 < \theta < 2$, and so on. For instance, $\Gamma(-1/2) = -2\sqrt{\pi}$, $\Gamma(-3/2) = 4\sqrt{\pi}/3$, $\Gamma(-5/2) = -8\sqrt{\pi}/15$, and so on.

6.1 Distribution $f(k)$ with index $\theta < 1$

Since $\tilde{Z}^{\mathrm{td}}(1, z) = 1/(1 - \tilde{f}(z))$, in Laplace space we have $\hat{Z}^{\mathrm{td}}(1, s) \approx 1/as^{\theta}$ which yields the expression of the partition function (see (4.6) in [22])

$$Z^{\mathrm{td}}(1, L) \underset{L\to\infty}{\approx} \frac{\theta \sin \pi\theta}{\pi c} L^{\theta-1}. \tag{6.5}$$

For instance, setting $\theta = 1/2$ and $c = 1/(2\sqrt{\pi})$ restores (4.3).

6.1.1 The number of intervals

We have (see (4.10) in [22]),

$$\langle N_L\rangle \approx \frac{A(\theta)}{c} L^{\theta}, \qquad A(\theta) = \frac{\Gamma(1+\theta)}{\Gamma(1-\theta)\Gamma(2\theta)}, \tag{6.6}$$

which can be easily deduced from (3.14). For the specific case of *Example 1* (see (4.1)), we have the exact result (see (2.47) in [22])

$$\langle N_L\rangle = \frac{1}{Z^{\mathrm{td}}(1, L)} - 1 = \frac{2^{2L}}{\binom{2L}{L}} - 1 \approx \sqrt{\pi L}, \tag{6.7}$$

which is in agreement with (6.6), with $\theta = 1/2, c = 1/(2\sqrt{\pi})$. We know from (3.9) that the distribution of N_L is given by the ratio

$$\mathrm{p}_n(L) = \frac{Z_n(L)}{Z^{\mathrm{td}}(1, L)}. \tag{6.8}$$

For n and L large $\mathrm{p}_n(L)$ has a scaling form. On the one hand, according to the generalised central limit theorem, the scaling form of the numerator is given by

$$Z_n(L) \approx \frac{1}{n^{1/\theta}} \mathcal{L}_{\theta,c}\left(\frac{L}{n^{1/\theta}}\right), \tag{6.9}$$

where $\mathcal{L}_{\theta,c}$ is the density of the stable law of index θ, tail parameter c and asymmetry parameter $\beta = 1$ (see e.g., [16]). Then using (6.5), we get, with $u = L/n^{1/\theta}$,

$$\mathrm{p}_n(L) \approx \frac{\pi c}{\theta \sin \pi\theta} \frac{1}{L^\theta} \frac{L}{n^{1/\theta}} \mathcal{L}_{\theta,c}\left(\frac{L}{n^{1/\theta}}\right) \approx \frac{\pi c}{\theta \sin \pi\theta} \frac{1}{L^\theta} u \mathcal{L}_{\theta,c}(u). \tag{6.10}$$

The Lévy distribution of index $\theta = 1/2$ has the explicit expression

$$\mathcal{L}_{1/2,c}(u) = \frac{c\,\mathrm{e}^{-\pi c^2/u}}{u^{3/2}}, \tag{6.11}$$

hence (see (2.49) in [22]), for *Example 1* (see (4.1)),

$$\mathrm{p}_n(L) \approx \frac{v}{2\sqrt{L}} \mathrm{e}^{-v^2/4}, \quad v = \frac{1}{\sqrt{u}} = \frac{n}{\sqrt{L}}. \tag{6.12}$$

Moreover, for this example, for n and L finite, $\mathrm{p}_n(L)$ is explicit since both $Z_n(L)$ given by (4.6) and $Z^{\mathrm{td}}(1, L)$, given by (4.3), are known explicitly.

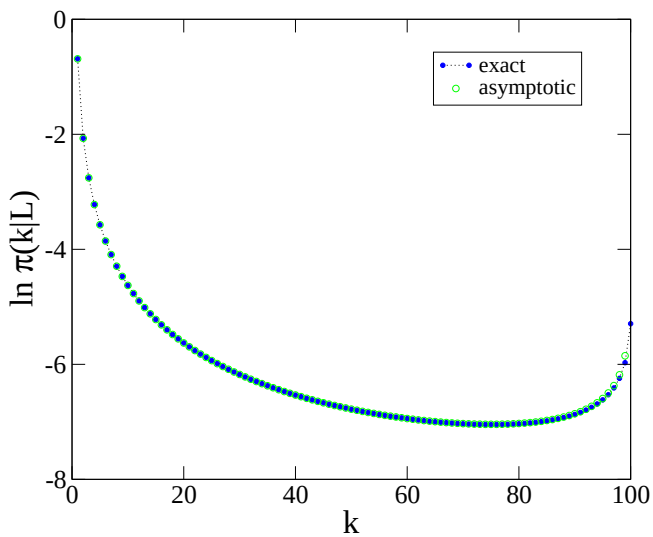

Fig. 8 TDRP: single interval distribution $\pi(k|L)$, with $L = 100$, for *Example 1* (see (4.1)), at criticality ($w = 1$). *Exact* refers to the middle expression in (6.13), *asymptotic* to the rightmost one. The $y-$axis is on a logarithmic scale.

6.1.2 Single interval distribution

Two regimes are to be considered.
(i) In all regimes where $\ell = L - k$ is large, using (6.5) we have

$$\pi(k|L) = f(k) \frac{Z^{\mathrm{td}}(1, L-k)}{Z^{\mathrm{td}}(1, L)} \approx f(k) \left(1 - \frac{k}{L}\right)^{\theta-1}. \tag{6.13}$$

For instance, if $1 \ll k \ll L$,

$$\pi(k|L) \approx f(k) \approx \frac{c}{k^{1+\theta}}, \tag{6.14}$$

while if $k = \lambda L$, with $\lambda \in (0,1)$,

$$\pi(k|L) \approx \frac{c}{\lambda^{1+\theta}(1-\lambda)^{1-\theta}} \frac{1}{L^{1+\theta}}, \tag{6.15}$$

which is minimum at $\lambda = (1+\theta)/2$.
(ii) On the other hand, if $\ell = L - k = O(1)$,

$$\pi(L-\ell|L) = f(L-\ell)\frac{Z^{\mathrm{td}}(1,\ell)}{Z^{\mathrm{td}}(1,L)} \approx \frac{\pi c^2 Z^{\mathrm{td}}(1,\ell)}{\theta \sin \pi\theta} \frac{1}{L^{2\theta}}. \tag{6.16}$$

In particular, for $k = L$,

$$\pi(L|L) = \frac{f(L)}{Z^{\mathrm{td}}(1,L)} \approx \frac{\pi c^2}{\theta \sin \pi\theta} \frac{1}{L^{2\theta}}. \tag{6.17}$$

In this regime the ratio of $\pi(k|L)$ to the estimate (6.13) reads

$$\frac{\pi(L-\ell|L)}{f(L-\ell)(1-k/L)^{\theta-1}} = \frac{\pi c\, \ell^{1-\theta} Z^{\mathrm{td}}(1,\ell)}{\theta \sin \pi\theta}, \tag{6.18}$$

which tends to one when ℓ becomes large, if one refers to (6.5).

All these results can be illustrated on *Example 1* (see (4.1)). For instance figure 8 gives a comparison between the exact expression of $\pi(k|L)$ computed for $L = 100$ by means of the middle expression in (6.13) and its asymptotic form given by the rightmost expression in (6.13).

The generating function of the mean interval $\langle X|L\rangle$ given in (3.32) yields the estimate, in Laplace space,

$$\sum_{L\geq 1} z^L \langle X|L\rangle_{|\mathrm{num}} \approx -\frac{\hat{f}'(s)}{1-\hat{f}(s)} \approx \frac{\theta}{s}, \tag{6.19}$$

hence (see (4.18) in [22])

$$\langle X|L\rangle \approx \frac{\theta}{Z^{\mathrm{td}}(1,L)} \approx \frac{\pi c}{\sin \pi\theta} L^{1-\theta}. \tag{6.20}$$

This result can be recovered by taking the average of the estimate (6.13). It predicts correctly that the product $\langle X|L\rangle\langle N_L\rangle \sim L$. For *Example 1*, (see (4.1)), the computation leads to the exact result (see (2.58) in [22])

$$\langle X|L\rangle = \frac{1}{2Z^{\mathrm{td}}(1,L)} \approx \frac{\sqrt{\pi L}}{2}. \tag{6.21}$$

6.1.3 The longest interval

The study of the statistics of the longest interval for the critical case, including the scaling analysis of $p(k|L)$ for $\theta < 1$, is done in [21,22] (see also [31] for a study of *Example 2*).

For $k > L/2$, in contrast with the case of random allocation models and ZRP where $p(k|L) = q(k|L) = n\pi(k|L)$ (see §2.1.1), the 'enhancement factor' n is now replaced by the factor $1 + \langle N_{L-k}\rangle$ (see (3.49)), which is equal to one for $k = L$, where $p(L|L) = \pi(L|L)$ are equal, see (3.44). Using (6.6) for $\langle N_L\rangle$, and (6.13) for $\pi(k|L)$ in (3.49) allows to recover the universal scaling expression valid for $k > L/2$ (see equation (4.48) in [22]), in the limit where $1 \ll k \sim L$, with $r = k/L$ kept fixed,

$$q(k|L) \approx \frac{1}{L}\frac{A(\theta)}{r^{1+\theta}(1-r)^{1-2\theta}}, \tag{6.22}$$

where $A(\theta)$ is given in (6.6).

Though there is no condensation at criticality, some features are precursors of this phenomenon. For instance, the mean longest interval $\langle X_{\max}\rangle$ scales as L while the typical interval $\langle X|L\rangle$ scales as $L^{1-\theta}$. However, not only $X_{\max} \equiv X^{(1)}$ scales as L but also all the following maxima $X^{(r)}$ $(k = 2, 3, \dots)$ do so [21,22]. Moreover $X_{\max}$ continues to fluctuate when $L \to \infty$ while for genuine condensation as in section 2 above or in section 7 below, its distribution is peaked. Finally, the dominant contribution to the weight of $\pi(k|L)$ comes from values of k less than a small cutoff.

