Condensation for a fluctuating number of independent random variables

Claude Godrèche
Université Paris-Saclay, CNRS, CEA, Institut de Physique Théorique, 91191 Gif-sur-Yvette, France

Abstract. The topics discussed in the present work are scattered in the literature and usually dressed up in other clothing. We give a stripped-down account of these topics in simple probabilistic terms in order to highlight the essentials. Besides reviewing the subject, this work investigates facets of the theory so far unexplored in previous studies.

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 $f(k) = \text{Prob}(X = k)$. Hereafter we consider the particular case of a subexponential distribution with asymptotic power-law decay $\dagger$

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

where the index $\theta$ 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,

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

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 thermodynamical limit,

$\dagger$ 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.
Condensation for a fluctuating number of independent random variables

$L, n \to \infty$, with $\rho = L/n$ fixed, the distribution of the size of this big step becomes narrow around $\Delta$. 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 steps $X_i$ are on the same footing, sharing the—now negative—excess difference $\Delta$. 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 steps $X_i$ are on the same footing, sharing the excess difference $\Delta$ (which is now either positive or negative).

Figure 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 $\Delta = L - \langle S_n \rangle = 4500$.

This scenario of condensation has been investigated in great detail and is basically understood. It is for example encountered in random allocation models where $n$ boxes (or sites) contain altogether $L$ particles, the $X_i$ representing the occupations of these boxes [1]. This situation in turn accounts for the stationary state of dynamical urn models such as zero range processes (ZRP) or variants [2, 3, 4]. 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 boxes 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,

$$S_{N_L} = \sum_{i=1}^{N_L} X_i,$$  \hspace{1cm} (1.2)


to satisfy either the inequality

$$S_{N_L} < L < S_{N_L+1},$$ \hspace{1cm} (1.3)


or the equality

$$S_{N_L} = L,$$ \hspace{1cm} (1.4)

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 (1.3) defines a free renewal process \cite{5, 6}, the process conditioned by the equality (1.4) defines a tied-down renewal process \cite{7, 8}. 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 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 (1.4) imposes $B_L = 0$.

A simple implementation of a TDRP is provided by the Bernoulli bridge, or tied-down random walk, made of $\pm 1$ steps, starting from the origin and ending at the origin at time $L$ \cite{7, 8}. 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 the Brownian bridge, also known as tied-down Brownian motion or else pinned Brownian motion.

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 weight parameter $w$
A tied-down random walk, or Bernoulli bridge, is a simple random walk, with steps \( \pm 1 \), 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, \( f(k) = \text{Prob}(X = k) \), is given by (3.23).

Figure 3. A tied-down random walk, or Bernoulli bridge, is a simple random walk, with steps \( \pm 1 \), 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, \( f(k) = \text{Prob}(X = k) \), is given by (3.23).

varies from larger values, favouring configurations with a large number of summands, to smaller ones, favouring atypical configurations with a smaller number of summands. Characterising this transition is the aim of the present work. Again, the occurrence of the phenomenon of condensation is due to: (i) subexponentiality of the distribution \( f(k) \), and: (ii) atypicality of the configurations.

TDRP fall into the class of linear systems considered by Fisher [9]. 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 [10], 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 boxes is varying [11], with an occupation variable \( X \) starting at \( k = 1 \). This model, as well as the spin domain model considered in [12] 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 [13, 14] 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 [13, 14]. 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.

Besides [13, 14], which give rigorous mathematical results, studies of the condensation phenomenon for 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.

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

This text is composed of three parts, which are respectively section 2, dealing with the class of random allocation models and ZRP, sections 3 to 6, dealing with TDRP, and finally, sections 7 and 8, dealing with free renewal processes. These three parts are both conceptually and analytically related.

Section 2 is a reminder of how condensation arises for random allocation models and ZRP, and will serve as a backdrop for the study contained in the two other parts. The first four subsections, §2.1 to §2.4, summarise the main features of the phenomenon of condensation. 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 [15] to which we refer for further bibliographical references. A complete mathematical review of the subject can be found in [16]. Subsection §2.5 contains novel aspects of the phenomenon of condensation for the class of models at hand. Namely, instead of considering the thermodynamical limit $L, n \to \infty$ with the ratio $\rho = L/n$ fixed, we investigate the situation where the number of summands $n$ is fixed and the value of the sum $S_n = L$ increases to infinity. Such a framework is precisely that considered in [17]. 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 $\rho_c$, 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 [17] ‘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 §2.5 and extend it to the case where $\theta < 1$, thus proving that even though the first moment $\langle X \rangle$ does not exist, there is however (total) condensation. We shall show that the correction of the mean largest summand to $L$ scales as $L^{1-\theta}$, with a known amplitude, given in (2.20).

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, 9 and 10, and the results (2.20), (2.21), (6.15), (6.16), (8.16) and (8.17).

Sections 3 to 6 are devoted to TDRP with power-law distribution of summands (1.1). Section 3 gives a systematic description of the formalism, followed, in sections 4 to 6 by the analysis of the behaviour of the system in the different phases. As said in the abstract, these topics are scattered in the literature. The analyses presented in these sections go deeper than previous studies, especially in the description of the condensation phenomenon, as detailed in section 6.

Likewise, sections 7 and 8, 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 9, together with tables 1 and 2, gives a summary of the present study.
2. Reminder of 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 with a reminder of how condensation arises for random allocation models and ZRP which will serve as a backdrop for the study to come. To this end we first give some elements of the formalism.

2.1. Statistical weight of a configuration

The joint conditional probability associated to a configuration \( \{X_1 = k_1, \ldots, X_n = k_n\} \), with \( S_n = L \) given, reads

\[
p(k_1, \ldots, k_n|L) = \text{Prob}(X_1 = k_1, \ldots, X_n = k_n|S_n = L) = \frac{1}{Z_n(L)} f(k_1) \cdots f(k_n) \delta\left(\sum_{i=1}^{n} k_i, L\right),
\]

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

\[
Z_n(L) = \sum_{\{k_i\}} f(k_1) \cdots f(k_n) \delta\left(\sum_{i=1}^{n} k_i, L\right) = (f^*)^n(L) = \langle \delta(S_n, L) \rangle = \text{Prob}(S_n = L).
\]

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 \( n \) boxes, \( L \) particles in total, and where the \( X_i \) are the occupation numbers of these boxes, i.e., the number of particles in each of them. Since these boxes can be empty, the occupation probability \( f(0) \) is non zero in general. It is therefore natural to initialise \( L \) to 0. In particular, the probability that the sum of occupations be zero is that all boxes are empty, i.e.,

\[
Z_n(0) = f(0)^n.
\]

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

\[
Z_0(L) = \delta(L, 0),
\]

which serves as an initial condition for the recursion

\[
Z_n(L) = \sum_{k=0}^{L} f(k) Z_{n-1}(L-k).
\]

As can be seen either from (2.2) or (2.4), the generating function of \( Z_n(L) \) with respect to \( L \) yields

\[
Z_n(z) = \sum_{L \geq 0} z^L Z_n(L) = \hat{f}(z)^n,
\]

\( \dagger \) For the sake of simplicity we restrict the study to the case where \( f(k) \) is a normalisable probability distribution.
where the generating function of $f(k)$ with respect to $k$ is
\[ \hat{f}(z) = \sum_{k \geq 0} z^k f(k). \]

The marginal distribution of the occupation of a generic site, say site 1, is
\[ p_n(k|L) = \text{Prob}(X_1 = k|S_n = L) = \langle \delta(X_1, k) \rangle = \frac{f(k) Z_{n-1} (L-k)}{Z_n(L)}, \tag{2.5} \]
where $\langle \cdot \rangle$ is the average with respect to (2.1). The mean conditional occupation is
\[ \langle X|L \rangle = \sum_{k \geq 0} k p_n(k|L) = \frac{L}{n} = \rho. \tag{2.6} \]

2.2. 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. Let $X_{\text{max}}$ be the largest summand (or occupation) under the conditioning $S_n = L$,
\[ X_{\text{max}} = \max(X_1, \ldots, X_n). \]

The distribution function of this variable is
\[ F_n^{(1)}(k|L) = \text{Prob}(X_{\text{max}} \leq k|S_n = L) = \frac{\text{Prob}(X_{\text{max}} \leq k, S_n = L)}{Z_n(L)}, \]
whose numerator is
\[ F_n^{(1)}(k|L)|_{\text{num}} = \sum_{k_1=0}^k f(k_1) \cdots \sum_{k_n=0}^k f(k_n) \delta\left(\sum_{i=1}^n k_i, L\right). \tag{2.7} \]

The generating function of the latter reads
\[ \sum_{L \geq 0} z^L F_n^{(1)}(k|L)|_{\text{num}} = \prod_{i=1}^n \left( \sum_{k_i=0}^k f(k_i) z^{k_i} \right) = \hat{f}(z,k)^n, \]
where
\[ \hat{f}(z,k) = \sum_{j=0}^k f(j) z^j. \]

The distribution of the largest occupation is thus given by the difference
\[ p_n^{(1)}(k|L) = \text{Prob}(X_{\text{max}} = k|S_n = L) = F_n^{(1)}(k|L) - F_n^{(1)}(k-1|L), \]
where
\[ F_n^{(1)}(0|L)|_{\text{num}} = f(0)^n \delta(L,0). \]

Its generating function is
\[ \sum_{L \geq 0} z^L p_n^{(1)}(k|L)|_{\text{num}} = \hat{f}(z,k)^n - \hat{f}(z,k-1)^n. \tag{2.8} \]

The numerator (2.7) obeys the recursion
\[ F_n^{(1)}(k|L)|_{\text{num}} = \sum_{j=0}^{\min(k,L)} f(j) F_{n-1}^{(1)}(k-j)|_{\text{num}}, \]
with initial condition
\[ F_0^{(1)}(k|L)_{\text{num}} = \delta(L, 0). \]

Let us note that if the occupation number \( X_1 \) is larger than \( L/2 \), then it is necessarily the largest one, \( X_{\text{max}} \). If so the probability distribution of the latter, \( p_n^{(1)}(k|L) \), is identical to \( np_n(k|L) \), since there are \( n \) possible choices of the generic summand \( X_1 \). We shall see later that this identity extends to an asymptotic equivalence for values beyond the strict range \( X_1 > L/2 \).

### 2.3. Phenomenology of condensation for zrp: the thermodynamical limit

This subsection is a short reminder of well-known facts on the phenomenon of condensation for a large system of size \( n \), with large total occupation \( L \), at fixed density \( \rho = L/n \). A more detailed account can be found, e.g., in [15]. We assume that \( \langle X \rangle = \rho_c \) is finite. Evidence for the existence of a condensate, i.e., a site with a macroscopic occupation, is given by the investigation of the single occupation distribution (2.5). There are three regimes to consider, according to the respective values of \( \rho \) and \( \rho_c \).