6.2 Distribution $f(k)$ with index $\theta > 1$

For $\theta > 1$, we have, using (3.5) (see (4.73) in [22]),

$$Z^{\mathrm{td}}(1, L) \approx \frac{1}{\langle X\rangle} + \frac{c}{\theta(\theta-1)\langle X\rangle^2}L^{1-\theta}. \tag{6.23}$$

The average value of N_L is obtained by means of (3.14)[4]

$$\langle N_L\rangle \approx \begin{cases} \frac{L}{\langle X\rangle} + \frac{c}{(\theta-1)(2-\theta)\langle X\rangle^2}L^{2-\theta} & 1 < \theta < 2 \\ \frac{L}{\langle X\rangle} + \frac{\mathrm{Var}\,X}{\langle X\rangle^2} & \theta > 2 \end{cases} \tag{6.24}$$

The subleading correction in the second line (i.e., for $\theta > 2$) is given by the correction term of the first line which is now negative and decreasing. The distribution of N_L reads

$$\mathrm{p}_n(L) = \frac{Z_n(L)}{Z^{\mathrm{td}}(1, L)} \approx \langle X\rangle Z_n(L). \tag{6.25}$$

The asymptotic estimate for $\langle X|L\rangle$ is obtained by analysing (3.32), yielding for $\theta > 1$,

$$\langle X|L\rangle \approx \langle X\rangle - \frac{c}{\theta-1}L^{1-\theta}, \tag{6.26}$$

which is the same expression as for a free renewal process [35]. The single interval distribution has the form

$$\pi(k|L) = f(k)\frac{Z^{\mathrm{td}}(1, L-k)}{Z^{\mathrm{td}}(1, L)} \underset{L\to\infty}{\approx} f(k), \tag{6.27}$$

[4] Equation (6.24) corrects the inaccurate expression (4.74) given in [22] for this quantity.

except for $L-k$ finite, where in particular,

$$\pi(L|L) = \frac{f(L)}{Z^{\mathrm{td}}(1,L)} \approx f(L)\langle X\rangle. \tag{6.28}$$

The distribution of the longest interval is analysed in [22] (see also [31] for the case of *Example 2* defined in (4.7)). The result is

$$F(k|L) \approx \mathrm{e}^{-[L/\langle X\rangle](k_0/k)^\theta}. \tag{6.29}$$

where k_0 is related to the tail coefficient by $c = \theta k_0^\theta$. Setting

$$X_{\max} = k_0 \left(\frac{L}{\langle X\rangle}\right)^{1/\theta} Y_L, \tag{6.30}$$

we have, as $L \to \infty$, $Y_L \to Y^F$, with limiting distribution

$$\mathrm{Prob}(Y^F < x) = \mathrm{e}^{-1/x^\theta} \tag{6.31}$$

which is the Fréchet law [45,44]. Therefore

$$\langle X_{\max}\rangle \approx k_0 \left(\frac{L}{\langle X\rangle}\right)^{1/\theta} \underbrace{\langle Y^F\rangle}_{\Gamma(1-1/\theta)}, \tag{6.32}$$

as was already the case for free renewal processes [36].

7 Condensed phase ($w < 1$) for tied-down renewal processes

The aim of this section—central in the present work—is to investigate the statistics of the number of intervals and characterise the fluctuations of the condensate. We start by analysing the large L behaviour of the quantities of interest which are functions of L only (partition function, moments and distribution of N_L). We then investigate the regimes for the distributions of the size of a generic interval, $\pi(k|L)$, and of the longest one, $p(k|L)$. Related material can be found in [33,31,32].

7.1 Asymptotic estimates at large L

Starting from (3.5) and linearising with respect to the singular part, we obtain, when $L \to \infty$, for any value of θ,

$$Z^{\mathrm{td}}(w,L) \approx \frac{w}{(1-w)^2}\frac{c}{L^{1+\theta}} \approx \frac{w}{(1-w)^2} f(L). \tag{7.1}$$

Alternatively, it suffices to notice that, for n fixed and L large, for any subexponential distribution [46],

$$Z_n(L) \approx n f(L), \tag{7.2}$$

as for example in (4.6), which entails

$$Z^{\mathrm{td}}(w,L) = \sum_{n\geq 0} w^n Z_n(L) \approx \sum_{n\geq 0} w^n n f(L), \tag{7.3}$$

hence restores (7.1). Likewise, we find

$$(Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w,L) \approx \frac{2w}{(1-w)^3} f(L). \tag{7.4}$$

If one substitutes (7.1) in the expression of the mean $\langle N_L \rangle$ (3.15), we obtain, the superuniversal result, independent of θ,

$$\langle N_L \rangle \underset{L\to\infty}{\to} \frac{1+w}{1-w}. \tag{7.5}$$

This result can be found alternatively using (3.17) together with (7.1) and (7.4), or else using (7.7) below. More generally, this superuniversality also holds for the asymptotic distribution

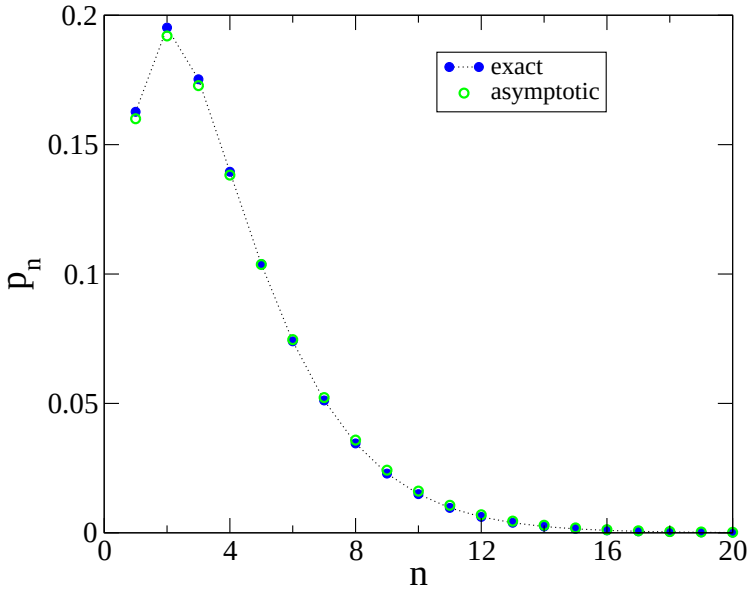

Fig. 9 TDRP: superuniversal asymptotic distribution p_n of the number N_L of intervals in the condensed phase for $w = 0.6$. *Exact* refers to $\mathrm{p}_n(L)$ extracted from (3.12) or (3.20) for $L = 200$ with *Example 1* (see (4.1)), *asymptotic* refers to (7.7). For this value of w, the asymptotic average $\langle N_L \rangle = 4$.

of N_L. Using (3.20), the latter reads

$$\langle v^{N_L} \rangle = \sum_{n\geq 0} v^n \mathrm{p}_n(L) = \frac{Z^{\mathrm{td}}(vw, L)}{Z^{\mathrm{td}}(w, L)} \approx \frac{y(1-w)^2}{(1-vw)^2}, \tag{7.6}$$

hence extracting the coefficient of order n in y of this expression leads to the asymptotic distribution, independent of θ,

$$\mathrm{p}_n(L) \underset{L\to\infty}{\to} \mathrm{p}_n = n(1-w)^2 w^{n-1}. \tag{7.7}$$

This distribution is depicted in figure 9. The same result can be found by noting that

$$\mathrm{p}_n(L) = \frac{w^n Z_n(L)}{Z^{\mathrm{td}}(w, L)} \approx \frac{Z_n(L)}{f(L)}(1-w)^2 w^{n-1} \underset{L\to\infty}{\to} n(1-w)^2 w^{n-1}, \tag{7.8}$$

using (7.2) again. The interpretation of (7.7) is simple: $N_L - 1$ is the sum of two independent geometric random variables (see (10.4)), which represent the fluid on either side of the remaining interval, which is the condensate.

The inverse moment $\langle 1/N_L \rangle$ can be obtained by using (7.7) above,

$$\left\langle \frac{1}{N_L} \right\rangle \to 1 - w. \tag{7.9}$$

As a consequence of (7.9) we have, for any value of θ,

$$\langle X|L \rangle \approx (1-w)L, \tag{7.10}$$

which can also be deduced from the asymptotic analysis of (3.32).

7.2 Regimes for the single interval distribution

For L large, the asymptotic expression of the single interval distribution is obtained by substituting (7.1) in (3.24), leading to

$$\pi(k|L) \underset{L\to\infty}{\approx} (1-w)^2 f(k) \frac{Z^{\mathrm{td}}(w, L-k)}{f(L)}. \tag{7.11}$$

Figure 10 depicts the distribution $\pi(k|L)$ (together with the distribution of the longest interval $p(k|L)$, see §7.3 below), for $L = 60$ and $w = 0.6$ computed from (3.26), with *Example 1* (see (4.1)). As can be seen on this figure, there are three distinct regions for $\pi(k|L)$, namely, from left to right, a downhill region, followed by a long dip region, then an uphill region which accounts for the fluctuations of the condensate.

Let us discuss the behaviour of $\pi(k|L)$ given by (7.11) in each of these regions successively.

1. *Downhill region.* For k finite we have, using again (7.1),

$$\pi(k|L) \approx w f(k). \tag{7.12}$$

Introducing a cutoff Λ, such that $1 \ll \Lambda \ll L$, the weight of this downhill region can be estimated as

$$\sum_{k=1}^{\Lambda} \pi(k|L) \approx \sum_{k=1}^{\Lambda} w f(k) \to \sum_{k=1}^{\infty} w f(k) = w. \tag{7.13}$$

2. *Dip region.* In the dip region, where k and $L-k$ are simultaneously large, setting $k = \lambda L$ in (7.12) ($0 < \lambda < 1$) and using (7.1) yields the estimate

$$\pi(k|L) \approx w f(k) \frac{f(L-k)}{f(L)} \approx \frac{w}{[\lambda(1-\lambda)]^{1+\theta}} \frac{c}{L^{1+\theta}}, \tag{7.14}$$

In this region the distribution is therefore U-shaped: the most probable configurations are those where almost all the particles are located on one of two sites. The dip centred around $k = L/2$ becomes deeper and deeper with L.

The weight of the dip region can be estimated using (7.14), as

$$\sum_{k=\Lambda}^{L-\Lambda} \pi(k|L) \approx L^{-\theta} w\, c \int_{\epsilon}^{1-\epsilon} \frac{\mathrm{d}\lambda}{[\lambda(1-\lambda)]^{1+\theta}}, \tag{7.15}$$

where, for the sake of simplicity, we chose $\Lambda = \epsilon L$. The two downhill and uphill regions are therefore well separated by the dip region, as is conspicuous on figure 10. Note the similarity between (7.15) and (2.52), with the correspondence

$$1 - \frac{1}{n} \hookrightarrow 1 - \left\langle \frac{1}{N_L} \right\rangle = w. \tag{7.16}$$

3. *Uphill region.* The uphill region corresponds to $\ell = L - k$ finite, where (7.11) simplifies into

$$\pi(L-\ell|L) \approx (1-w)^2 Z^{\rm td}(w,\ell). \tag{7.17}$$

The weight of this region can be estimated as

$$\begin{aligned}\sum_{\ell=0}^{\Lambda} \pi(L-\ell|L) \;&\approx\; (1-w)^2 \sum_{\ell=0}^{\Lambda} Z^{\rm td}(w,\ell) \\ &\to (1-w)^2 \sum_{\ell=0}^{\infty} Z^{\rm td}(w,\ell) = 1-w \end{aligned} \tag{7.18}$$

where the last step is obtained by setting $z = 1$ in the expression of the generating function (3.5). The right side of (7.18), $1-w$, is precisely the limiting value of $\langle 1/N_L \rangle$, see (7.9). This result is therefore the analogue of (2.54) with the correspondence given in (7.16).

In view of (7.13) and (7.18) we conclude that the weights of the downhill and uphill regions add to one, in line with the fact that the contribution of the dip region is subdominant, as shown above.

7.3 Regimes for the distribution of the longest interval

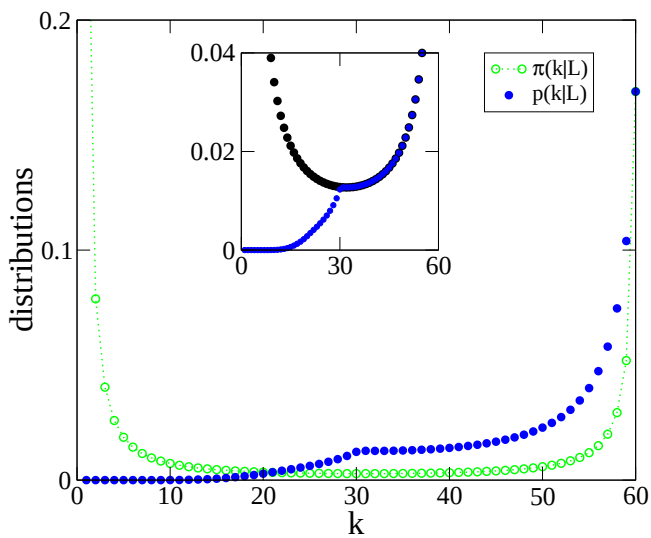

Fig. 10 TDRP: the main graph depicts the exact distributions of the size of a generic interval, $\pi(k|L)$ (in green), and of the longest one, $p(k|L)$ (in blue), in the condensed phase, for $L = 60$ and $w = 0.6$ computed from (3.26) and (3.43), with *Example 1* (see (4.1)). The two curves join at $wf(L)/Z^{\rm td}(w,L) \approx (1-w)^2$ on the y-axis, for $k = 60$. The inset depicts $\pi(k|L)(1+\langle N_{L-k}\rangle)$ (in black) and $p(k|L)$ (in blue). For this value of w, the asymptotic average number $\langle N_L \rangle \to 4$.

As can be seen on figure 10, there are two main regions for the distribution of the maximum, $p(k|L)$. For $k \leq L/2$, the contribution of $p(k|L)$ to the total weight is vanishingly small. The argument is the same as in §2.3.3. Hence we restrict the rest of the discussion to the region $(L/2 < k \leq L)$, where $p(k|L)$ has the simpler expression $q(k|L)$ given by (3.47). Using (7.1), its asymptotic estimate is

$$q(k|L) \underset{L\to\infty}{\approx} (1-w)^2 f(k) \frac{(Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w, L-k)}{f(L)}, \tag{7.19}$$

or, equivalently, for $\ell = L - k \in (0, L/2 - 1)$,

$$q(L-\ell|L) \underset{L\to\infty}{\approx} (1-w)^2 f(L-\ell) \frac{(Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w, \ell)}{f(L)}. \tag{7.20}$$

Let us note that the ratio

$$r(\ell) = \frac{q(L-\ell|L)}{\pi(L-\ell|L)} = \frac{(Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w, \ell)}{Z^{\mathrm{td}}(w, \ell)} = 1 + \langle N_\ell \rangle \tag{7.21}$$

is an increasing function of ℓ, with first values

$$r(0) = 1, \quad r(1) = 2, \quad r(2) = \frac{3wf(1)^2 + 2f(2)}{wf(1)^2 + f(2)}, \tag{7.22}$$

and so on, reaching the limit, for large ℓ,

$$r(\ell) \to \frac{2}{1-w}. \tag{7.23}$$

The inset in figure 10 depicts the prediction $q(k|L) = \pi(k|L)(1 + \langle N_{L-k} \rangle)$ (see (3.49)) which coincides perfectly with $p(k|L)$ in the second half $k > L/2$. This graph is very similar to the inset in figure 5.

We now discuss the behaviour of $q(L-\ell|L)$ in the two regions of interest.

1. For ℓ finite, we have, see (7.20),

$$q(L-\ell|L) \approx (1-w)^2 (Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w, \ell). \tag{7.24}$$

   In particular for $\ell = 0$,

$$p(L|L) = q(L|L) = \pi(L|L) = \mathrm{Prob}(N_L = 1) \approx (1-w)^2. \tag{7.25}$$

   If $1 \ll \ell \ll L$, (7.24) simplifies to

$$q(L-\ell|L) \approx \frac{2w}{1-w} f(\ell). \tag{7.26}$$

2. If ℓ et $L - \ell$ are simultaneously large, with $\ell = \lambda L$, we have

$$q(L-\ell|L) \approx \frac{2w}{1-w} f(\ell) \frac{f(L-\ell)}{f(L)} \approx \frac{2w}{1-w} \frac{1}{[\lambda(1-\lambda)]^{1+\theta}} \frac{c}{L^{1+\theta}}, \tag{7.27}$$

   which is proportional to (7.14), with ratio $2/(1-w)$. Note the similarity of (7.27) with (2.56) with the correspondence

$$n - 1 \hookrightarrow \langle N_L \rangle - 1 = \frac{2w}{1-w}. \tag{7.28}$$

The weight under the peak of the condensate tends to unity,

$$\sum_{\ell=0}^{\Lambda} q(L-\ell|L) \approx (1-w)^2 \sum_{\ell=0}^{\Lambda} (Z^{\rm td} \star Z^{\rm td})(w,\ell)$$
$$\to (1-w)^2 \sum_{\ell=0}^{\infty} (Z^{\rm td} \star Z^{\rm td})(w,\ell) = 1, \tag{7.29}$$

by a computation similar to (7.18), using now the fact that the last sum is equal to $1/(1-w)^2$, as can be seen by setting $z=1$ in the expression of the generating function $\tilde{Z}^{\rm td}(w,z)^2$.

We can further analyse the fluctuations of the condensate by considering the width of the peak,

$$L - \langle X_{\max} \rangle = \sum_{\ell=0}^{L-1} \ell\, q(L-\ell|L) \approx \sum_{\ell=0}^{L/2} \ell\, q(L-\ell|L). \tag{7.30}$$

The dominant contribution to this sum depends on whether θ is smaller or larger than one.

• If $\theta < 1$, the dominant contribution comes from (7.27), i.e., for ℓ comparable to L,

$$L - \langle X_{\max} \rangle \approx \frac{2wc}{1-w} L^{1-\theta} \int_0^{1/2} {\rm d}\lambda \frac{\lambda}{[(\lambda(1-\lambda)]^{1+\theta}}$$
$$\approx \frac{2wc}{1-w} {\rm B}_{\frac{1}{2}}\Big(1-\theta, -\theta\Big) L^{1-\theta}, \tag{7.31}$$

where ${\rm B}(\cdot)$ is the incomplete beta function, which, for example, is equal to 2 for $\theta = 1/2$.

• If $\theta > 1$, the main contribution comes from (7.24),

$$L - \langle X_{\max} \rangle \approx (1-w)^2 \sum_{\ell=0}^{\Lambda} \ell (Z^{\rm td} \star Z^{\rm td})(w,\ell)$$
$$\to (1-w)^2 \sum_{\ell=0}^{\infty} \ell (Z^{\rm td} \star Z^{\rm td})(w,\ell) \to \frac{2w}{1-w} \langle X \rangle, \tag{7.32}$$

using (7.4). This last result (7.32) has a simple interpretation. It says that the correction $L - \langle X_{\max} \rangle$ is made of $\langle N_L \rangle - 1 = 2w/(1-w)$ intervals, of mean length $\langle X \rangle$. It is therefore the perfect parallel of the result (2.48). Likewise, (7.31) is the perfect parallel of (2.46). In the present case the expression in the right side of this equation is proportional to $\langle N_L \rangle - 1$ times the critical mean interval $\langle X|L \rangle$, given in (6.20).

7.4 Discussion

In view of the above analysis the following picture emerges. The 'contrast' between the dip and the condensate region increases with L, i.e., the dip centred around $L/2$ becomes deeper and deeper as $L^{-1-\theta}$ (see (7.14)) relatively to the height of the peak, which is of order one. An estimate of the contribution of the condensate to the total weight can thus be operationally obtained by summing $\pi(L-\ell|L)$ for ℓ in the layer $(0,\Lambda)$. This sum is asymptotically equal to $1-w$, according to (7.18), which turns out to be also the asymptotic estimate of $\langle 1/N_L \rangle$. The interpretation of this result is clear. When X_1 is larger than $L/2$, this interval is necessarily

the longest one, i.e., $X_1 = X_{\max}$. Furthermore, since this interval is chosen amongst N_L intervals, we expect that, in average,

$$\sum_{\ell=0}^{\Lambda} \pi(L-\ell|L) \approx \underbrace{\left\langle \frac{1}{N_L} \right\rangle}_{1-w} \underbrace{\sum_{\ell=0}^{\Lambda} p(L-\ell|L)}_{\approx 1}, \tag{7.33}$$

But the sum on the right side, namely the weight of $X_{\max}$ in the same layer is asymptotically equal to one. This simple heuristic reasoning therefore recovers (7.18).

In the condensed phase ($w < 1, L \to \infty$) the number of intervals is finite and fluctuates around its mean, which is a superuniversal constant independent of θ. This situation is akin to the case of random allocation models and ZRP, when $L \to \infty$ and n is kept fixed. Note that, for the latter, results were independent of the value of θ, too. In both situations condensation is total, the condensed fraction is asymptotically equal to unity.

Table 1 summarises the results found in sections 5, 6 and 7, which demonstrate a large degree of universality.

Table 1 Dominant asymptotic behaviours at large L for tied-down renewal processes with power-law distribution (1.1) for $f(k)$, in the different phases.

	disordered	critical $\theta < 1$	critical $\theta > 1$	condensed
$\langle N_L \rangle$	$\frac{L}{\langle X\|L \rangle}$	L^{θ}	$\frac{L}{\langle X \rangle}$	$\frac{1+w}{1-w}$
$\langle X\|L \rangle$	constant	$L^{1-\theta}$	$\langle X \rangle$	$(1-w)L$
$\langle X_{\max} \rangle$	$\ln L$	L	$L^{1/\theta}$	L
$Z^{\mathrm{td}}(w,L)$	$\mathrm{e}^{L/\xi}$	$L^{\theta-1}$	$\frac{1}{\langle X \rangle}$	$L^{-1-\theta}$

8 General statements on free renewal processes

We now turn to the case of free renewal processes. The random number N_L of intervals up to L is defined through the condition (1.4), $S_{N_L} < L < S_{N_L+1}$. The size of the current (unfinished) interval—named the *backward recurrence time* in renewal theory—is denoted by $B_L = L - S_{N_L}$, see figure 2. As will appear shortly, free renewal processes are more complicated to analyse than TDRP, essentially because now there are two kinds of intervals to consider, the intervals X_i on one hand, and the last unfinished interval B_L, on the other hand.

In the present section, $f(k) = \mathrm{Prob}(X = k)$ is any arbitrary distribution of the positive random variable X. Later on, in sections 9 and 10, $f(k)$ will obey the form (1.1).

8.1 Joint distribution

As for TDRP a weight w is attached to each interval. The joint probability of the configuration $\{X_1 = k_1, \ldots, X_{N_L} = k_n, B_L = b, N_L = n\}$, reads

$$\begin{aligned} p(k_1,\ldots,k_n,b,n|L) &= \mathrm{Prob}(\{X_i = k_i\}, B_L = b, N_L = n) \\ &= \frac{1}{Z^{\mathrm{f}}(w,L)} w^n f(k_1)\ldots f(k_n)\, g(b)\, \delta\Big(\sum_{i=1}^{n} k_i + b, L\Big), \end{aligned} \tag{8.1}$$

where $g(b)$ is the tail probability (or complementary distribution function) defined in (6.4),

$$g(b) = \mathrm{Prob}(X > b) = \sum_{k>b} f(k) = f(b+1) + f(b+2) + \cdots \tag{8.2}$$

As mentioned in section 3, since the summands X_i have the interpretation of the sizes of the intervals, we take $f(0) = 0$. For $n = 0$,

$$p(\{\}, b, 0|L)_{|\mathrm{num}} = g(b)\delta(b, L), \tag{8.3}$$

corresponding to the event of no renewal occurring between 0 and L, i.e., $B_L = L$, and where $\{\}$ means empty. The generating function of $g(b)$ is

$$\tilde{g}(z) = \sum_{b\geq 0} z^b g(b) = \frac{1}{1-z} - \sum_{b\geq 1} z^b \sum_{k=1}^{b} f(k) = \frac{1-\tilde{f}(z)}{1-z}, \tag{8.4}$$

with $\tilde{g}(0) = g(0) = 1$.