1. **Subcritical regime** (\( \rho < \rho_c \))

The asymptotic estimate of the partition function \( Z_n(L) \) is given by the saddle-point method
\[
Z_n(L) = \oint \frac{dz}{2\pi i z^{L+1}} \tilde{f}(z)^n \sim \frac{\tilde{f}(z_0)^n}{z_0^L},
\]
where \( z_0 \) obeys the saddle-point (sp) equation
\[
\frac{z_0 \tilde{f}'(z_0)}{\tilde{f}(z_0)} = \rho. \tag{2.9}
\]
This equation has a solution \( z_0(\rho) \) for any \( \rho < \rho_c \). It follows that
\[
p_n(k|L)_{\text{sp}} \approx \frac{z_0^k f(k)}{\tilde{f}(z_0)}, \tag{2.10}
\]
which is no longer dependent on \( L \) and \( n \) separately and only depends on their ratio \( \rho \). Note that (2.9) and (2.10) entail that
\[
\langle X | L \rangle_{\text{sp}} = \sum_{k \geq 0} kp_n(k|L)_{\text{sp}} \approx \frac{z_0 \tilde{f}'(z_0)}{\tilde{f}(z_0)} = \rho,
\]
consistently with (2.6).

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. The bulk of the partition function is given by the central limit theorem and
\[
p_n(k|L) \approx f(k), \tag{2.11}
\]
up to finite-size corrections. At criticality the equality \( \langle X | L \rangle = \langle X \rangle \) holds identically.

3. **Supercritical regime** (\( \rho > \rho_c \))

In this regime the saddle-point equation (2.9) can no longer be satisfied because \( z_0 \) sticks to the head of the cut of \( \tilde{f}(z) \). The excess difference,
\[
\Delta = L - n\langle X \rangle = n(\rho - \rho_c),
\]
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 given by its right tail (see (2.19) for an example and [15] for more details),

$$Z_n(L) \approx \frac{nc}{\Delta^{1+\theta}}. \quad (2.12)$$

In the supercritical regime, the marginal distribution $p_n(k|L)$ 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 (2.11) 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 (2.12), is asymptotically equal to

$$\frac{f(k)}{Z_n(L)} \approx \frac{c/\Delta^{1+\theta}}{nc/\Delta^{1+\theta}} = \frac{1}{n}. \quad (2.13)$$

On the other hand, $Z_{n-1}(L - k)$, is given by its bulk since $L - k \approx n\rho_c$. Hence, if $1 < \theta < 2$, $Z_{n-1}(L - k)$ is given by the right tail (2.12). So, for any $\theta > 1$, obtained from (2.13) or (2.14),

$$\sum_{k \in \text{hump}} p_n(k|L)_{\text{cond}} \approx \frac{1}{n}, \quad (2.15)$$

which demonstrates that the excess difference $\Delta$ is borne by only one summand.

This hump becomes peaked in the thermodynamical 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 $k \gg 1, \Delta - k \gg 1$, interpolates between the critical part of $p_n(k|L)$, for $k$ of order 1, and the condensate, for $k$ close to $\Delta$. It corresponds to the **dip region** on figure 4. In this region, $Z_{n-1}(L - k)$ is given by its right tail (2.12). So, for any $\theta > 1$,

$$p_n(k|L)_{\text{dip}} \approx c \left[ \frac{\Delta}{k(\Delta - k)} \right]^{1+\theta} \approx \frac{f(k)f(\Delta - k)}{f(\Delta)}. \quad (2.16)$$

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 configurations where the excess difference is shared by two summands.
Setting \( k = \lambda \Delta \) in (2.16) and introducing a cutoff \( \epsilon \Delta \), the weight of these configurations can be estimated as

\[
\sum_{k=\epsilon \Delta}^{(1-\epsilon)\Delta} p_n(k|L)_{\text{dip}} \sim \Delta^{-\theta} \sim n^{-\theta}.
\]

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

### 2.4. An illustration

![Figure 4](image-url)

**Figure 4.** Random allocations models and ZRP: comparison of \( np_n(k|L) \) (single occupation distribution) with \( p_n^{(1)}(k|L) \) (distribution of the maximum) for the example (2.17), 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 \( np_n(k|L) \) and \( p_n^{(1)}(k|L) \), for \( k > L/2 \), as explained in §2.2. Moreover, there is already excellent numerical coincidence between the two curves as soon as \( k > 58 \gtrsim \Delta/2 \).

As said above, in the condensed phase the fluctuations of the condensate become narrow around \( \Delta \) in the thermodynamical limit. The weight of the distribution of the maximum in the left region, \( k < \Delta/2 \), is asymptotically negligible, because this region corresponds to those rare events where the condensate (i.e., the maximum) is less than \( \Delta/2 \). All the weight of this distribution thus lies in the region \( k > \Delta/2 \).

Figure 4 depicts a comparison between \( np_n(k|L) \), obtained from (2.5), and \( p_n^{(1)}(k|L) \) obtained from (2.8), on the following example, defined by the normalised distribution

\[
f(k) = \frac{1}{\zeta(1+\theta)} \frac{1}{(k+1)^{1+\theta}}, \quad (k \geq 0),
\]

(2.17)
Condensation for a fluctuating number of independent random variables

where \( \zeta(s) = \sum_{n \geq 1} 1/n^s \) is Riemann zeta function. In this figure \( \theta = 3, \rho_c = 0.1106 \ldots, n = 600, L = 168 \). This choice of parameters correspond to a density \( \rho = 0.28 \) slightly larger than \( 2\rho_c \), where \( \Delta \) and \( L/2 \) coincide. The top of the hump is approximately located at \( \Delta \approx 102 \) and 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.

If \( \rho \) increases, the peak of the condensate moves towards the right end, at \( L \), as detailed in the next subsection.

2.5. Condensation when \( n \) is fixed and \( L \to \infty \)

If \( \rho \gg \rho_c \), condensation becomes total, in the sense that the peak of the condensate is asymptotically located at \( L \). This occurs irrespective of the existence of a first moment \( \langle X \rangle \), or in other words, irrespective of whether \( \theta \) is smaller or larger than one.

Before analysing this phenomenon, we start by giving an illustration of the phenomenon. Figure 5 depicts the case where \( \theta = 1/2 \) for the following example,

\[
\hat{f}(z) = \frac{1 - \sqrt{1 - z}}{z},
\]

(2.18)

So

\[
f(k) = (-1)^k \left( \frac{1}{k+1} \right) = \frac{1}{2^{2k+1}} \frac{(2k)!}{k!(k+1)!} \approx \frac{c}{k^{3/2}},
\]

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

\[
Z_n(L) = \frac{n}{2^{2L+n}(2L+n)} \binom{2L+n}{L} \approx \frac{nc}{L^{3/2}},
\]

(2.19)
for $n$ fixed, which is a particular case of (2.12).

Coming back to the general case, the width of the peak can be analysed by considering

$$L - \langle X_{\text{max}} \rangle = \sum_{\ell=0}^{L-1} \ell \ p^{(1)}_n(L - \ell | L) \approx \sum_{\ell=0}^{L/2} \ell \ p^{(1)}_n(L - \ell | L) \approx \sum_{\ell=0}^{L/2} \ell n p_n(L - \ell | L).$$

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

If $\theta < 1$, the dominant contribution comes from values of $\ell = L - k$ comparable to $L$. Setting $\ell = \lambda L$, we have

$$L - \langle X_{\text{max}} \rangle \approx (n-1) c L^{1-\theta} \int_0^{1/2} d\lambda \frac{\lambda}{[(\lambda(1-\lambda)]^{1+\theta}} \approx (n-1) c B\left(\frac{1}{2}; 1-\theta, -\theta\right) L^{1-\theta}, \quad (2.20)$$

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

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

$$L - \langle X_{\text{max}} \rangle \approx n \sum_{\ell=0}^{L/2} \ell \ f(L - \ell) Z_{n-1}(\ell) \approx \sum_{\ell=0}^{L/2} \ell Z_{n-1}(\ell) \approx (n-1) \sum_{\ell=0}^{L/2} \ell f(\ell) = (n-1) \langle X \rangle. \quad (2.21)$$

This last result (2.21) has a simple interpretation. It says that the correction $L - \langle X_{\text{max}} \rangle$ is made of $n-1$ intervals, all of sizes equal to the average interval $\langle X \rangle$.

In the present situation where $\rho \gg \rho_c$, since the two distributions $p^{(1)}_n(k|L)$ and $np_n(k|L)$ coincide for $k > L/2$, it is easy to recover the fact that the weight under the peak of $p_n(k|L)$ is equal to $1/n$,

$$1 \approx \sum_{k=L/2+1}^{L} p^{(1)}_n(k|L) = \sum_{k=L/2+1}^{L} np_n(k|L). \quad (2.22)$$

This property is manifest in figure 4, but it is all the more true in the present case since the peak is located at $L$.

All the discussion above is a preparation for respectively subsections §6.3 and §8.4, where figures 9 and 10 are to be compared to figure 5, and equations (6.15), (6.16), (8.16) and (8.17) to (2.20) and (2.21).

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

$$f(0) = 0. \quad (3.1)$$

In the temporal language $L$ is the total duration, 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 [11, 12], $w$ has the interpretation of the ratio $y/y_c$, where $y$ is a fugacity, and $y_c$ its value at criticality.
3.1. Joint distribution

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

$$p(k_1, \ldots, k_n, n|L) = \text{Prob}(X_1 = k_1, \ldots, X_{N_L} = k_n, N_L = n|S_{N_L} = L) = \frac{1}{Z^{td}(w, L)} w^n f(k_1) \ldots f(k_n) \delta \left( \sum_{i=1}^n k_i, L \right),$$

(3.2)

where the denominator is the tied-down partition function

$$Z^{td}(w, L) = \sum_{n \geq 0} w^n \sum_{\{k_i\}} f(k_1) \ldots f(k_n) \delta \left( \sum_{i=1}^n k_i, L \right)$$

$$= \sum_{n \geq 0} w^n Z_n(L) = \delta(L, 0) + \sum_{n \geq 1} w^n (f^*)^n(L).$$

(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 (2.3). The first values of $Z^{td}(w, L)$ are

$$Z^{td}(w, 0) = 1, \quad Z^{td}(w, 1) = wf(1), \quad Z^{td}(w, 2) = wf(2) + w^2 f(1)^2,$$

$$Z^{td}(w, 3) = wf(3) + 2w^2 f(1)f(2) + w^3 f(1)^3,$$

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

$$\hat{Z}^{td}(w, z) = \sum_{L \geq 0} z^L Z^{td}(w, L) = \sum_{n \geq 0} w^n \hat{f}(z)^n = \frac{1}{1 - wf(z)}.$$

(3.4)

Note that $Z^{td}(w, L)$ can be seen as the grand canonical partition function of the system with respect to $N_L$. Finally, for $w = 1$, the tied-down partition function,

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

(3.5)

is the probability that a renewal occurs at $L$.

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 $n$,

$$p_n(L) = \text{Prob}(N_L = n) = \frac{w^n Z_n(L)}{Z^{td}(w, L)}.$$

(3.6)

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

$$p_0(L) = \frac{\delta(L, 0)}{Z^{td}(w, L)}, \quad p_1(L) = \frac{wf(L)}{Z^{td}(w, L)}, \quad p_2(L) = \frac{w^2 \sum_{k_1} f(k_1)f(L-k_1)}{Z^{td}(w, L)} \ldots$$

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

$$p_n(L) = \frac{w^n (f^*)^n(L)}{Z^{td}(w, L)} = \frac{w^n [\hat{f}(z)^n]}{Z^{td}(w, L)}.$$
where the notation $[.]_L$ stands for the $L-$th coefficient of the series inside the brackets. Hence the generating function with respect to $L$ where the notation \[ \text{Condensation for a fluctuating number of independent random variables} \]

\[
\sum_{L \geq 0} z^L p_n(L)_{\text{num}} = w^n \hat{f}(z)^n. \tag{3.7}
\]

Taking the sum of the right side upon $n \geq 0$ yields back $\hat{Z}^{\text{td}}(w, z)$ given in (3.4).

The first moment of this distribution is by definition \[ \langle N_L \rangle = \sum_{n \geq 1} n p_n(L). \]

The generating function of its numerator reads, using (3.7) \[ \sum_{L \geq 0} z^L \langle N_L \rangle_{\text{num}} = \sum_{n \geq 1} n(w \hat{f}(z))^n = \frac{w \hat{f}(z)}{1 - w \hat{f}(z)} = w \frac{d \hat{Z}^{\text{td}}(w, z)}{dw}, \tag{3.8} \]

hence \[ \langle N_L \rangle = w \frac{d \ln Z^{\text{td}}(w, L)}{dw}, \tag{3.9} \]

as expected in the grand canonical ensemble with respect to $N_L$.

More generally, the generating function of the moments of $N_L$ is given by \[ \langle v^{N_L} \rangle = \sum_{n \geq 0} v^n p_n(L). \]

Taking the generating function of its numerator, using (3.7), \[ \sum_{L \geq 0} z^L \langle v^{N_L} \rangle_{\text{num}} = \sum_{n \geq 0} v^n(w \hat{f}(z))^n = \frac{1}{1 - vw \hat{f}(z)} = \hat{Z}^{\text{td}}(vw, z), \]

yields \[ \langle v^{N_L} \rangle = \frac{Z^{\text{td}}(vw, L)}{Z^{\text{td}}(w, L)}. \tag{3.10} \]

Likewise, the inverse moment $\langle 1/N_L \rangle$ is \[ \langle \frac{1}{N_L} \rangle = \sum_{n \geq 1} \frac{p_n(L)}{n} = \frac{1}{Z^{\text{td}}(w, L)} \sum_{n \geq 1} \frac{[w \hat{f}(z)]^n}{n} L = \frac{1}{Z^{\text{td}}(w, L)} \left[ \ln(1 - w \hat{f}(z)) \right]_L. \]

### 3.3. Distribution of the size of a generic interval

The marginal distribution of one of the summands, say $X_1$, is by definition \[ p(k|L) = \text{Prob}(X_1 = k|S_{N_L} = L) = \langle \delta(X_1, k) \rangle, \]

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 \[ p(k|L)_{\text{num}} = \sum_{n \geq 1} \sum_{k_1} \delta(k_1, k) \sum_{k_2, \ldots} w^{n-1}f(k_2) \cdots \delta(k_1 + \sum k_i, L) \]

\[ = \sum_{k_1} \delta(k_1, k)w f(k_1)\delta(k_1, L) + \sum_{k_1, k_2} \delta(k_1, k)w^2 f(k_1) f(k_2) \delta(k_1 + k_2, L) + \cdots \]

\[ = w f(k) \delta(k, L) + w f(k) Z^{\text{td}}(w, L-k)(1 - \delta(k, L)). \tag{3.11} \]
Finally

\[ p(k|L) = \frac{w f(k)}{Z^{td}(w, L)} \delta(k, L) + w f(k) \frac{Z^{td}(w, L - k)}{Z^{td}(w, L)} (1 - \delta(k, L)), \]  

(3.12)

where the first term corresponds to \( n = 1 \), i.e.,

\[ p(L|L) = \frac{w f(L)}{Z^{td}(w, L)} = \text{Prob}(N_L = 1). \]  

(3.13)

Also, since \( Z^{td}(w, 0) = 1 \), (3.12) can be more compactly written as

\[ p(k|L) = \frac{w f(k) Z^{td}(w, L - k)}{Z^{td}(w, L)}. \]  

(3.14)

The generating function of the numerator of (3.14) with respect to \( L \) yields

\[ \sum_{L \geq k} z^L p(k|L)_{\text{num}} = w z^k f(k) Z^{td}(w, z) = \frac{w z^k f(k)}{1 - w f(z)}, \]  

(3.15)

Summing (3.14) upon \( k \) we obtain

\[ Z^{td}(w, L) = \sum_{k=1}^{L} w f(k) Z^{td}(w, L - k), \quad L \geq 1, \]  

(3.16)

which can also be obtained by multiplying the recursion (2.4) for \( Z_n(L) \) by \( w^n \) and summing on \( n \).

3.4. Mean interval \( \langle X|L \rangle \)

This is, by definition,

\[ \langle X|L \rangle = \sum_{k \geq 1} k p(k|L). \]

Multiplying (3.15) by \( k \) and summing upon \( k \) yields

\[ \sum_{L \geq k} z^L \langle X|L \rangle_{\text{num}} = \frac{w z^k \tilde{f}(z)}{1 - w f(z)}, \]  

(3.17)

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,

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

(3.18)

then, taking the generating function of the right side gives

\[ \sum_{L \geq 0} z^L \langle X|L \rangle_{\text{num}} = \sum_{L \geq 1} z^L L \left[ - \ln(1 - w \tilde{f}(z)) \right]_L = z \frac{d}{dz} \left( - \ln(1 - w \tilde{f}(z)) \right) = \frac{w z \tilde{f}(z)}{1 - w f(z)}. \]
3.5. The longest interval

By definition, the longest interval is
\[ X_{\text{max}} = \max(X_1, \ldots, X_{N_L}), \]
the distribution function of which is defined as
\[ F^{(1)}(k|L) = \text{Prob}(X_{\text{max}} \leq k|L) = \sum_{n \geq 0} \sum_{k_1=0}^{k} \ldots \sum_{k_n=1}^{k} p\{k_i\}, n|L) = \frac{F^{(1)}(k|L)_{\text{num}}}{Z^{td}(w, L)}, \]
with initial value
\[ F^{(1)}(k|0)_{\text{num}} = 1. \] \hspace{1cm} (3.19)

Note that \( F^{(1)}(L|L)_{\text{num}} = Z^{td}(w, L), \) hence \( F^{(1)}(L|L) = 1. \) The generating function of the numerator is
\[
\sum_{L \geq 0} z^L F^{(1)}(k|L)_{\text{num}} = 1 + \sum_{n \geq 1} \prod_{i=1}^{n} \left( \sum_{k_i=1}^{k} w f(k_i) z^{k_i} \right) \\
= 1 + \sum_{n \geq 1} \left( w \hat{f}(z, k) \right)^n = \frac{1}{1 - w \hat{f}(z, k)},
\]
where
\[ \hat{f}(z, k) = \sum_{j=1}^{k} z^j f(j). \]
The numerator obeys the recursion (renewal) equation, which generalises (2.4) of [8],
\[ F^{(1)}(k|L)_{\text{num}} = \sum_{j=1}^{\min(k, L)} w f(j) F^{(1)}(k|L-j)_{\text{num}}, \]
with initial condition (3.19).

The distribution of \( X_{\text{max}} \) is given by the difference
\[ p^{(1)}(k|L) = \text{Prob}(X_{\text{max}} = k) = F^{(1)}(k|L) - F^{(1)}(k-1|L), \]
where \( F^{(1)}(0|L) = \delta(L, 0), \) with generating function
\[
\sum_{L \geq 0} z^L p^{(1)}(k|L)_{\text{num}} = \frac{1}{1 - w \hat{f}(z, k)} - \frac{1}{1 - w \hat{f}(z, k-1)} \\
= \frac{w z^k f(k)}{[1 - w \hat{f}(z, k)][1 - w \hat{f}(z, k-1)]}. \hspace{1cm} (3.20)
\]
Note that (3.20) can be obtained by multiplying (2.8) by \( w^n \) and summing on \( n. \) The end point value is the same as \( p(L|L) \) (3.13), i.e.,
\[ p^{(1)}(L|L) = \text{Prob}(N_L = 1) = \frac{w f(L)}{Z^{td}(w, L)}. \] \hspace{1cm} (3.21)

When \( X_{\text{max}} = k > L/2, \) the longest interval is unique. Generalising the reasoning made in [7, 8] we can decompose a configuration into three contributions to obtain
\[ p^{(1)}(k|L)_{\text{num}} = \sum_{i=0}^{L-k} Z^{td}(w, i) w f(k) Z^{td}(w, L-k-i). \]

Hence, denoting the restriction of \( p^{(1)}(k|L) \) to \( k > L/2 \) by \( q(k|L) \) we have, for \( k > L/2,
\[ q(k|L) = \frac{w f(k)}{Z^{td}(w, L)} \sum_{i=0}^{L-k} Z^{td}(w, i) Z^{td}(w, L-k-i) = \frac{w f(k)(Z^{td} * Z^{td})(w, L-k)}{Z^{td}(w, L)}. \] (3.22)
3.6. 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 7, 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, \( f(k) \), represents the probability of first return of the walk after 2\( k \) steps, or equivalently after \( k \) tick marks on figure 3,

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

with generating function

\[
\tilde{f}(z) = 1 - \sqrt{1 - z}.
\]

The partition function (3.5) for \( w = 1 \) represents the probability that the walk returns at the origin after 2\( L \) steps, or equivalently after \( L \) tick marks,

\[
Z_{td}(1, L) = \frac{1}{2^{2L}} \binom{2L}{L} \approx \frac{1}{\sqrt{\pi L}},
\]

with generating function

\[
\tilde{Z}_{td}(1, z) = \frac{1}{\sqrt{1 - z}}.
\]