The denominator of (8.1) is the free partition function, obtained by summing on n, on the k_i and on b,

$$\begin{aligned}
Z^{\mathrm{f}}(w, L) &= \sum_{n\geq 0}\sum_{b\geq 0}\sum_{\{k_i\}} p(k_1, \ldots, k_n, b, n|L) \\
&= \sum_{n\geq 0} w^n \sum_{b\geq 0}\sum_{\{k_i\}} f(k_1)\cdots f(k_n) g(b)\delta\big(\sum_{i=1}^{n} k_i + b, L\big) \\
&= \sum_{b\geq 0} g(b)\Big[\delta(b, L) + \sum_{k_1} w f(k_1)\delta(k_1 + b, L) \\
&+ \sum_{k_1,k_2} w^2 f(k_1) f(k_2)\delta(k_1 + k_2 + b, L) + \ldots\Big] \\
&= \underbrace{g(L)}_{n=0} + \underbrace{w\, g \star f}_{n=1} + \underbrace{w^2\, g \star f \star f}_{n=2} + \cdots = \sum_{n\geq 0} w^n \big(g \star (f\star)^n\big)(L).
\end{aligned} \tag{8.5}$$

For instance,

$$Z^{\mathrm{f}}(w, 0) = 1, \qquad Z^{\mathrm{f}}(w, 1) = g(1) + w f(1), \tag{8.6}$$

$$Z^{\mathrm{f}}(w, 2) = g(2) + w f(2) + w f(1) g(1) + w^2 f(1)^2, \tag{8.7}$$

and so on. From (8.5) we have

$$\tilde{Z}^{\mathrm{f}}(w, z) = \sum_{L\geq 0} z^L Z^{\mathrm{f}}(w, L) = \sum_{n\geq 0} \big(w\tilde{f}(z)\big)^n \tilde{g}(z) = \frac{\tilde{g}(z)}{1 - w\tilde{f}(z)}. \tag{8.8}$$

Let us note that

$$Z^{\mathrm{f}}(w, L) = \sum_{n\geq 0} w^n \sum_{b\geq 0} \mathrm{Prob}(S_n = L - b) g(b) = (g \star Z^{\mathrm{td}})(w, L), \tag{8.9}$$

where $Z^{\mathrm{td}}(w, L)$ is the partition function (3.3) for TDRP.

The case of usual (unweighted) free renewal processes is recovered by setting $w = 1$ in these expressions. This yields $Z^{\mathrm{f}}(1, L) = 1$, as can be seen from (8.8), and the joint probability distribution (8.1) simplifies accordingly.

8.2 Distribution of the number of intervals

As for TDRP we denote this distribution as[5]

$$\mathrm{p}_n(L) = \mathrm{Prob}(N_L = n). \tag{8.10}$$

We read on the successive terms of $Z^{\mathrm{f}}(w, L)$ that

$$\mathrm{p}_0(L) = \frac{g(L)}{Z^{\mathrm{f}}(w, L)} = \mathrm{Prob}(B_L = L), \quad \mathrm{p}_1(L) = \frac{w(g \star f)(L)}{Z^{\mathrm{f}}(w, L)}, \tag{8.11}$$

and so on. More generally,

$$\mathrm{p}_n(L) = \frac{w^n \big(g \star (f\star)^n\big)(L)}{Z^{\mathrm{f}}(w, L)} = \frac{w^n (g \star Z_n)(L)}{(g \star Z^{\mathrm{td}})(w, L)} \tag{8.12}$$

Summing (8.1) on b and on the k_i, and taking the generating function with respect to L yields

$$\sum_{L\geq 0} z^L \mathrm{p}_n(L)_{|\mathrm{num}} = \Big(w\tilde{f}(z)\Big)^n \tilde{g}(z), \tag{8.13}$$

to be compared to (8.8). Therefore

$$\sum_{L\geq 0} z^L \langle N_L\rangle_{|\mathrm{num}} = \frac{w\tilde{f}(z)\tilde{g}(z)}{(1 - w\tilde{f}(z))^2} = w\frac{\mathrm{d}}{\mathrm{d}w}\tilde{Z}^{\mathrm{f}}(w, z), \tag{8.14}$$

and

$$\langle N_L\rangle = w\frac{\mathrm{d}\ln Z^{\mathrm{f}}(w, L)}{\mathrm{d}w}, \tag{8.15}$$

as for TDRP, see (3.15). More generally,

$$\sum_{L\geq 0} z^L \left\langle v^{N_L}\right\rangle_{|\mathrm{num}} = \sum_{n\geq 0} v^n \Big(w\tilde{f}(z)\Big)^n \tilde{g}(z) = \tilde{Z}^{\mathrm{f}}(vw, z), \tag{8.16}$$

so we obtain, as for TDRP, see (3.20),

$$\left\langle v^{N_L}\right\rangle = \sum_{n\geq 0} v^n \mathrm{p}_n(L) = \frac{Z^{\mathrm{f}}(vw, L)}{Z^{\mathrm{f}}(w, L)}. \tag{8.17}$$

Finally, comparing (8.14) to (3.14), we note the relationship between the free and tied-down cases,

$$\langle N_L^{\mathrm{f}}\rangle_{|\mathrm{num}} = \sum_{k=0}^{L} g(k)\langle N_{L-k}^{\mathrm{td}}\rangle_{|\mathrm{num}} \tag{8.18}$$

hence using (3.17), we have

$$\langle N_L^{\mathrm{f}}\rangle_{|\mathrm{num}} = (g \star Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w, L) - Z^{\mathrm{f}}(w, L) \tag{8.19}$$

or, equivalently,

$$(g \star Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w, L) = Z^{\mathrm{f}}(w, L)(1 + \langle N_L\rangle). \tag{8.20}$$

[5] Whenever no ambiguity arises, we use the same notations for the observables of the tied-down and free renewal processes. Otherwise, when necessary, we add a superscript, as e.g., for Z^{td}, Z^{f} or in (8.18).

8.3 Distribution of S_{N_L}

We recall that this quantity is the sum of the N_L intervals before L, see (1.3) and figure 2. By definition,

$$\mathrm{Prob}(S_{N_L}=j)=\langle\delta(S_{N_L},j)\rangle=\sum_{n\geq 0}\sum_{b\geq 0}\sum_{\{k_i\}}\delta(S_{N_L},j)p(k_1,\ldots,k_n,b,n|L)$$

Thus, using (8.1), we have

$$\sum_{L\geq 0}z^L\sum_{j=0}^{L}x^j\,\mathrm{Prob}(S_{N_L}=j)_{|\mathrm{num}}=\frac{\tilde{g}(z)}{1-w\tilde{f}(xz)},\tag{8.21}$$

which generalises the expression for this quantity when $w=1$ [35]. By derivation with respect to x then setting $x=1$, leads to

$$\sum_{L\geq 0}z^L\langle S_{N_L}\rangle_{|\mathrm{num}}=\frac{wz\tilde{f}'(z)\tilde{g}(z)}{(1-w\tilde{f}(z))^2},\tag{8.22}$$

whose summation with (8.30) below leads to the equality

$$\sum_{L\geq 0}z^L\big(\langle S_{N_L}\rangle_{|\mathrm{num}}+\langle B_L\rangle_{|\mathrm{num}}\big)=z\frac{\mathrm{d}\tilde{Z}^{\mathrm{f}}(w,z)}{\mathrm{d}z},\tag{8.23}$$

which expresses the sum rule

$$\langle S_{N_L}\rangle+\langle B_L\rangle=L.\tag{8.24}$$

The asymptotic behaviours of these quantities for $w=1$ are simple [35]. If $\theta<1$, then $\langle S_{N_L}\rangle\approx\theta L$, $\langle B_L\rangle\approx(1-\theta)L$. If $1<\theta<2$, then $\langle S_{N_L}\rangle\approx L-cL^{2-\theta}/[\theta(2-\theta)\langle X\rangle]$, and $\langle B_L\rangle$ follows by difference.

8.4 Distribution of B_L

The distribution of B_L is obtained by summing (8.1) on the k_i and on n

$$\mathrm{Prob}(B_L=b)=\frac{1}{Z^{\mathrm{f}}(w,L)}g(b)\sum_{n\geq 0}w^n\sum_{\{k_i\geq 1\}}f(k_1)\cdots f(k_n)\delta\Big(\sum_{i=1}^{n}k_i+b,L\Big).\tag{8.25}$$

This entails that

$$\mathrm{Prob}(B_L=b)_{|\mathrm{num}}=g(b)\sum_{n\geq 0}w^n\,\mathrm{Prob}(S_n=L-b)=g(b)Z^{\mathrm{td}}(w,L-b),\tag{8.26}$$

hence, using (8.9)

$$\mathrm{Prob}(B_L=b)=\frac{g(b)Z^{\mathrm{td}}(w,L-b)}{(g\star Z^{\mathrm{td}})(w,L)}.\tag{8.27}$$

The generating function with respect to L of (8.26) reads

$$\sum_{L\geq 0}z^L\,\mathrm{Prob}(B_L=b)_{|\mathrm{num}}=\frac{z^bg(b)}{1-w\tilde{f}(z)},\tag{8.28}$$

which summed upon b gives $\tilde{Z}^{\mathrm{f}}(w,z)$ back. Taking now the generating function of (8.28) with respect to b yields

$$\sum_{L\geq 0} z^L \sum_{b\geq 0} y^b \,\mathrm{Prob}(B_L = b)_{|\mathrm{num}} = \frac{\tilde{g}(yz)}{1-w\tilde{f}(z)}. \tag{8.29}$$

The mean $\langle B_L\rangle$ ensues by taking the derivative of the right side of this equation with respect to y and setting y to one,

$$\sum_{L\geq 0} z^L \langle B_L\rangle_{|\mathrm{num}} = \frac{z\tilde{g}'(z)}{1-w\tilde{f}(z)}. \tag{8.30}$$

8.5 Single interval distribution

As for TDRP, the single interval distribution,

$$\pi(k|L) = \langle \delta(X_1,k)\rangle, \tag{8.31}$$

is obtained by summing $p(k_1,\ldots,k_n,b,n|L)$ on $k_1,\ldots,k_n$, b and $n\geq 1$

$$\begin{aligned}
\pi(k|L)_{|\mathrm{num}} &= \sum_{b\geq 0} g(b)\Big[\sum_{k_1}\delta(k_1,k)wf(k_1)\delta(k_1+b,L)\\
&\quad+\sum_{k_1,k_2}\delta(k_1,k)w^2f(k_1)f(k_2)\delta(k_1+k_2+b,L)+\cdots\Big]\\
&= \sum_{b\geq 0} g(b)\Big[wf(k)\delta(k+b,L)+\sum_{k_2}w^2f(k)f(k_2)\delta(k+k_2+b,L)+\ldots\Big]\\
&= wf(k)Z^{\mathrm{f}}(w,L-k).
\end{aligned} \tag{8.32}$$

So

$$\pi(k|L) = \frac{wf(k)Z^{\mathrm{f}}(w,L-k)}{Z^{\mathrm{f}}(w,L)}. \tag{8.33}$$

The generating function of the numerator is therefore

$$\sum_{L\geq 0} z^L \pi(k|L)_{|\mathrm{num}} = wz^k f(k)\tilde{Z}^{\mathrm{f}}(w,z) = \frac{wz^k f(k)\tilde{g}(z)}{1-w\tilde{f}(z)}, \tag{8.34}$$

to be compared to (3.32). Though (8.33) is formally identical to (3.26) the normalisations of these two distributions are different. Indeed,

$$\sum_{L\geq 0} z^L \sum_{k=1}^{L}\pi(k|L)_{|\mathrm{num}} = w\tilde{f}(z)\tilde{Z}^{\mathrm{f}}(w,z) = \frac{w\tilde{f}(z)\tilde{g}(z)}{1-w\tilde{f}(z)} = \tilde{Z}^{\mathrm{f}}(w,z)-\tilde{g}(z), \tag{8.35}$$

which means that

$$1-\sum_{k=1}^{L}\pi(k|L) = \frac{g(L)}{Z^{\mathrm{f}}(w,L)} = \mathrm{p}_0(L) = \mathrm{Prob}(B_L = L). \tag{8.36}$$

In other words

$$\sum_{k=1}^{L}\mathrm{Prob}(X=k|L)+\mathrm{Prob}(B_L=L) = 1. \tag{8.37}$$

The distribution $\pi(k|L)$ is thus defective. The recursion relation for $Z^{\mathrm{f}}(w,L)$ follows from (8.32) and (8.36)

$$Z^{\mathrm{f}}(w,L) = g(L) + \sum_{k=1}^{L} w f(k) Z^{\mathrm{f}}(w, L-k). \tag{8.38}$$

At the end point, $k = L$, we have

$$\pi(L|L) = \frac{w f(L)}{Z^{\mathrm{f}}(w,L)}, \tag{8.39}$$

which corresponds to the event $\{X_1 = L, B_L = 0, N_L = 1\}$, and which is formally the same as (3.25).

Both sides of (8.38) are equal to unity if $w = 1$. Note that the computation of $\pi(k|L)$ for $w = 1$ is given in [35], yielding $\pi(k|L) = f(k)$ with $k \leq L$, which shows that this distribution is already defective for $w = 1$, with $\sum_{k=1}^{L} \pi(k|L) = 1 - g(L)$.

8.6 Mean interval $\langle X|L\rangle$

We proceed as in §3.4. The mean interval is, by definition,

$$\langle X|L\rangle = \sum_{k\geq 1} k\pi(k|L). \tag{8.40}$$

Multiplying (8.34) by k and summing upon k yields

$$\sum_{L\geq 1} z^L \langle X|L\rangle_{|\mathrm{num}} = \frac{wz\tilde{f}'(z)\tilde{g}(z)}{1 - w\tilde{f}(z)}, \tag{8.41}$$

to be compared to (3.32) for TDRP.

8.7 The longest interval

In the present case the longest interval is defined as

$$X_{\max} = \max(X_1, X_2, \ldots, B_L). \tag{8.42}$$

Its distribution function is

$$F(k|L) = \mathrm{Prob}(X_{\max} \leq k|L) = \sum_{n\geq 0} \sum_{b=0}^{k} \sum_{k_1=1}^{k} \cdots \sum_{k_n=1}^{k} p(\{k_i\}, b, n|L) = \frac{F(k|L)_{|\mathrm{num}}}{Z^{\mathrm{f}}(w,L)}, \tag{8.43}$$

with initial value

$$F(k|0)_{|\mathrm{num}} = 1. \tag{8.44}$$

As for TDRP, $F(L|L)_{|\mathrm{num}} = Z^{\mathrm{f}}(w,L)$, hence $F(L|L) = 1$. The generating function of the numerator is

$$\begin{aligned}\sum_{L\geq 0} z^L F(k|L)_{|\mathrm{num}} &= \tilde{g}(z,k)\left(1 + \sum_{n\geq 1} \prod_{i=1}^{n} \Big(\sum_{k_i=1}^{k} w f(k_i) z^{k_i} \Big)\right) \\ &= \tilde{g}(z,k)\left(1 + \sum_{n\geq 1} \Big(w\tilde{f}(z,k)\Big)^n\right) = \frac{\tilde{g}(z,k)}{1 - w\tilde{f}(z,k)},\end{aligned} \tag{8.45}$$

where

$$\tilde{f}(z,k) = \sum_{j=1}^{k} z^j f(j), \qquad \tilde{g}(z,k) = \sum_{b=0}^{k} z^b g(b), \tag{8.46}$$

are related by

$$1 - \tilde{f}(z,k) = z^{k+1} g(k) + (1-z)\tilde{g}(z,k). \tag{8.47}$$

Note that

$$\begin{aligned} F^{\mathrm{f}}(k|L)_{|\mathrm{num}} &= \sum_{b=0}^{k} g(b) F^{\mathrm{td}}(k|L-b)_{|\mathrm{num}} \\ &= \sum_{b=0}^{k} g(b) \sum_{n=0}^{L} w^n F_n(k|L-b)_{|\mathrm{num}}, \end{aligned} \tag{8.48}$$

where we used (3.38) in the last step.

The distribution of $X_{\max}$ is given by the difference

$$p(k|L) = \mathrm{Prob}(X_{\max} = k) = F(k|L) - F(k-1|L), \tag{8.49}$$

where $F(0|L) = \delta(L,0)$, with generating function

$$\begin{aligned} \sum_{L\geq 0} z^L p(k|L)_{|\mathrm{num}} &= \tilde{g}(z,k)\Big(\frac{1}{1-w\tilde{f}(z,k)} - \frac{1}{1-w\tilde{f}(z,k-1)}\Big) \\ &= \frac{wz^k f(k)\tilde{g}(z,k)}{[1-w\tilde{f}(z,k)][1-w\tilde{f}(z,k-1)]}. \end{aligned} \tag{8.50}$$

At the end point, $k = L$, we have

$$p(L|L) = \frac{wf(L) + g(L)}{Z^{\mathrm{f}}(w,L)} = \pi(L|L) + \mathrm{Prob}(B_L = L), \tag{8.51}$$

where the last two terms correspond respectively to the events $\{X_1 = L, B_L = 0, N_L = 1\}$, cf (8.39) and $\{B_L = L, N_L = 0\}$, cf (8.11). The mean is given by the sum

$$\langle X_{\max}\rangle = \sum_{k=0}^{L} \Big(1 - F(k|L)\Big) = L - \sum_{k=1}^{L-1} F(k|L), \tag{8.52}$$

which implies a relation between the generating functions

$$\sum_{L\geq 0} z^L \langle X_{\max}\rangle_{|\mathrm{num}} = \sum_{k\geq 0} \Big(\tilde{Z}^{\mathrm{f}}(w,z) - \tilde{F}(k,z)_{|\mathrm{num}}\Big) \tag{8.53}$$

$$= \sum_{k\geq 0} \Big(\frac{\tilde{g}(z)}{1-w\tilde{f}(z)} - \frac{\tilde{g}(z,k)}{1-w\tilde{f}(z,k)}\Big), \tag{8.54}$$

where $\tilde{F}(k,z)_{|\mathrm{num}}$ is given by (8.45).

Denoting again the restriction of $p(k|L)$ to $k > L/2$ by $q(k|L)$, we can obtain an expression of this quantity by a similar reasoning as done for (3.47) in §3.5. We thus obtain

$$\begin{aligned} q(k|L) &= \mathrm{Prob}(B_L = k) + \frac{wf(k)}{Z^{\mathrm{f}}(w,L)} \sum_{b=0}^{L-k} g(b)(Z^{\mathrm{td}} \star Z^{\mathrm{td}})(w, L-k-b) \\ &= \mathrm{Prob}(B_L = k) + \frac{wf(k)\big(g \star Z^{\mathrm{td}} \star Z^{\mathrm{td}}\big)(w, L-k)}{Z^{\mathrm{f}}(w,L)}, \end{aligned} \tag{8.55}$$

expressing the fact that the longest interval can be either B_L or a generic interval X_i. Equivalently, using (8.20), this reads

$$q(k|L) = \mathrm{Prob}(B_L = k) + \frac{wf(k)Z^{\mathrm{f}}(w, L-k)}{Z^{\mathrm{f}}(w, L)}(1 + \langle N_{L-k}\rangle) \tag{8.56}$$

or else

$$q(k|L) = \mathrm{Prob}(B_L = k) + \pi(k|L)(1 + \langle N_{L-k}\rangle), \tag{8.57}$$

to be compared to (3.48) and (3.49). For $k = L$, (8.51) is recovered.

Remarks

1. We can prove (8.55) otherwise. We start from the first line of (8.48) and take the discrete derivative with respect to k of both sides, which yields

$$\Delta_k F^{\mathrm{f}}(k|L)_{|\mathrm{num}} = g(k)F^{\mathrm{td}}(k|L-k)_{|\mathrm{num}} + \sum_{b=0}^{k} g(b)\Delta_k F^{\mathrm{td}}(k|L-b)_{|\mathrm{num}}. \tag{8.58}$$

If $k > L/2$, using (3.54), we recognise the numerator of the first term of (8.55) in the first term of this equation. Likewise, if $k > L/2$, it can be shown that the second term of the above equation is equal to the numerator of the second term of (8.55).

2. From (3.54) we infer that, if $k > L/2$,

$$F^{\mathrm{f}}(k|L-k)_{|\mathrm{num}} = Z^{\mathrm{f}}(w, L-k). \tag{8.59}$$

Lastly, consider the probability Q_L that the last unfinished interval is the longest one, that is

$$Q_L = \mathrm{Prob}(B_L \geq \max(X_1, \ldots, X_{N_L})) = \sum_{n\geq 0}\sum_{b\geq 0}\sum_{k_1=1}^{b}\cdots\sum_{k_n=1}^{b} p(\{k_i\}, b, n|L). \tag{8.60}$$

The generating function with respect to L of its numerator reads

$$\sum_{L\geq 0} z^L Q_{L|\mathrm{num}} = \sum_{b\geq 0}\frac{z^b g(b)}{1 - w\tilde{f}(z, b)}, \tag{8.61}$$

hence

$$\tilde{Z}^{\mathrm{f}}(w, z) - \sum_{L\geq 0} z^L Q_{L|\mathrm{num}} = \frac{\tilde{g}(z)}{1 - w\tilde{f}} - \sum_{b\geq 0}\frac{z^b g(b)}{1 - w\tilde{f}(z, b)} \tag{8.62}$$

$$= \sum_{b\geq 0} z^b g(b)\Big(\frac{1}{1 - w\tilde{f}} - \frac{1}{1 - w\tilde{f}(z, b)}\Big). \tag{8.63}$$

9 Critical regime ($w = 1$) for free renewal processes

In this section and in the following one (section 10) we specialise the discussion to the case of a subexponential distribution $f(k) = \mathrm{Prob}(X = k)$ with asymptotic power-law decay (1.1).

The critical regime is thoroughly described in [35,36] and builds upon previous studies [17,47–49]. The results are summarised in table 2, which also presents the main outcomes for the disordered regime ($w > 1$).

The initial analysis of the distribution of the longest interval is due to Lamperti [49]. Let us just recover the universal asymptotic expression of $q(k|L)$, for $1 \ll k \sim L$, with $r = k/L$ fixed, when $\theta < 1$, given in [49],

$$q(k|L) \approx \frac{1}{L}\frac{\sin \pi\theta}{\pi}\frac{1}{r^{1+\theta}(1-r)^{1-\theta}}. \tag{9.1}$$

This result can be simply inferred from (8.57). For the first term of (8.57), we obtain for L large, using (6.5),

$$\mathrm{Prob}(B_L = k) = g(k)Z^{\mathrm{td}}(1, L-k) \underset{L\to\infty}{\approx} \frac{\sin \pi\theta}{\pi}\frac{1}{k^{\theta}(L-k)^{1-\theta}}, \tag{9.2}$$

i.e., an arcsine law in the variable $r = k/L$, which is a well-known result [47,17,35]. For the second term of (8.57), we need

$$\langle N_{L-k}\rangle \underset{L\to\infty}{\approx} \frac{\sin \pi\theta}{\pi c}(L-k)^{\theta}, \tag{9.3}$$

which is obtained using (8.14). Adding the two terms of (8.57) yields (9.1).