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

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

with \( n \leq L \). Note that

\[
f(L) = Z_{td}(1, L - 1) - Z_{td}(1, L).
\]

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

\[
f(k) = \frac{1}{\zeta(1 + \theta)} k^{1+\theta}, \quad k > 0.
\]

So

\[
\tilde{f}(z) = \frac{\text{Li}_{1+\theta}(z)}{\zeta(1 + \theta)},
\]

where \( \text{Li}_s(z) \) is the polylogarithm function,

\[
\text{Li}_s(z) = \sum_{k \geq 1} \frac{z^k}{k^s}.
\]

The mean \( \langle X \rangle = \zeta(\theta)/\zeta(1 + \theta) \). This is the distribution used, e.g., in [1, 11, 18, 12, 19, 20].
3.7. Regimes and condensation transition for tied-down renewal processes

Demonstrating the existence of a phase transition in the model defined by (3.2), when \( w \) crosses the value one, is a classical subject. This model is a particular instance of a linear system, as described in [9], where the mechanism of the transition is explained in simple terms. This transition is also studied in [11, 12, 19, 20] for Example 2. Let us first analyse the large \( L \) behaviour of \( Z^{td}(w, L) \). Recalling (3.4) we have, for a contour encircling the origin,

\[
Z^{td}(w, L) = \oint \frac{dz}{2\pi i} \hat{Z}^{td}(w, z) = \oint \frac{dz}{2\pi i z^{L+1}} \frac{1}{1 - w \hat{f}(z)}.
\]

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

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

\[
Z^{td}(w, L) \approx z_0^{-L} \frac{w z_0 \hat{f}'(z_0)}{w_0 \hat{f}'(z_0)}.
\]

(b) If \( w = 1 \), then \( z_0 = 1 \). The asymptotic estimate of \( Z^{td}(1, L) \) is given in (5.4).

(c) If \( w < 1 \), the denominator \( 1 - w \hat{f}(z) \) has no zero, but it is singular for \( z = z_0 = 1 \) (which is the singularity of \( \hat{f}(z) \)). Hence \( z_0 \) sticks to 1. The asymptotic estimate of \( Z^{td}(w, L) \) is given in (6.1).

This is the switch mechanism of Fisher [9]: the condition determining \( z_0 \) switches from \( z_0 \) being the smallest root of the equation \( 1 - w \hat{f}(z) = 0 \) to being the closest real singularity of \( \hat{f}(z) \), which is a cut at \( z = z_0 = 1 \). This non analytical switch signals the phase transition. The reduced free energy reads [9]

\[
f = \lim_{L \to \infty} -\frac{1}{L} \ln Z^{td}(w, L) = \ln z_0.
\]

The three cases above are successively reviewed in the next sections.

4. 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 (3.27) in (3.14), which leads to

\[
p(k|L) \approx w f(k) z_0^k = w f(k) e^{-k/\xi}, \quad \xi = \frac{1}{|\ln z_0|}.
\]

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

\[
(X|L) \approx w_0 \hat{f}'(z_0),
\]

an expression which can also be inferred from (3.17). Thus (3.27) can be recast as

\[
Z^{td}(w, L) \approx \frac{z_0^{-L}}{(X|L)}.
\]

The distribution of \( N_L \) is given by (3.6)

\[
p_n(L) = \frac{w^n Z_n(L)}{Z^{td}(w, L)} \approx w^n Z_n(L) w \hat{f}'(z_0) z_0^{L+1}.
\]
This distribution obeys the central limit theorem, as illustrated on the example below. Using (3.9), we obtain the asymptotic expression of $\langle N_L \rangle$,

$$\langle N_L \rangle \approx -L \frac{w \, dz_0}{z_0} \, dw \approx \frac{L}{\langle X|L \rangle},$$

which means that

$$\frac{1}{\langle N_L \rangle} \approx \left( \frac{1}{N_L} \right).$$

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

$$\nu_L = \frac{\langle N_L \rangle}{L}, \quad (4.2)$$

then, asymptotically, we have

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

We illustrate these general statements on Example 1 (see (3.23)), for which $z_0$ is explicit,

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

hence

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

and the following asymptotic expressions hold,

$$Z^{td}(w, L) \approx \frac{2(w - 1)w^{2L}}{(2w - 1)^{L+1}},$$

$$\langle X|L \rangle \approx \frac{2w - 1}{2(w - 1)},$$

$$\langle N_L \rangle \approx \frac{L}{\langle X|L \rangle} + \frac{w}{(w - 1)(2w - 1)},$$

$$\text{Var} \, N_L \approx \frac{L}{2(2w - 1)^2} - \frac{w(2w^2 - 1)}{(2w - 1)^2(w - 1)^2},$$

$$p_n(L) \approx \frac{1}{\sqrt{2\pi \text{Var} \, N_L}} \, \exp \left( -\frac{(n - \langle N_L \rangle)^2}{2 \text{Var} \, N_L} \right),$$

$$\nu \approx \frac{2(w - 1)}{2w - 1}. \quad (4.3)$$

Figure 6 depicts a comparison between the exact finite-size expression of the density $\nu_L$ obtained by means of (3.8) for $L = 1000$ as a function of $w$, and the asymptotic expression (4.3). It vanishes at the transition $w = 1$, where the system becomes critical.

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

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

as can be easily inferred by means of the expansion (5.1), a result already present in [11]. If $\theta > 1$ the density $\nu$ tends to $1/\langle X \rangle$ for $w \to 1$, as can be seen using the
expansion (5.2). It is therefore discontinuous at the transition, as noted in [11, 12]. Likewise, it is easy to see that

$$\xi^{-1} \sim \begin{cases} 
(w - 1)^{1/\theta} & \theta < 1 \\
w - 1 & \theta > 1.
\end{cases} \quad (4.4)$$

The correlation length diverges at the transition, while the order parameter \( \nu \) is either continuous (\( \theta < 1 \)) or discontinuous (\( \theta > 1 \)). The transition is therefore of mixed order [11, 12].

Finally we note that the intervals \( X_i \) behave essentially as iid random variables, with distribution (4.1), hence the statistics of the longest interval belongs to the Gumbel class. This is detailed on Example 2 in [19].

5. 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 \( \theta \) 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 [7, 8] and is also addressed in [19, 20]. 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 \( p(k|L) \) of the size of a generic interval.

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

$$
\hat{f}(s) \approx \begin{cases} 
1 - |a|s^\theta, & \theta < 1 \\
1 - s\langle X \rangle + \cdots + as^\theta, & \theta > 1
\end{cases}
$$

(5.1)

(5.2)

with

$$a = c \Gamma(-\theta) = \theta \Gamma(-\theta)k_0^\theta,
$$

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

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

(5.3)

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.

### 5.1. Distribution $f(k)$ with index $\theta < 1$

Since $\tilde{Z}_{td}(1, z) = 1/(1 - \tilde{f}(z))$, in Laplace space we have $\hat{Z}_{td}(1, s) \approx 1/as^\theta$ which yields the expression of the partition function (see (4.6) in [8])

$$Z_{td}(1, L) \approx \frac{\theta \sin \pi \theta}{\pi c L} L^{\theta - 1}.
$$

(5.4)

For instance, setting $\theta = 1/2$ and $c = 1/(2\sqrt{\pi})$ restores (3.24).

#### 5.1.1. The number of intervals

We have (see (4.10) in [8]),

$$\langle N_L \rangle \approx \frac{\Gamma(\theta)}{\Gamma(1 - \theta) \Gamma(2\theta)} \left( \frac{L}{k_0} \right)^\theta,
$$

(5.5)

which can be easily deduced from (3.8). For the specific case of Example 1 (see (3.23)), we have the exact result (see (2.47) in [8])

$$\langle N_L \rangle = \frac{1}{Z_{td}(1, L)} - 1 = \frac{2^{2L}}{(2L)^L} - 1 \approx \sqrt{\pi} L,
$$

which is in agreement with (5.5), with $\theta = 1/2, c = 1/(2\sqrt{\pi})$. We know from (3.6) that the distribution of $N_L$ is given by the ratio

$$p_n(L) = \frac{Z_n(L)}{Z_{td}(1, L)}.
$$

For $n$ and $L$ large $p_n(L)$ has a scaling form. On the one hand, according to the central limit theorem, the scaling form of the numerator is given by

$$Z_n(L) \approx \frac{1}{n^{1/\theta}} L_{\theta,c} \left( \frac{L}{n^{1/\theta}} \right)^\theta,
$$

where $L_{\theta,c}$ is the density of the stable law of index $\theta$, tail parameter $c$ and asymmetry parameter $\beta = 1$ (see e.g., [15]). Then using (5.4), we get, with $u = L/n^{1/\theta}$,

$$p_n(L) \approx \frac{\pi c}{\theta \sin \pi \theta} \frac{1}{L^\theta} \frac{L}{n^{1/\theta}} L_{\theta,c} \left( \frac{L}{n^{1/\theta}} \right)^\theta \approx \frac{\pi c}{\theta \sin \pi \theta} L^\theta u L_{\theta,c}(u).
$$

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

$$L_{1/2,c}(u) = \frac{c e^{-\pi c^2/u}}{u^{3/2}},
$$
hence (see (2.49) in [8]), for Example 1 (see (3.23)),
\[ p_n(L) \approx \frac{v}{2\sqrt{L}} e^{-v^2/4}, \quad v = \frac{1}{\sqrt{u}} = \frac{n}{\sqrt{L}}. \]
Moreover, for this example, for \( n \) and \( L \) finite, \( p_n(L) \) is explicit since both \( Z_n(L) \) given by (3.25) and \( Z^\text{td}(1, L) \), given by (3.24), are known explicitly.

\[ \text{Figure 7. TD} \text{rp}: \text{distribution of the size of a generic interval } p(k|L), \text{with } L = 100, \text{ for Example 1 (see (3.23)), at criticality (} w = 1). \text{ Exact refers to the middle expression in (5.6), asymptotic to the rightmost one. The } y\text{–axis is on a logarithmic scale.} \]

5.1.2. The distribution of the size of a generic interval 

(i) In all regimes where \( \ell = L - k \) is large, using (5.4) we have
\[ p(k|L) = f(k) \frac{Z^\text{td}(1, L - k)}{Z^\text{td}(1, L)} \approx f(k) \left( 1 - \frac{k}{L} \right)^{\theta - 1}. \] (5.6)
For instance, if \( 1 \ll k \ll L \),
\[ p(k|L) \approx f(k) \approx \frac{c}{k^{1+\theta}}, \]
while if \( k = \lambda L \), with \( \lambda \in (0, 1) \),
\[ p(k|L) \approx \frac{c}{\lambda^{1+\theta}(1 - \lambda)^{1-\theta} L^{1+\theta}}, \]
which is minimum at \( \lambda = (1 + \theta)/2 \).

(ii) On the other hand, if \( \ell = L - k = O(1) \),
\[ p(L - \ell|L) = f(L - \ell) \frac{Z^\text{td}(1, \ell)}{Z^\text{td}(1, L)} \approx \frac{\pi e^2 Z^\text{td}(1, \ell)}{\theta \sin \pi \theta} \frac{1}{L^{2\theta}}. \]
In particular, for $k = L$,

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

In this regime the ratio of $p(k|L)$ to the estimate (5.6) reads

$$\frac{p(L-\ell|L)}{f(L-\ell)(1 - k/L)^{\theta - 1}} = \frac{\pi c}{\theta \sin \pi \theta} \approx \frac{\pi c}{\theta \sin \pi \theta},$$

which tends to one when $\ell$ becomes large, if one refers to (5.4).

All these results can be illustrated on Example 1 (see (3.23)). For instance figure 7 gives a comparison between the exact expression of $p(k|L)$ computed for $L = 100$ by means of the middle expression in (5.6) and its asymptotic form given by the rightmost expression in (5.6).

The generating function of the mean interval $\langle X | L \rangle$ given in (3.17) yields the estimate, in Laplace space,

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

hence (see (4.18) in [8])

$$\langle X | L \rangle \approx \frac{\theta}{Z^{td}(1, L)} \approx \frac{\pi c}{\sin \pi \theta} L^{1-\theta}. \quad (5.7)$$

This result can be recovered by taking the average of the estimate (5.6). It predicts correctly that the product $\langle X | L \rangle \langle N_L \rangle \sim L$. For Example 1, (see (3.23)), the computation leads to the exact result (see (2.58) in [8])

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

5.1.3. The longest interval  

The study of the longest interval for the case $w = 1, \theta < 1$ is done in [7, 8] (see also [19]). In contrast with the case of random allocation models and ZRP where $p^{(1)}(k|L) = np(k|L)$ for $k > L/2$ (see §2.1), these two quantities are quite different for all values of $k$, except for $k = L$, where they are equal, see (3.21).

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 $X^{(k)} (k = 2, 3, \ldots)$ do so [7, 8]. Moreover $X_{max}$ continues to fluctuate when $L \to \infty$ while for genuine condensation as in section 2 above or in section 6 below, its distribution is peaked. Finally, the dominant contribution to the weight of $p(k|L)$ comes from values of $k$ less than a small cutoff.

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

We have, using (3.4) (see (4.73) in [8]), for $\theta > 1$,

$$Z^{td}(1, L) \approx \frac{1}{\langle X \rangle} + \frac{c}{\theta(\theta - 1)\langle X \rangle^2} L^{1-\theta}.$$
The average value of $N_L$ is obtained by means of (3.8)\)

$$\langle N_L \rangle \approx \begin{cases} \frac{L}{\langle X \rangle} + \frac{c}{(\theta - 1)(1 - \theta)} \langle X \rangle^2 L^{2 - \theta} & 1 < \theta < 2 \\ \frac{L}{\langle X \rangle} + \frac{\text{Var} X}{\langle X \rangle^2} & \theta > 2 \end{cases}$$ \quad (5.8)$$

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

$$p_n(L) = \frac{Z_n(L)}{Z_{\text{tot}}(1, L)} \approx \langle X \rangle Z_n(L).$$

The asymptotic estimate for $\langle X | L \rangle$ is obtained by analysing (3.17), yielding for $\theta > 1$,

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

which is the same expression as for a free renewal process [21]. The single interval distribution has the form

$$p(k | L) = f(k) \frac{Z_{\text{tot}}(1, L - k)}{Z_{\text{tot}}(1, L)} \approx f(k) L \to \infty$$

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

$$p(L | L) = \frac{f(L)}{Z_{\text{tot}}(1, L)} \approx f(L) \langle X \rangle.$$

The distribution of the longest interval is analysed in [8]. The result is

$$F^{(1)}(k | L) \approx e^{-L/\langle X \rangle (k/k_0)^{-\theta}},$$

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

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

we have, as $L \to \infty$, $Y_L \to Y_F$, with limiting distribution

$$\text{Prob}(Y_F < x) = e^{-1/x^\theta}$$

which is the Fréchet law. Therefore

$$\langle X_{\text{max}} \rangle \approx k_0 \left( \frac{L}{\langle X \rangle} \right)^{1/\theta} \langle Y_F \rangle_{\Gamma(1-1/\theta)},$$

as was already the case for free renewal processes [22].

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

The aim of this section—central in this work—is to investigate the statistics of the number of intervals and characterise the fluctuations of the size 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, $p(k | L)$, and of the longest one, $p^{(1)}(k | L)$. Related material can be found in [13, 19, 20].

\| Equation (5.8) corrects the improper expression (4.74) given in [8] for this quantity.
6.1. Asymptotic estimates

Starting from (3.4) and linearising with respect to the singular part, we obtain, when \( L \to \infty \), for any value of \( \theta \),

\[
Z_{td}(w, L) \approx \frac{w}{(1-w)^2} \frac{c}{L^{1+\theta}} \approx \frac{w}{(1-w)^2} f(L),
\]

(6.1)

Alternatively, it suffices to notice that, for \( n \) fixed and \( L \) large, for any subexponential distribution \cite{23},

\[
Z_n(L) \approx nf(L),
\]

(6.2)
as for example in (3.25), which entails

\[
Z_{td}(w, L) = \sum_{n \geq 0} w^n Z_n(L) \approx \sum_{n \geq 0} w^n nf(L),
\]

which restores (6.1). Likewise, we find

\[
(Z_{td} * Z_{td})(w, L) \approx \frac{2w}{(1-w)^3} f(L).
\]

(6.3)

If one substitutes (6.1) in the expression of the mean \( \langle N_L \rangle \) (3.9), we obtain, for any value of \( \theta \), the universal result

\[
\langle N_L \rangle \to \frac{1 + w}{1 - w}.
\]

This result can be found alternatively using (6.4) below. More generally, using (3.10), the asymptotic distribution of \( N_L \) reads

\[
\langle v^{N_L} \rangle = \sum_{n \geq 0} v^n p_n(L) = \frac{Z_{td}(vw, L)}{Z_{td}(w, L)} \approx \frac{y(1-w)^2}{(1-vw)^2},
\]
hence extracting the coefficient of order \( n \) in \( y \) of this expression leads to the asymptotic distribution

\[
p_n(L) \to n(1-w)^2 w^{n-1},
\]

(6.4)

which is independent of \( \theta \) and therefore completely universal, see figure 8. The same result can be found by noting that

\[
p_n(L) = \frac{w^n Z_n(L)}{Z_{td}(w, L)} \approx \frac{Z_n(L)}{f(L)} \frac{(1-w)^2 w^{n-1}}{L \to \infty} \to n(1-w)^2 w^{n-1},
\]

using (6.2) again. The inverse moment \( \langle 1/N_L \rangle \) can be obtained by using (6.4) above,

\[
\langle \frac{1}{N_L} \rangle \to 1 - w.
\]

(6.5)

As a consequence of (6.5) we have, for any value of \( \theta \),

\[
\langle X|\rangle \approx (1 - w)L,
\]

(6.6)

which can also be found from the asymptotic analysis of (3.17).
6.2. Regimes for the distribution of the size of a generic interval

For $L$ large, the marginal distribution of the size of a generic interval is obtained by substituting (6.1) in (3.12), leading to

$$p(k|L) \approx w^2 f(k) \frac{Z_{td}^{L}(w, L-k)}{f(L)}.$$  (6.7)

Figure 9 depicts the distribution $p(k|L)$ (together with the distribution of the longest interval $p^{(1)}(k|L)$, see $\S6.3$ below, for $L = 60$ and $w = 0.6$ computed from (3.14), with Example 1 (see (3.23)). As can be seen on this figure, there are three distinct regions for $p(k|L)$, namely, a downhill region, followed by a long dip region, then an uphill region which accounts for the fluctuations of the condensate. We shall discuss the behaviour of $p(k|L)$ given by (6.7) in each of these regions successively, assuming that $L$ is large.

(i) For $k$ finite we have, using again (6.1),

$$p(k|L) \approx w f(k).$$  (6.8)

(ii) In the dip region, where $k$ and $L - k$ are simultaneously large, setting $k = \lambda L$ in (6.8) ($0 < \lambda < 1$) yields the estimate

$$p(k|L) \approx w f(k) \frac{f(L-k)}{f(L)} \approx \frac{w}{\lambda(1-\lambda)} \frac{c}{L^{1+\theta}},$$  (6.9)
showing that the dip centred around \( L/2 \) becomes deeper and deeper with \( L \).

(iii) The condensate region corresponds to \( \ell = L - k \) finite, where (6.7) simplifies into
\[
p(L - \ell|L) \approx (1 - w)^2 Z^{td}(w, \ell).
\]

Let us now estimate the contributions of each of these regions to the total weight.

(1) Introducing a cutoff \( \Lambda \), such that \( 1 \ll \Lambda \ll L \), the weight of the downhill region can be estimated by the sum
\[
\sum_{k=1}^{\Lambda} p(k|L) \approx \sum_{k=1}^{\Lambda} w f(k) \to \sum_{k=1}^{\infty} w f(k) = w. \quad (6.10)
\]

(2) The weight of the uphill region can be estimated likewise,
\[
\sum_{\ell=0}^{\Lambda} p(L - \ell|L) \approx (1 - w)^2 \sum_{\ell=0}^{\Lambda} Z^{td}(w, \ell)
\to (1 - w)^2 \sum_{\ell=0}^{\infty} Z^{td}(w, \ell) = 1 - w \quad (6.11)
\]
where the last step is obtained by setting \( z = 1 \) in the expression of the generating function (3.4). The right side of (6.11), \( 1 - w \), is precisely the limiting value of \( \langle 1/N_L \rangle \), see (6.5). This result is therefore the analogue of (2.15) predicting a weight of the condensate equal to \( 1/n \) for the random allocation models and ZRP.

From (6.10) and (6.11) we conclude that the weights of the downhill and uphill regions add to one, hence that the contribution of the dip region is necessarily subdominant, as we now show.

(3) The weight of the dip region can be estimated using (6.9), as
\[
\sum_{k=1}^{L-\Lambda} p(k|L) \approx L^{-\theta} wc \int_{1-\epsilon}^{1} \frac{d\lambda}{[\lambda(1-\lambda)]^{1+\theta}},
\]
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 9.