10 Condensed phase ($w < 1$) for free renewal processes

We now focus on the case of most interest, namely the condensed phase, $w < 1$, for subexponential distributions (1.1). As in section 7, we investigate the statistics of the number of intervals, the single-interval distribution and we characterise the fluctuations of the condensate. We also address the statistics of the last interval B_L.

10.1 Asymptotic estimates at large L

The asymptotic analysis of (8.8) yields, for large L,

$$Z^{\mathrm{f}}(w, L) \approx \frac{g(L)}{1-w}\Big(1 - \frac{wc\Gamma(1-\theta)^2}{(1-w)\Gamma(1-2\theta)\theta L^{\theta}}\Big), \quad \theta < 1 \tag{10.1}$$

$$Z^{\mathrm{f}}(w, L) \approx \frac{g(L)}{1-w}\Big(1 + \frac{2w\theta\langle X\rangle}{(1-w)L}\Big), \quad \theta > 1. \tag{10.2}$$

where $g(\cdot)$ is defined in (6.4). As a consequence, (8.17) yields

$$\big\langle v^{N_L}\big\rangle \underset{L\to\infty}{\to} \frac{1-w}{1-vw}, \tag{10.3}$$

leading asymptotically to a geometric distribution for N_L,

$$\mathrm{Prob}(N_L = n) \to \mathrm{p}_n = (1-w)w^n, \tag{10.4}$$

independent of θ, from which entails, in the same limit,

$$\langle N_L \rangle \to \frac{w}{1-w}. \tag{10.5}$$

For L large but finite, we find, using (8.15) and (10.1) or (10.2),

$$\langle N_L \rangle \approx \frac{w}{1-w}\Big(1 - \frac{c}{1-w}\frac{\theta\,\Gamma(1-\theta)^2}{\Gamma(1-2\theta)L^\theta}\Big), \quad \theta < 1.$$

$$\langle N_L \rangle \approx \frac{w}{1-w}\Big(1 + \frac{2\theta\langle X\rangle}{(1-w)L}\Big), \quad \theta > 1. \tag{10.6}$$

The asymptotic estimate of the mean interval can be obtained from the analysis of (8.41),

$$\langle X|L \rangle \approx \frac{wc\Gamma(1-\theta)^2}{\Gamma(2-2\theta)}L^{1-\theta}, \quad \theta < 1, \tag{10.7}$$

$$\langle X|L \rangle \approx w(1+\theta)\langle X\rangle, \quad \theta > 1. \tag{10.8}$$

These expressions can also be obtained by means of the marginal distribution $\pi(k|L)$, see below. Likewise, the asymptotic estimate of the mean sum can be obtained from the analysis of (8.22),

$$\langle S_{N_L} \rangle \approx \frac{wc}{1-w}\frac{\Gamma(1-\theta)^2}{\Gamma(2-2\theta)}L^{1-\theta}, \qquad \theta < 1, \tag{10.9}$$

$$\langle S_{N_L} \rangle \approx \frac{w}{1-w}(1+\theta)\langle X\rangle, \qquad \theta > 1, \tag{10.10}$$

hence

$$\langle B_L \rangle \approx L - \frac{wc}{1-w}\frac{\Gamma(1-\theta)^2}{\Gamma(2-2\theta)}L^{1-\theta}, \qquad \theta < 1, \tag{10.11}$$

$$\langle B_L \rangle \approx L - \frac{w}{1-w}(1+\theta)\langle X\rangle, \qquad \theta > 1. \tag{10.12}$$

10.2 Regimes for the single interval distribution

We proceed as was done for TDRP (see section 7.2). Using (10.1) we have the estimate, at large L,

$$\pi(k|L) \underset{L\to\infty}{\approx} w(1-w)\frac{f(k)Z^{\mathrm{f}}(w, L-k)}{g(L)}. \tag{10.13}$$

Figure 11 depicts the distribution $\pi(k|L)$ (together with the distribution $\mathrm{Prob}(B_L = k)$, see §10.3 below), for $L = 60$ and $w = 0.8$ computed with *Example 1* (see (4.1)). As can be seen on this figure, there are three distinct regions for $\pi(k|L)$, that we consider in turn.

1. *Downhill region.* For k finite, using (10.1) again, we have

$$\pi(k|L) \approx \frac{wf(k)g(L-k)}{g(L)} \approx wf(k). \tag{10.14}$$

2. *Dip region.* When k and $L-k$ are simultaneously large, setting $k = \lambda L$ in (10.13) $(0 < \lambda < 1)$ yields the estimate, at large L,

$$\pi(k|L) \approx \frac{wf(k)g(L-k)}{g(L)} \approx \frac{1}{\lambda^{1+\theta}(1-\lambda)^\theta}\frac{wc}{L^{1+\theta}} \approx \frac{1}{\lambda^{1+\theta}(1-\lambda)^\theta}wf(L), \tag{10.15}$$

with a dip centred around $k_{\min} = L(1+\theta)/(1+2\theta)$.

3. In the region corresponding to $L-k$ finite, (10.13) simplifies into

$$\pi(k|L) \approx \frac{w(1-w)\theta}{L} Z^{\mathrm{f}}(w, L-k). \tag{10.16}$$

In particular, for $k = L$, $\pi(L|L) \approx w(1-w)\theta/L$, cf (8.39).

The weight of the downhill region is found to be asymptotically equal to w, using the same reasoning as in §7.2. The complement is borne by $g(L)/Z^{\mathrm{f}}(w, L) \approx 1-w$, see (8.36) and (10.21). The two other regions therefore do not contribute to the total weight, asymptotically.

In order to complete the picture we now investigate the distribution of B_L, the last unfinished interval.

10.3 Regimes for the distribution of B_L

According to (8.26), and in view of (10.1), for L large we have

$$\mathrm{Prob}(B_L = b) \approx (1-w)\frac{g(b)Z^{\mathrm{td}}(w, L-b)}{g(L)}. \tag{10.17}$$

Let us discuss the different regimes of this expression according to the magnitude of b.

1. If b is finite, the asymptotic estimate of $Z^{\mathrm{td}}(w, L-b)$ is given by (7.1), hence (10.17) becomes

$$\mathrm{Prob}(B_L = b) \approx \frac{w}{1-w}\frac{g(b)f(L-b)}{g(L)}. \tag{10.18}$$

2. If $b \sim L$, the same estimate (10.18) still holds, then setting $L-b = \lambda L$, we get

$$\mathrm{Prob}(B_L = b) \approx \frac{w}{1-w}\frac{g(b)f(L-b)}{g(L)} \approx \frac{wc}{1-w}\frac{L^{-1-\theta}}{\lambda^{1+\theta}(1-\lambda)^{\theta}}, \tag{10.19}$$

which has its minimum at $k_{\min} = L\,\theta/(1+2\theta)$.

3. If $L-b$ is finite, (10.17) becomes

$$\mathrm{Prob}(B_L = b) \approx (1-w)Z^{\mathrm{td}}(w, L-b), \tag{10.20}$$

in particular, see (8.36),

$$\mathrm{Prob}(B_L = L) = \mathrm{p}_0(L) = \frac{g(L)}{Z^{\mathrm{f}}(w, L)} \to 1-w. \tag{10.21}$$

Let us estimate, for later use, the probability that B_L is less than $L/2$. The result depends on the value of θ.

• If $\theta < 1$, using (10.19), we have

$$\mathrm{Prob}(B_L \le L/2) \approx \frac{wc}{1-w}L^{-\theta}\int_{1/2}^{1}\frac{\mathrm{d}\lambda}{\lambda^{1+\theta}(1-\lambda)^{\theta}} \approx \frac{wc}{1-w}\mathrm{B}_{\frac{1}{2}}\Big(1-\theta, -\theta\Big)L^{-\theta}. \tag{10.22}$$

• If $\theta > 1$, using (10.18), we have

$$\mathrm{Prob}(B_L \le L/2) \approx \frac{w\,\theta}{(1-w)L}\sum_{b=0}^{L/2} g(b) \approx \frac{w\,\theta}{(1-w)L}\sum_{b=0}^{L} g(b) \approx \frac{w\,\theta\langle X\rangle}{(1-w)L}. \tag{10.23}$$

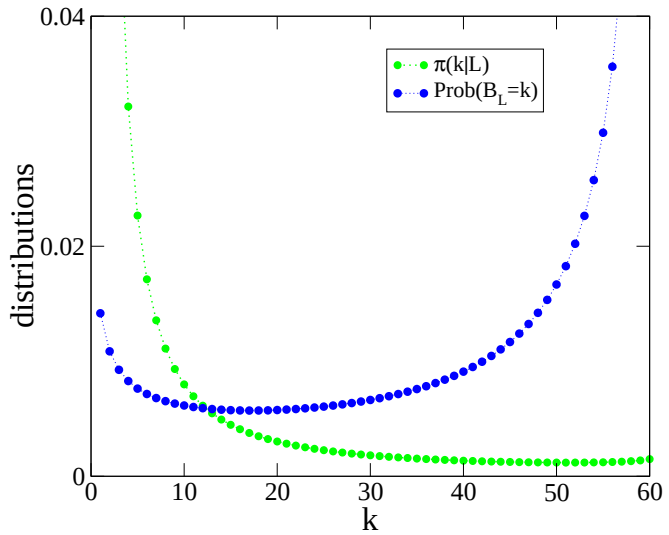

Fig. 11 Free renewal processes: distributions $\pi(k|L)$ (in green) and $\mathrm{Prob}(B_L = k)$ (in blue) for *Example 1* (see (4.1)), with $w = 0.8$ and $L = 60$. For this value of w, the asymptotic average number $\langle N_L \rangle \to 4$.

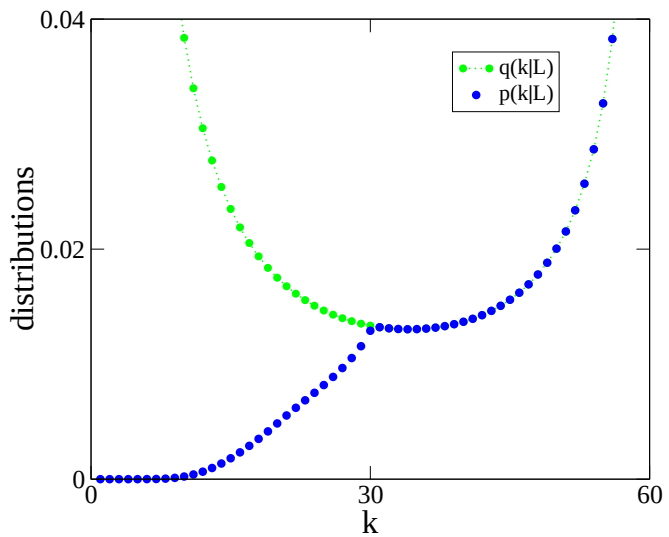

Fig. 12 Free renewal processes: distributions $\mathrm{Prob}(B_L = k) + \pi(k|L)(1 + \langle N_{L-k} \rangle)$ (in green) and $p(k|L)$ (in blue) for *Example 1* (see (4.1)), with $w = 0.8$ and $L = 60$. The distribution $q(k|L)$ is the restriction to $k > L/2$ of $\mathrm{Prob}(B_L = k) + \pi(k|L)(1 + \langle N_{L-k} \rangle)$ according to (8.57). For this value of w, the asymptotic average number $\langle N_L \rangle \to 4$. This figure is to be compared to the insets of figures 5 and 10.

10.4 Regimes for the distribution of the longest interval

The bulk of the distribution of $X_{\max}$ lies in the region $k > L/2$, and is therefore given by $q(k|L)$, see (8.55) or (8.57).