6.3. Regimes for the distribution of the longest interval

As can be seen on figure 9, there are two main regions for the distribution of the maximum, \( p^{(1)}(k|L) \). For \( k \leq L/2 \), as we shall see shortly, the contribution of \( p^{(1)}(k|L) \) to the total weight is vanishingly small. Hence we restrict the rest of the discussion to the region \( (L/2 < k \leq L) \), where \( p^{(1)}(k|L) \) has the simpler expression \( q(k|L) \) given by (3.22). Using (6.1), its asymptotic estimate is
\[
q(k|L) \approx_{L \to \infty} (1 - w)^2 f(k) \frac{(Z^{td} * Z^{td})(w, L-k)}{f(L)},
\]
or, equivalently, for \( \ell = L - k \in (0, L/2 - 1) \),
\[
q(L - \ell|L) \approx_{L \to \infty} (1 - w)^2 f(L - \ell) \frac{(Z^{td} * Z^{td})(w, \ell)}{f(L)}. \quad (6.12)
\]
As a preliminary remark let us note that the ratio
\[
r(\ell) = \frac{q(L - \ell|L)}{p(L - \ell|L)} = \frac{(Z^{td} * Z^{td})(w, \ell)}{Z^{td}(w, \ell)}
\]
Figure 9. tdrp: exact distributions of the size of a generic interval, \( p(k|L) \), and of the longest one, \( p^{(1)}(k|L) \), in the condensed phase, for \( L = 60 \) and \( w = 0.6 \) computed from (3.14) and (3.20), with Example 1 (see (3.23)). The two curves join at \( wz(L)/Z_{td}(w, L) \approx (1 - w)^2 \) on the y-axis, for \( k = 60 \). The cusp of \( p^{(1)}(k|L) \) at \( L/2 \) is more visible in the inset.

is an increasing function of \( \ell \), with reaches the limit, for large \( \ell \),

\[
r(\ell) \rightarrow \frac{2}{1 - w}.
\]

Its first values are

\[
r(0) = 1, \quad r(1) = 2, \quad r(2) = \frac{3wf(1)^2 + 2f(2)}{wf(1)^2 + f(2)},
\]

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

(i) For \( \ell \) finite, we have, see (6.12),

\[
q(L - \ell|L) \approx (1 - w)^2(Z_{td} \ast Z_{td})(w, \ell).
\]  

(6.13)

In particular for \( \ell = 0 \),

\[
p^{(1)}(L|L) = q(L|L) = p(L|L) = \text{Prob}(N_L = 1) \approx (1 - w)^2.
\]

If \( 1 \ll \ell \ll L \), (6.13) simplifies to

\[
q(L - \ell|L) \approx \frac{2w}{1 - w} f(\ell).
\]

(ii) If \( \ell \) et \( L - \ell \) are simultaneously large, with \( \ell = \lambda L \), we have

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

(6.14)

which is proportional to (6.9), with ratio \( 2/(1 - w) \).
The weight under the peak of the condensate tends to unity,
\[ \sum_{\ell=0}^{L} \mathcal{P}_{\ell}^{(1)}(L - \ell | L) \approx (1 - w)^2 \sum_{\ell=0}^{L} (Z_{\text{td}}^{\dagger} \ast Z_{\text{td}})(w, \ell) \]
\[ \rightarrow (1 - w)^2 \sum_{\ell=0}^{\infty} (Z_{\text{td}}^{\dagger} \ast Z_{\text{td}})(w, \ell) = 1, \]
by a computation similar to (6.11), using now the fact that the last sum is equal to \( \frac{1}{1 - w} \), as can be seen by setting \( z = 1 \) in the expression of the generating function \( \tilde{Z}_{\text{td}}(w, z)^2 \). We can further analyse the nature of the condensate by considering the width of the peak,
\[ L - \langle X_{\text{max}} \rangle = \sum_{\ell=0}^{L-1} \ell \mathcal{P}_{\ell}^{(1)}(L - \ell | L) \approx \sum_{\ell=0}^{L/2} \ell \mathcal{P}_{\ell}^{(1)}(L - \ell | L). \]

The dominant contribution to this sum depends on whether \( \theta \) is smaller or larger than one.

If \( \theta < 1 \), the dominant contribution comes from (6.14), i.e., for \( \ell \) comparable to \( L \),
\[ L - \langle X_{\text{max}} \rangle \approx 2wc \frac{1}{1 - w} L^{1 - \theta} \int_{0}^{1/2} d\lambda \frac{\lambda}{(\lambda(1 - \lambda))^{1+\theta}} \]
\[ \approx 2wc \frac{1}{1 - w} B\left(\frac{1}{2}; 1 - \theta, -\theta\right) L^{1 - \theta}, \quad (6.15) \]
where \( B(\cdot) \) is the incomplete beta function, which, for example, is equal to 2 for \( \theta = 1/2 \).

If \( \theta > 1 \), the main contribution comes from (6.13),
\[ L - \langle X_{\text{max}} \rangle \approx (1 - w)^2 \sum_{\ell=0}^{L} \ell (Z_{\text{td}} \ast Z_{\text{td}})(w, \ell) \]
\[ \rightarrow (1 - w)^2 \sum_{\ell=0}^{\infty} \ell (Z_{\text{td}}^{\dagger} \ast Z_{\text{td}})(w, \ell) \rightarrow \frac{2w}{1 - w} \langle X \rangle, \quad (6.16) \]
using (6.3). This last result (6.16) has a simple interpretation. It says that the correction \( L - \langle X_{\text{max}} \rangle \) is made of \( \langle N_{L} \rangle - 1 \) (which is equal to \( 2w/(1 - w) \)) intervals, all of sizes equal to the average interval \( \langle X \rangle \). It is therefore the perfect parallel of the result (2.21). Likewise, the right side of (6.15) can be interpreted as being proportional to \( \langle N_{L} \rangle - 1 \) times the critical mean interval \( \langle X | L \rangle \), given in (5.7).

6.4. 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 (6.9)) 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 \( p(L - \ell | L) \) for \( \ell \) in the layer \((0, \Lambda)\). This sum is asymptotically equal to \( 1 - w \) according to (6.11), which turns out to be also the asymptotic estimate of \( (1/N_{L}) \). The interpretation of this result is clear. When \( X_{1} \) is larger than \( L/2 \), this interval is necessarily the longest one,
Condensation for a fluctuating number of independent random variables

i.e., \(X_1 = X_{\text{max}}\). Furthermore, since this interval is chosen amongst \(N_L\) intervals, we expect that, in average,

\[
\sum_{\ell=0}^{A} p(L - \ell|L) \approx \left( \frac{1}{N_L} \right) \sum_{\ell=0}^{A} p^{(1)}(L - \ell|L),
\]

But the sum on the right side, namely the weight of \(X_{\text{max}}\) in the same layer is asymptotically equal to one. This simple heuristic reasoning, which generalises that done for the case of random allocation models and ZRP, see (2.22), therefore recovers (6.11).

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

Table 1 summarises the results found in sections 4, 5 and 6, which demonstrate a large degree of universality.

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

7. General statements on free renewal processes

We now turn now to the case of free renewal processes. The random number \(N_L\) of intervals up to \(L\) is defined through the condition (1.3), \(S_{N_L} < L < S_{N_L+1}\). 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.

7.1. Weighted joint distribution for free renewal processes

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

\[
p(k_1, \ldots, k_n, b, n|L) = \text{Prob}(\{X_i = k_i\}, B_L = b, N_L = n) = \frac{1}{Z^f(w, L)} w^n f(k_1) \cdots f(k_n) g(b) \delta\left(\sum_{i=1}^{n} k_i + b, L\right),
\]

(7.1)
Condensation for a fluctuating number of independent random variables

where $g(b)$ is the tail probability (or complementary distribution function) defined in (5.3),

$$g(b) = \text{Prob}(X > b) = \sum_{k>b} f(k) = f(b+1) + f(b+2) + \cdots \tag{7.2}$$

Since the summands $X_i$ have the interpretation of the sizes of the intervals, we take $f(0) = 0$. For $n = 0$,

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

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

$$\tilde{g}(z) = \sum_{b \geq 0} z^b g(b) = \frac{1}{1 - z - \sum_{b \geq 1} z^b \sum_{k=1}^b f(k) = \frac{1 - \tilde{f}(z)}{1 - z}, \tag{7.3}$$

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

The denominator of (7.1) is the free partition function, obtained by summing on $n$, on the $k_i$ and on $b$,

$$Z_f(w, L) = \sum_{n \geq 0} \sum_{b \geq 0} \sum_{\{k_i\}} p(k_1, \ldots, k_n, b, n | L)$$

$$= \sum_{n \geq 0} w^n \sum_{b \geq 0} \sum_{\{k_i\}} f(k_1) \cdots f(k_n) g(b) \delta(\sum_{i=1}^n k_i + b, L)$$

$$= \sum_{b \geq 0} g(b) \left[ \delta(b, L) + \sum_{k_1} w f(k_1) \delta(k_1 + b, L) \right.$$  

$$+ \sum_{k_1, k_2} w^2 f(k_1) f(k_2) \delta(k_1 + k_2 + b, L) + \ldots \right]$$

$$= g(L) + w g * f + w^2 g * f * f + \ldots = \sum_{n \geq 0} w^n (g * (f *)^n)(L). \tag{7.4}$$

For instance,

$$Z_f(w, 0) = 1, \quad Z_f(w, 1) = g(1) + w f(1),$$

$$Z_f(w, 2) = g(2) + w f(2) + w f(1) g(1) + w^2 f(1)^2,$$

and so on. From (7.4) we have

$$\tilde{Z}_f(w, z) = \sum_{L \geq 0} z^L Z_f(w, L) = \sum_{n \geq 0} \left( w \tilde{f}(z) \right)^n \tilde{g}(z) = \frac{\tilde{g}(z)}{1 - w \tilde{f}(z)}. \tag{7.5}$$

It is interesting to note that

$$Z_f(w, L) = \sum_{n \geq 0} w^n \sum_{b \geq 0} \text{Prob}(S_n = L - b) g(b) = (g * Z_{\text{td}})(w, L), \tag{7.6}$$

where $Z_{\text{td}}(w, L)$ 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 $Z_f(1, L) = 1$, as can be seen from (7.5), and the joint probability distribution (7.1) simplifies accordingly.
Condensation for a fluctuating number of independent random variables

7.2. Distribution of $N_L$

As for TDRP we denote this distribution as

$$p_n(L) = \text{Prob}(N_L = n).$$

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

$$p_0(L) = \frac{g(L)}{Z^f(w, L)} = \text{Prob}(B_L = L), \quad p_1(L) = \frac{w(g \ast f)(L)}{Z^f(w, L)}, \quad (7.7)$$

and so on. More generally,

$$p_n(L) = \frac{w^n(g \ast (f \ast Z_n)(L))}{Z^f(w, L)} = \frac{w^n(g \ast Z^a)(w, L)}{(g \ast Z^a)(w, L)}Z^f(w, L) = w^n(g \ast Z^a)(w, L)Z^f(w, L).$$

Summing (7.1) on $b$ and on the $k_i$, and taking the generating function with respect to $L$ yields

$$\sum_{L \geq 0} z^L p_n(L)_{\text{num}} = (w \tilde{f}(z))^n \tilde{g}(z),$$

to be compared to (7.5). Therefore

$$\sum_{L \geq 0} z^L \langle N_L \rangle_{\text{num}} = \sum_{n \geq 0} w^n \tilde{f}(z) \tilde{g}(z) = \tilde{Z}^i(w, z),$$

and

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

as for TDRP, see (3.9). More generally,

$$\sum_{L \geq 0} z^L \langle v^{N_L} \rangle_{\text{num}} = \sum_{n \geq 0} v^n (w \tilde{f}(z))^n \tilde{g}(z) = \tilde{Z}^i(vw, z),$$

so we obtain, as for TDRP, see (3.10),

$$\langle v^{N_L} \rangle = \sum_{n \geq 0} v^n p_n(L) = \frac{Z^i(vw, L)}{Z^f(w, L)}. \quad (7.8)$$

7.3. Distribution of $S_{NL}$

We recall that this quantity is the sum of the $N_L$ intervals before $L$, see (1.2) and figure 2. Summing (7.1) upon $n$, $b$, and $\{k_i\}$ implies

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

which generalises the expression for this quantity when $w = 1$ [21]. By derivation with respect to $x$ then setting $x = 1$, leads to

$$\sum_{L \geq 0} z^L (S_{NL})_{\text{num}} = \frac{wz \tilde{f}(z) \tilde{g}(z)}{(1 - w \tilde{f}(z))^2}, \quad (7.9)$$

whose summation with (7.12) below leads to the equality

$$\sum_{L \geq 0} z^L (\langle S_{NL} \rangle_{\text{num}} + \langle B_L \rangle_{\text{num}}) = z \frac{d \tilde{Z}^i(w, z)}{dz},$$
which expresses the sum rule
\[ \langle S_{NL} \rangle + \langle B_L \rangle = L. \]

The asymptotic behaviours of these quantities for \( w = 1 \) are simple [21]. If \( \theta < 1 \), then \( \langle S_{NL} \rangle \approx \theta L, \langle B_L \rangle \approx (1 - \theta)L \). If \( 1 < \theta < 2 \), then \( \langle S_{NL} \rangle \approx L - cL^{2-\theta}/[\theta(2-\theta)(X)] \), and \( \langle B_L \rangle \) follows by difference.

### 7.4. Distribution of \( B_L \)

The distribution of \( B_L \) is obtained by summing (7.1) on the \( k_i \) and on \( n \)

\[
\text{Prob}(B_L = b) = \frac{1}{Z^L(w, L)} g(b) \sum_{n \geq 0} w^n \sum_{\{k_i \geq 1\}} f(k_1) \cdots f(k_n) \delta \left( \sum_{i=1}^{n} k_i + b, L \right).
\]

This entails that

\[
\text{Prob}(B_L = b)_{\text{num}} = \sum_{n \geq 0} w^n \text{Prob}(S_n = L - b)g(b) = Z^{id}(w, L - b)g(b), \tag{7.10}
\]

hence, using (7.6)

\[
\text{Prob}(B_L = b) = \frac{Z^{id}(w, L - b)g(b)}{(g \ast Z^{id})(w, L)}.
\]

The generating function with respect to \( L \) of (7.10) reads

\[
\sum_{L \geq 0} z^L \text{Prob}(B_L = b)_{\text{num}} = \frac{z^b g(b)}{1 - w f(z)}, \tag{7.11}
\]

which summed upon \( b \) gives \( Z^L(w, z) \) back. Note the analogy with (3.15). Taking now the generating function of (7.11) with respect to \( b \) yields

\[
\sum_{L \geq 0} z^L \sum_{b \geq 0} y^b \text{Prob}(B_L = b)_{\text{num}} = \frac{\hat{g}(yz)}{1 - w f(z)}.
\]

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

\[
\sum_{L \geq 0} z^L \langle B_L \rangle_{\text{num}} = \frac{z \hat{g}'(z)}{1 - w f(z)}. \tag{7.12}
\]

### 7.5. Single interval distribution

As for TDRP, the single interval distribution,

\[ p(k|L) = \langle \delta(X_1, k) \rangle, \]

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

\[
p(k|L)_{\text{num}} = \sum_{b \geq 0} g(b) \left[ \sum_{k_1} \delta(k_1, k) w f(k_1) \delta(k_1 + b, L) \right.
\]

\[
+ \sum_{k_1, k_2} \delta(k_1, k) w^2 f(k_1) f(k_2) \delta(k_1 + k_2 + b, L) + \cdots \]

\[
= \sum_{b \geq 0} g(b) \left[ w f(k) \delta(k + b, L) + \sum_{k_2} w^2 f(k) f(k_2) \delta(k + k_2 + b, L) + \cdots \right]
\]

\[
= w f(k) Z^f(w, L - k). \tag{7.13}
\]
So
\[ p(k|L) = \frac{w f(k) Z^f(w, L-k)}{Z^f(w, L)}. \tag{7.14} \]
The generating function of the numerator is therefore
\[ \sum_{L \geq 0} z^L p(k|L)_{\text{num}} = wz^k f(k) Z^f(w, z) = \frac{wz^k f(k) \hat{g}(z)}{1-wf(z)}, \tag{7.15} \]
to be compared to (3.17). Though (7.14) is formally identical to (3.14) the normalisations of these two distributions are different. Indeed,
\[ \sum_{L \geq 0} z^L \sum_{k=1}^L p(k|L)_{\text{num}} = w \hat{f}(z) Z^f(w, z) = \frac{w \hat{f}(z) \hat{g}(z)}{1-wf(z)} = \hat{Z}^f(w, z) - \hat{g}(z), \]
which means that
\[ 1 - \sum_{k=1}^L p(k|L) = \frac{g(L)}{Z^f(w, L)} = p_0(L) = \text{Prob}(B_L = L). \tag{7.16} \]
In other words
\[ \sum_{k=1}^L \text{Prob}(X = k|L) + \text{Prob}(B_L = L) = 1. \]
The distribution \( p(k|L) \) is thus defective. The recursion relation for \( Z^f(w, L) \) follows from (7.13) and (7.16)
\[ Z^f(w, L) = g(L) + \sum_{k=1}^L w f(k) Z^f(w, L-k). \tag{7.17} \]
At the end point, \( k = L \), we have
\[ p(L|L) = \frac{w f(L)}{Z^f(w, L)}. \tag{7.18} \]
which corresponds to the event \( \{X_1 = L, B_L = 0, N_L = 1\} \), and which is formally identical to (3.13).
Both sides of (7.17) are equal to unity if \( w = 1 \). Note that the computation of \( p(k|L) \) for \( w = 1 \) is given in [21], yielding \( p(k|L) = f(k) \) with \( k \leq L \), which shows that this distribution is already defective for \( w = 1 \), with \( \sum_{k=1}^L p(k|L) = 1 - g(L) \).

7.6. Mean interval \( \langle X|L \rangle \)

We proceed as in §3.4. The mean interval is, by definition,
\[ \langle X|L \rangle = \sum_{k \geq 1} kp(k|L). \]
Multiplying (7.15) by \( k \) and summing upon \( k \) yields
\[ \sum_{L \geq 1} z^L \langle X|L \rangle_{\text{num}} = \frac{wz \hat{f}(z) \hat{g}(z)}{1-wf(z)}, \tag{7.19} \]
to be compared to (3.17) for TDRP.
which implies a relation between the generating functions

As for

where the last two terms correspond respectively to the events

In the present case the longest interval is defined as

its distribution function is

In the present case the longest interval is defined as

where

The distribution of

are related by

The distribution of

where

At the end point,

are related by

The distribution of

where

At the end point,

where the last two terms correspond respectively to the events

and

The mean is given by the sum

which implies a relation between the generating functions

|
Condensation for a fluctuating number of independent random variables

where $\tilde{F}^{(1)}(k, z)_{\text{num}}$ is given by (7.20).

Denoting again the restriction of $p^{(1)}(k|L)$ to $k > L/2$ by $g(k|L)$, we can obtain an expression of this quantity by a similar reasoning as done for (3.22) in §3.5. We thus obtain

$$q(k|L) = \text{Prob}(B_L = k) + \frac{w f(k)}{Z^f(w, L)} \sum_{b=0}^{L} g(b)(Z^{\text{id}})_{k-b}^{}(w, L - k - b)$$

$$= \text{Prob}(B_L = k) + \frac{w f(k)(g * Z^{\text{id}})_{w, L - k}}{Z^f(w, L)}, \quad (7.22)$$

expressing the fact that the longest interval can be either $B_L$ or a generic interval $X_i$. For $k = L$, (7.21) with (7.18) are recovered.

To close, we introduce the probability $Q_L$ that the last unfinished interval is the longest one, which is thus

$$Q_L = \text{Prob}(B_L \geq \max(X_1, \ldots, X_{N_L})) = \sum_{n \geq 0} \sum_{b \geq 0} \sum_{k_1=1}^{b} \ldots \sum_{k_n=1}^{b} p\{\{k_i\}, b, n|L\}. \quad (7.23)$$

The generating function with respect to $L$ of its numerator reads

$$\sum_{L \geq 0} z^L Q_L_{\text{num}} = \sum_{b \geq 0} \frac{z^b g(b)}{1 - w f(z, b)},$$

hence

$$\tilde{Z}^f(w, z) - \sum_{L \geq 0} z^L Q_L_{\text{num}} = \frac{\tilde{g}(z)}{1 - w f} - \sum_{b \geq 0} \frac{z^b g(b)}{1 - w f(z, b)}$$

$$= \sum_{b \geq 0} z^b g(b) \left( \frac{1}{1 - w f} - \frac{1}{1 - w f(z, b)} \right).$$

8. Condensed phase for free renewal processes

We now focus on the case of most interest, namely the condensed phase, $w < 1$, for subexponential distributions (1.1). The case $w = 1$ is thoroughly described in [21, 22] and summarised in table 2, which also presents the main outcomes for the $w > 1$ case.