We start by giving an illustration. The distribution $p(k|L)$ of $X_{\max}$ on the whole interval $(0, L)$ is depicted in figure 12 for *Example 1* (see (4.1)). The restriction of this distribution to the second half $k > L/2$, that is $q(k|L)$, computed from (8.57) is also depicted. The corresponding figure for $\theta > 1$ is qualitatively alike.

Let us now compare the respective contributions of each of the two terms in (8.55) to the total weight in the region $k > L/2$. The first term is investigated in §10.3 above. The asymptotic estimate of the second term is as follows. We start with the case $\theta < 1$.

• If $\theta < 1$, we consider two regimes.
(i) The main contribution of the second term to the total weight comes from the regime $L - k \sim L$. Using the asymptotic estimate for large L,

$$\big(g \star Z^{\mathrm{td}} \star Z^{\mathrm{td}}\big)(w, L) \approx \frac{g(L)}{(1-w)^2}\Big(1 - \frac{2wc}{1-w}\frac{\theta\Gamma(-\theta)^2}{\Gamma(1-2\theta)L^{\theta}}\Big), \tag{10.24}$$

and setting $L - k = \lambda L$, the second term reads

$$\frac{wf(k)}{Z^{\mathrm{f}}(w, L)}\frac{c}{(1-w)^2\theta(L-k)^{\theta}} \approx \frac{wc}{1-w}\frac{L^{-1-\theta}}{\lambda^{\theta}(1-\lambda)^{1+\theta}}, \tag{10.25}$$

which is similar to (10.19). Summing this expression upon λ from 0 to 1/2 yields (10.22), that is $\mathrm{Prob}(B_L \leq L/2)$. Therefore adding this contribution to the first one, namely $\mathrm{Prob}(B_L > 1/2)$, gives unity, up to small corrections, in agreement with the fact that the weight of $p(k|L)$ in the left domain $k < L/2$ is negligible.
(ii) If $L - k$ is finite, then the second term reads

$$\frac{wf(k)\big(g \star Z^{\mathrm{td}} \star Z^{\mathrm{td}}\big)(w, L-k)}{Z^{\mathrm{f}}(w, L)} \approx \frac{(1-w)w\theta}{L}\big(g \star Z^{\mathrm{td}} \star Z^{\mathrm{td}}\big)(w, L-k), \tag{10.26}$$

which is subdominant compared to (10.20).

In order to get the correction of $\langle X_{\max}\rangle$ to L, we take the average of (8.55), by integrating each of the terms from 0 to 1/2 upon λ, using (10.19) and (10.25). Adding the contributions coming from the two terms, we finally obtain for the dominant correction,

$$\begin{aligned} L - \langle X_{\max}\rangle &\approx \frac{wc}{1-w}\Big(\mathrm{B}_{\frac{1}{2}}(1-\theta, 1-\theta) + \mathrm{B}_{\frac{1}{2}}(2-\theta, -\theta)\Big)L^{1-\theta} \\ &= \frac{wc}{1-w}\mathrm{B}_{\frac{1}{2}}\Big(1-\theta, -\theta\Big)L^{1-\theta}, \end{aligned} \tag{10.27}$$

which has the same structure as (7.31) or (2.46). The comment made below (7.32) also holds here. In the present case the expression in the right side of (10.27) is proportional to $\langle N_L\rangle = w/(1-w)$ times the critical mean interval $\langle X|L\rangle \sim L^{1-\theta}$, as given in [35].

• Likewise, for $\theta > 1$, the weight of the first term in (8.55) dominates upon the second one, and we find for the correction of the mean to L,

$$L - \langle X_{\max}\rangle \approx \frac{w}{1-w}\langle X\rangle, \tag{10.28}$$

showing that this correction is made of $\langle N_L\rangle = w/(1-w)$ intervals of size $\langle X\rangle$. This expression is therefore the perfect parallel of (7.32) or (2.48).

10.5 Probability for the last interval to be the longest one

Lastly, we investigate the behaviour of Q_L defined in (8.60) as the probability that B_L is the longest interval. An estimate of Q_L for $w < 1$ can be obtained by means of the inequality

$$1 - Q_L \lesssim \mathrm{Prob}(B_L \leq L/2). \tag{10.29}$$

In view of (10.22) and (10.27), we infer that asymptotically for L large, if $\theta < 1$,

$$Q_L \approx \frac{\langle X_{\max} \rangle}{L}, \tag{10.30}$$

while, if $\theta > 1$, in view of (10.23) and (10.28),

$$1 - Q_L \approx \frac{\theta}{L}(L - \langle X_{\max} \rangle). \tag{10.31}$$

In other words, for $w < 1$, $Q_L \to 1$. At criticality, $w = 1$, $Q_L \to Q_\infty = 0.626\ldots$, if $\theta < 1$, while if $\theta > 1$, $Q_L \sim L^{-(1-1/\theta)}$ [36]. For $w > 1$, $Q_L \to 0$.

Figure 13 depicts Q_L as a function of w for *Example 1* (see (4.1)) and for three different sizes, crossing at the universal critical value $Q_\infty = 0.626\ldots$ for $w = 1$ [36] and the data collapse obtained by using the scaling variable $x = (w-1)L^{1/2}$.

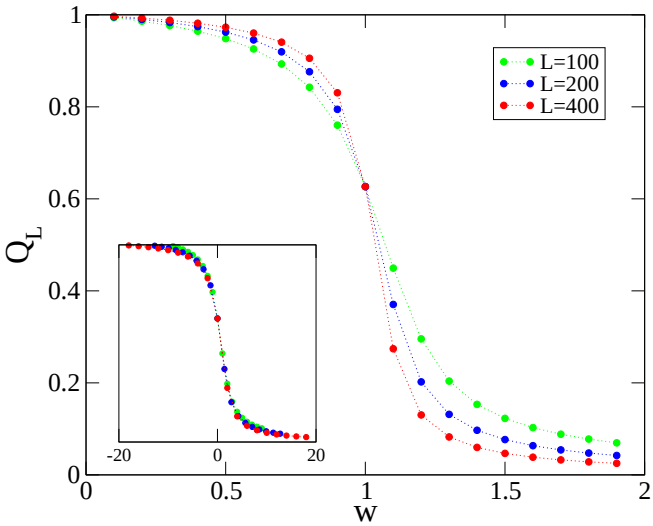

Fig. 13 Free renewal processes: probability Q_L that the last interval, B_L, is the longest one, for *Example 1* (see (4.1)), for three different values of L. The curves cross at the universal critical value $Q_\infty = 0.626\ldots$ for $w = 1$. Inset: after rescaling, Q_L against the scaling variable $x = (w-1)L^{1/2}$.

Table 2 summarises the results found in section 10 and recapitulates the results for the two other phases (disordered and critical). This table demonstrates a large degree of universality of the results, as was the case of table 1, with which it should be put in perspective.

Table 2 Dominant asymptotic behaviours at large L for free renewal processes with power-law distribution (1.1) for $f(k)$, in the different phases. The results in columns 2 and 3 (critical phase) are taken from [35,36]. In the last column, $\langle X|L\rangle \sim L^{1-\theta}$ if $\theta < 1$, or $\langle X|L\rangle \approx$ constant if $\theta > 1$.

	disordered	critical $\theta < 1$	critical $\theta > 1$	condensed
$\langle N_L\rangle$	$\frac{L}{\langle X\rangle}$	L^{θ}	$\frac{L}{\langle X\rangle}$	$\frac{w}{1-w}$
$\langle X\|L\rangle$	constant	$L^{1-\theta}$	$\langle X\rangle$	$L^{1-\theta}$ or constant
$\langle B_L\rangle$	constant	L	$L^{2-\theta}$	L
$\langle X_{\max}\rangle$	$\ln L$	L	$L^{1/\theta}$	L
$Z^{\mathrm{f}}(w,L)$	$\mathrm{e}^{L/\xi}$	1	1	$L^{-\theta}$

11 Conclusion

Let us summarise the salient aspects of this study.

We first recalled the main features of the condensation transition taking place in the thermodynamic limit ($L, n \to \infty$ with fixed ratio $\rho = L/n$) of random allocation models and ZRP, when the distribution of occupations is subexponential. These occupations are independent and identically distributed random variables conditioned by the value of their sum. The phase diagram is made of three phases: disordered, critical, and condensed. The critical line $\rho = \rho_c(\theta)$, where $\theta > 1$, separates the disordered phase at low density from the condensed phase at high density. Condensation manifests itself by the occurrence, in the thermodynamic limit, of a unique site with macroscopic occupation. In the language of particles and boxes (or sites), the condensate is by definition the site with the largest occupation. In the language of sums of random variables used all throughout the present work, the condensate $X_{\max}$ is the unique summand with extensive value. In the thermodynamic limit, the fraction $X_{\max}/L$ no longer fluctuates and takes the asymptotic value $1 - \rho_c/\rho$.

A second scenario for the same class of models consists in taking the $L \to \infty$ limit keeping the number of sites (or summands) fixed. In this limit the system condenses. There is again a single extensive summand $X_{\max}$, but now the fraction $X_{\max}/L$ tends to unity, which means that condensation is *total*. The novelty is that this occurs irrespective of the existence of a first moment $\langle X\rangle$, or in other words, irrespective of whether θ is smaller or larger than one. If L is large but finite, the distribution of $X_{\max}$ is peaked, with a width $L - \langle X_{\max}\rangle$ scaling as $L^{1-\theta}$, with a known amplitude, if $\theta < 1$, and asymptotically equal to $(n-1)\langle X\rangle$, if $\theta > 1$. Note that, in contrast to the previous case, one can no longer speak of a phase transition, nor even of a phase, since the system is made of a finite number n of summands (or sites).

This scenario is a good preparation for the study of condensation in free and tied-down renewal processes, with power-law distribution of intervals (1.1), which is the main aim of the present work. Instead of particle occupations and sites one speaks in terms of renewal events and intervals, whose sizes sum up to a fixed value L. The novelty—and complication—is that the number of these renewal events, or equivalently of intervals, N_L, fluctuates. For instance these renewal points are the passages by the origin of a random walk, as depicted in figure 3. A weight w is attached to each renewal event. In the language of random walks (or of polymer chains) w represents the reward or penalty when the walk touches the origin [26,33,34]. A high value of w favours configurations with a large number of intervals N_L, i.e., a disordered phase—or localised phase in the language of random walks. A low value of w favours configurations with a small number of intervals N_L, i.e., a condensed phase—or delocalised phase in the language of random walks. It is therefore intuitively clear that the same scenario of total condensation as seen above should prevail, where now the driving force is no longer a change in the density, ρ, but a change in the value of the weight w attached to each interval (or summand). In this respect it is worth noting the similarity between equations (2.46), (7.31) and (10.27) on one hand, and the similarity between equations (2.48), (7.32) and (10.28) on the other hand.

It turns out that, in the condensed phase, when $L \to \infty$, the distribution of the number of intervals, N_L, is completely universal, i.e., model independent, since it only depends on w and not even on the index θ. This distribution is geometric for free renewal processes, while it is a deformation of the latter for TDRP. More generally, an important distinction is to be made according to whether θ is less or larger than unity. In the first case the distribution $f(k)$ has no first moment, atypical events play a major role and the system becomes self-similar at criticality. In the second case the observables of interest depend on the first moment $\langle X \rangle$, which is finite.

In closing, let us broaden the perspective. The phase transition occurring when w passes through unity is second order for the density of intervals ν (defined in (5.8)) if $\theta < 1$, and first order if $\theta > 1$. On the other hand the correlation length diverges at the transition, see (5.13). The transition is therefore mixed order as was pointed out for the particular case of TDRP with *Example 2* (see (4.7)) in [29,30]. Furthermore, the magnetisation, defined as the alternating sum $m = (X_1 - X_2 + X_3 - \cdots)/L$, changes, when $L \to \infty$, from the value 0 in the disordered phase to ± 1 in the condensed phase since condensation is total. More on this can be found in [30]. If $\theta < 1$ the distribution of the magnetisation at criticality is broad and self-similar, both for free [48,35] and tied-down renewal processes [23]. At criticality, for $\theta < 1$, the non stationary two-time (or two-space) correlation function is also self-similar, again for both processes [35,23].[6]

Acknowledgements It is a pleasure to thank G Giacomin and J-M Luck for enlightening discussions. I am also indebted to S Grosskinsky and S Janson for useful correspondence.

Appendix

A Weight of the maximum for a Lévy $\frac{1}{2}$ stable law

We want to determine the weight on the left of the maximum considered in §2.3.3,

$$P_n = \mathrm{Prob}(X_{\max} < L/2) = \sum_{k=0}^{L/2} p_n(k|L) = 1 - \sum_{k=L/2+1}^{L} p_n(k|L) = 1 - \sum_{k=L/2+1}^{L} n\pi_n(k|L). \tag{A.1}$$

We use a continuum formalism where distributions are densities and variables are real numbers for the particular case where $f(k)$ is the Lévy $\frac{1}{2}$ stable density (6.11),

$$f(k) = \mathcal{L}_{\frac{1}{2},c}(k) = \frac{c}{k^{3/2}} \mathrm{e}^{-\pi c^2/k}. \tag{A.2}$$

Likewise, considering L as a real number and $Z_n(L)$ as a density,

$$Z_n(L) = \mathcal{L}_{\frac{1}{2},nc}(k) = \frac{nc}{L^{3/2}} \mathrm{e}^{-n^2\pi c^2/L}. \tag{A.3}$$

Thus

$$\pi_n(k|L) = \frac{f(k) Z_{n-1}(L-k)}{Z_n(L)} \tag{A.4}$$

is explicit. Setting $k = L/t$, we obtain

$$P_n = 1 - \frac{(n-1)c}{\sqrt{L}} \int_1^2 \mathrm{d}t \, \frac{t \, \mathrm{e}^{-\frac{\pi c^2 (n-t)^2}{L(t-1)}}}{(t-1)^{3/2}}. \tag{A.5}$$

[6] After submission of the present work, a study devoted to the statistics of $X_{\max}$ in the range $(L/2, L)$ for tied-down or free renewal processes at criticality ($w = 1$) was presented in [50]. For the TDRP case the result (3.48) with $w = 1$ is obtained. For the free renewal case, [50] predicts, if $w = 1$,

$$q(k|L) = g(k)(\langle N^{\mathrm{f}}_{L-k} \rangle - \langle N^{\mathrm{f}}_{L-k-1} \rangle) + f(k)(1 + \langle N^{\mathrm{f}}_{L-k} \rangle), \tag{11.1}$$

which is (8.56), with $w = 1$, noting that $Z^{\mathrm{td}}(1, L) = \langle N^{\mathrm{f}}_L \rangle - \langle N^{\mathrm{f}}_{L-1} \rangle$, as is clear by taking the generating functions of both sides.

Setting $c/\sqrt{L} = b/(2\sqrt{\pi})$, we finally get

$$P_n = \frac{1}{2}\left(2 - n\,\mathrm{erfc}\left(\frac{(n-2)b}{2}\right) + (n-2)\mathrm{e}^{(n-1)b^2}\,\mathrm{erfc}\left(\frac{nb}{2}\right)\right). \tag{A.6}$$

For L large, expanding in powers of b, we obtain

$$P_n = (n-1)(n-2)\left(\frac{b^2}{2} - n\frac{b^3}{3\sqrt{\pi}} + (n-1)\frac{b^4}{4} + \cdots.\right) \tag{A.7}$$

For *Example 1* (see (4.1)), $c = 1/(2\sqrt{\pi})$ hence $b = 1/\sqrt{L}$. The prediction

$$P_n \approx \frac{(n-1)(n-2)}{2L} \tag{A.8}$$

is in very good agreement with an exact numerical computation of P_n, as in figure 6.

With some more work, this result can be extended to any value of $\theta < 1$, as

$$P_n \approx (n-1)(n-2)\frac{4\pi c^2 A(\theta)}{L}, \tag{A.9}$$

where $A(\theta)$ is a known amplitude with a complicated expression. For instance, for $\theta = 1/3$, $A(\theta) = \sqrt{3}\Gamma(2/3)/(4\pi)$.

References

1. Bialas P, Burda Z and Johnston D 1997 *Nucl. Phys. B* **493** 505 *Condensation in the Backgammon model*
2. Janson S 2012 *Prob. Surveys* **9** 103 *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation*
3. Spitzer F 1970 Advances in Math. **5** 246 *Interaction of Markov Processes*
4. Andjel ED 1982 Ann. Prob. **10** 525 *Invariant measures for the zero-range process*
5. Drouffe JM, Godrèche C and Camia F 1998 *J. Phys. A* **31** L19 *A simple stochastic model for the dynamics of condensation*
6. Bialas P, Bogacz L, Burda Z and Johnston D 2000 *Nucl. Phys. B* **575** 599 *Finite size scaling of the balls in boxes model*
7. Jeon I, March P and Pittel B 2000 *Ann. Probab.* **28** 1162 *Size of the largest cluster under zero-range invariant measures*
8. Grosskinsky S, Schütz GM and Spohn H 2003 *J. Stat. Phys.* **113** 389 *Condensation in the zero range process: stationary and dynamical properties*
9. Evans MR and Hanney T 2005 *J. Phys. A* **38** R195 *Nonequilibrium Statistical Mechanics of the Zero-Range Process and Related Models*
10. Godrèche C 2007 *Lect. Notes Phys.* Springer **716** 261 *From urn models to zero-range processes: statics and dynamics*
11. Evans MR, Majumdar SN and Zia RKP 2006 *J. Stat. Phys.* **123** 357 *Canonical Analysis of Condensation in Factorised Steady States*
12. Ferrari PA, Landim C and Sisko VV 2007 *J. Stat. Phys.* **128** 1153 *Condensation for a Fixed Number of Independent Random Variables*
13. Evans MR and Majumdar SN 2008 *J. Stat. Mech.* P05004 *Condensation and Extreme Value Statistics*
14. Armendariz I and Loulakis M 2009 Probab. Theory Relat. Fields **145** 175 *Thermodynamic limit for the invariant measures in supercritical zero range processes*
15. Armendariz I and Loulakis M 2011 Stochastic Process. Appl. **121** 1138 *Conditional distribution of heavy tailed random variables on large deviations of their sum*
16. Godrèche C 2019 *J. Stat. Mech.* 063207 *Condensation for random variables conditioned by the value of their sum*
17. Feller W 1968 1971 *An Introduction to Probability Theory and its Applications* volumes 1&2 (New York: Wiley)
18. Doob JL 1948 *Trans. Amer. Math. Soc.* **63** 422 *Renewal theory from the point of view of the theory of probability*
19. Smith WL 1958 *J. Roy. Statist. Soc. Ser. B* **20** 243 *Renewal Theory and its Ramifications*
20. Cox DR 1962 *Renewal theory* (London: Methuen)
21. Wendel JG 1964 *Math. Scand.* **14** 21 *Zero-Free Intervals of Semi-Stable Markov Processes*
22. Godrèche C 2017 *J. Phys. A* **50** 195003 *Longest interval between zeros of the tied-down random walk, the Brownian bridge and related renewal processes*

23. Godrèche C 2017 *J. Stat. Mech.* P073205 *Two-time correlation and occupation time for the Brownian bridge and tied-down renewal processes*
24. Labarbe JM and Marckert JF 2007 *Elect. Journ. Prob.* **12** 229 *Asymptotics of Bernoulli random walks, bridges, excursions and meanders with a given number of peaks*
25. Vervaat W 1979 *Ann. Probab.* **7** 143 *A relation between Brownian bridge and Brownian excursion*
26. Fisher M 1984 *J. Stat. Phys.* **34** 667 *Walks, Walls, Wetting and Melting*
27. Poland D and Scheraga HA 1966 *J. Chem. Phys.* **45** 1456 *Phase Transitions in One Dimension and the Helix-Coil Transition in Polyamino Acids*
28. Poland D and Scheraga HA 1966 *J. Chem. Phys.* **45** 1464 *Occurrence of a Phase Transition in Nucleic Acid Models*
29. Bialas P, Burda Z and Johnston D 1999 *Nucl. Phys. B* **542** 413 *Phase diagram of the mean field model of simplicial gravity*
30. Bar A and Mukamel D 2014 *J. Stat. Mech.* P11001 *Mixed order transition and condensation in an exactly soluble one dimensional spin model*
31. Bar A, Majumdar SN, Schehr G and Mukamel D 2016 *Phys. Rev.* **93** 052130 *Exact extreme-value statistics at mixed-order transitions*
32. Barma M, Majumdar SN and Mukamel D 2019 *J. Phys. A* **52** 254001 *Fluctuation-dominated phase ordering at a mixed order transition*
33. Giacomin G 2007 *Random polymer models* (Imperial College Press: World Scientific)
34. Giacomin G 2011 *Disorder and critical phenomena through basic probability models* École d'été de probabilités de Saint-Flour *Lect. Notes Math.* Springer
35. Godrèche C and Luck JM 2001 *J. Stat. Phys.* **104** 489 *Statistics of the occupation time of renewal processes*
36. Godrèche C Majumdar SN and Schehr G 2015 *J. Stat. Mech.* P03014 *Statistics of the longest interval in renewal processes*
37. Bardou F, Bouchaud J P, Aspect A and Cohen-Tannoudji C 2002 *Lévy Statistics and Laser Cooling* (Cambridge: Cambridge University Press)
38. Barkai E 2003 *Phys. Rev. Lett.* **90** 104101 *Aging in Subdiffusion Generated by a Deterministic Dynamical System*
39. J. Beltrán J and C. Landim C 2012 *Probab. Theory Relat. Fields* **152** 781 *Metastability of reversible condensed zero range processes on a finite set*
40. Gnedenko B V and Kolmogorov A 1954 *Limit Distributions for Sums of Independent Random Variables* (Cambridge, Mass: Addison-Wesley)
41. Godrèche C and Luck J M 2005 J. Phys. A **38** 7215 *Dynamics of the condensate in zero-range processes*
42. Godrèche C and Luck JM 2001 *Eur. Phys. J. B* **23**, 473 *Nonequilibrium dynamics of the zeta urn model*
43. Gumbel E J 1958 *Statistics of Extremes* (New York: Columbia University Press)
44. Gnedenko B V 1943 *Ann. Math.* **44** 423 *Sur la distribution limite du terme maximum d'une série aléatoire*
45. Fréchet M 1927 *Ann. Soc. Polon. Math.* **6** 93 *Sur la loi de probabilité de l'écart maximum*
46. Chistyakov VP 1964 *Theor. Probab. Appl.* **9** 640 *A theorem on sums of independent positive random variables and its applications to branching random processes*
47. Dynkin EB 1955 *Izv. Akad. Nauk. SSSR Ser. Math.* **19** 247; 1961 *Selected Translations Math. Stat. Prob.* **1** 171 *Limit theorems for sums of independent random quantities*
48. Lamperti J 1958 *Trans. Amer. Math. Soc.* **88** 380 *An occupation time theorem for a class of stochastic processes*
49. Lamperti J 1961 *Am. Math. Soc.* **12** 724 *A contribution to renewal theory*
50. Höll M, Wang W and Barkai E arXiv:2006.06253 *Extreme value theory for constrained physical systems*