8.1. Asymptotic estimates at large $L$

The asymptotic analysis of (7.5) yields, for large $L$,

$$Z^f(w, L) \approx \frac{g(L)}{(1-w)} \left( 1 - \frac{w^2c(1-\theta)^2}{(1-w)\Gamma(1-2\theta\theta L^{\theta})} \right), \quad \theta < 1$$

$$Z^f(w, L) \approx \frac{g(L)}{(1-w)} \left( 1 + \frac{2w\theta(X)}{(1-w)L} \right), \quad \theta > 1. \quad (8.1)$$

where $g(\cdot)$ is defined in (5.3). As a consequence, (7.8) yields

$$\langle v^{N_L} \rangle \xrightarrow{L \rightarrow \infty} \frac{1-w}{1-vw},$$

leading asymptotically to a geometric distribution for $N_L$,

$$\text{Prob}(N_L = n) \rightarrow p_n = (1-w)w^n,$$
independent of \( \theta \), from which entails, in the same limit,

\[
\langle N_L \rangle \rightarrow \frac{w}{1-w},
\]

and

\[
\left\langle \frac{1}{N_L} \right\rangle \rightarrow -(1-w) \ln(1-w).
\]

The asymptotic estimate of the mean interval can be obtained from the analysis of (7.19),

\[
\langle X|L \rangle \approx wc \Gamma(1-\theta)^2 L^{1-\theta}, \quad \theta < 1,
\]

\[
\langle X|L \rangle \approx w(1+\theta)\langle X \rangle, \quad \theta > 1.
\]

These expressions can also be obtained by means of the marginal distribution \( p(k|L) \), see below. Likewise, the asymptotic estimate of the mean sum can be obtained from the analysis of (7.9),

\[
\langle S_N L \rangle \approx wc \frac{(1-\theta)^2}{1-w} \frac{\Gamma(2-2\theta)}{\Gamma(2-2\theta)} L^{1-\theta}, \quad \theta < 1,
\]

\[
\langle S_N L \rangle \approx w(1+\theta)\langle X \rangle, \quad \theta > 1,
\]

hence

\[
\langle B_L \rangle \approx L - \frac{wc}{1-w} \frac{\Gamma(1-\theta)^2}{\Gamma(2-2\theta)} L^{1-\theta}, \quad \theta < 1,
\]

\[
\langle B_L \rangle \approx L - \frac{w}{1-w} (1+\theta)\langle X \rangle, \quad \theta > 1.
\]

8.2. Regimes for the distribution of the size of a generic interval

We proceed as was done for TDRP. Using (8.1) we have the estimate, at large \( L \),

\[
p(k|L) \approx w(1-w) \frac{f(k)Z^f(w, L-k)}{g(L)}.
\]

(i) For \( k \) finite, using (8.1) again, we have

\[
p(k|L) \approx \frac{wf(k)g(L-k)}{g(L)} \approx wf(k).
\]

(ii) When \( k \) and \( L - k \) are simultaneously large, setting \( k = \lambda L \) in (8.5) \( 0 < \lambda < 1 \) yields the estimate, at large \( L \),

\[
p(k|L) \approx \frac{wf(k)g(L-k)}{g(L)} \approx \frac{1}{\lambda^{1+\theta}(1-\lambda)^{\theta}} \frac{wc}{L^{1+\theta}} \approx \frac{1}{\lambda^{1+\theta}(1-\lambda)^{\theta}} wf(L),
\]

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

(iii) In the region corresponding to \( L - k \) finite, (8.5) simplifies into

\[
p(k|L) \approx \frac{w(1-w)\theta}{L} Z^f(w, L-k).
\]

In particular, for \( k = L \), \( p(L|L) \approx w(1-w)\theta/L \), cf (7.18).

The weight of the downhill region (from 0 to \( k_{\text{min}} \)) is found to be asymptotically equal to \( w \), using the method of §6.2. The complement is borne by \( g(L)/Z^f(w, L) \approx 1-w \), see (7.16) and (8.11). The uphill region therefore does 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.
8.3. Regimes for the distribution of $B_L$

According to (7.10), and in view of (8.1), for $L$ large we have

$$\text{Prob}(B_L = b) \approx (1 - w) \frac{g(b)Z^{td}(w, L - b)}{g(L)}.$$  \hfill (8.7)

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

(i) If $b$ is finite, the asymptotic estimate of $Z^{td}(w, L - b)$ is given by (6.1), hence (8.7) becomes

$$\text{Prob}(B_L = b) \approx \frac{w}{1 - w} \frac{g(b)f(L - b)}{g(L)}.$$  \hfill (8.8)

(ii) If $b \sim L$, the same estimate (8.8) still holds, then setting $L - b = \lambda L$, we get

$$\text{Prob}(B_L = b) \approx \frac{wc}{1 - w} \frac{L^{-\theta}}{\lambda^{1+\theta}(1 - \lambda)^{\theta}},$$  \hfill (8.9)

which has its minimum at $k_{\text{min}} = \frac{L\theta}{1 + 2\theta}$.

(iii) If $L - b$ is finite, (8.7) becomes

$$\text{Prob}(B_L = b) \approx (1 - w)Z^{td}(w, L - b),$$  \hfill (8.10)

in particular, see (7.16),

$$\text{Prob}(B_L = L) = p_0(L) = \frac{g(L)}{Z^{td}(w, L)} \to 1 - w.$$  \hfill (8.11)

Let us estimate, for later use, the probability that $B_L$ is less than $L/2$. The result depends on the value of $\theta$. If $\theta < 1$, using (8.9), we have

$$\text{Prob}(B_L \leq L/2) \approx \frac{wc}{1 - w} \frac{L^{-\theta}}{\lambda^{1+\theta}(1 - \lambda)^{\theta}} \int_{1/2}^{1} \frac{d\lambda}{\lambda^{1+\theta}(1 - \lambda)^{\theta}} \approx \frac{wc}{1 - w} B\left(\frac{1}{2}; 1 - \theta, -\theta\right)L^{-\theta}. \hfill (8.12)$$

If $\theta > 1$, using (8.8), we have

$$\text{Prob}(B_L \leq L/2) \approx \frac{w\theta}{(1 - w)L}\sum_{b=0}^{L/2} g(b) \approx \frac{w\theta}{(1 - w)L}\sum_{b=0}^{L} g(b) \approx \frac{w\theta\langle X\rangle}{(1 - w)L}. \hfill (8.13)$$

8.4. The longest interval

The bulk of the distribution of $X_{\text{max}}$ lies in the region $k > L/2$, and is therefore given by (7.22). We start by giving an illustration. The distributions of $X_{\text{max}}$ and $B_L$ are depicted in figure 10 for Example 1 (see (3.23)). At $k = L$, the two quantities have the same limit $1 - w$, because $p(L|L) \to 0$, see (7.21) and (8.11). The corresponding figure for $\theta > 1$ is qualitatively alike.

Let us now compare the respective contributions of each of the two terms in (7.22) to the total weight in the region $k > L/2$. The first term is investigated in §8.3 above. The asymptotic estimate of the second term is as follows. We first treat the case $\theta < 1$.

(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 $L$

$$(g * Z^{td} * Z^{td})(w, L) \approx \frac{g(L)}{(1 - w)^2} \left(1 - \frac{2wc}{1 - w} \frac{\theta\Gamma(-\theta)^2}{\Gamma(1 - 2\theta)L^{\theta}}\right). \hfill (8.14)$$
Condensation for a fluctuating number of independent random variables

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

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

which is similar to (8.9). Summing this expression upon $\lambda$ from 0 to 1/2 yields (8.12), that is $\text{Prob}(B_L \leq L/2)$. Therefore adding this contribution to the first one, namely $\text{Prob}(B_L > 1/2)$, gives unity, up to small corrections, in agreement with the fact that the weight of $p^{(1)}(k|L)$ in the left domain $k < L/2$ is negligible.

(ii) If $L - k$ is finite, then the second term reads

$$\frac{wf(k)(g \ast Z^{td} \ast Z^{td})(w, L - k)}{Z^I(w, L)} \approx \frac{(1 - w)w\theta}{L} (g \ast Z^{td} \ast Z^{td})(w, L - k),$$

which is subdominant compared to (8.10).

In order to get the correction of $\langle X_{\text{max}} \rangle$ to $L$, we take the average of (7.22), by integrating each of the terms from 0 to 1/2 upon $\lambda$, using (8.9) and (8.15). Adding the contributions coming from the two terms, we finally obtain for the dominant correction,

$$L - \langle X_{\text{max}} \rangle \approx \frac{wc}{1 - w} \left( B\left(\frac{1}{2}; 1 - \theta, 1 - \theta \right) + B\left(\frac{1}{2}; 2 - \theta, -\theta \right) \right) L^{1 - \theta}$$

$$= \frac{wc}{1 - w} B\left(\frac{1}{2}; 1 - \theta, -\theta \right) L^{1 - \theta},$$

which has the same structure as (6.15) or (2.20).

Figure 10. Free renewal processes: distributions of $B_L$ and $X_{\text{max}}$ for Example 1 (see (3.23)), with $w = 0.6$ and $L = 60$. The discrepancy between the two curves, for $k > L/2$ is predicted in (7.22).
Likewise, for $\theta > 1$, the weight of the first term in (7.22) dominates upon the second one, and we find for the correction of the mean to $L$,

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

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 (6.16) or (2.21).

To close, we investigate the behaviour of $Q_L$, the probability that $B_L$ is the longest interval, defined in (7.23). An estimate of $Q_L$ for $w < 1$ can be obtained by means of the inequality

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

In view of (8.12) and (8.16), we infer that asymptotically for $L$ large, if $\theta < 1$,

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

while, if $\theta > 1$, in view of (8.13) and (8.17),

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

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 $\theta > 1$, $Q_L \sim L^{-(1-1/\theta)}$ [22]. For $w > 1$, $Q_L \to 0$.

Figure 11 depicts $Q_L$ as a function of $w$ for Example 1 (see (3.23)) and for three different sizes, crossing at the universal critical value $Q_\infty = 0.626 \ldots$ for $w = 1$ [22] and the data collapse obtained by using the scaling variable $x = (w - 1)L^{1/2}$.

Figure 11. Free renewal processes: probability $Q_L$ that the last interval, $B_L$, is the longest one, for Example 1 (see (3.23)), 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 8 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 in the different phases. The results in columns 2 and 3 (critical phase) are taken from [21, 22]. In the last column, $\langle X|L \rangle \sim L^{1-\theta}$ if $\theta < 1$, or $\langle X|L \rangle \approx \text{constant}$ if $\theta > 1$.

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

9. Conclusion

Let us summarise the salient aspects of this study.

We first recalled the main features of the condensation transition taking place in the thermodynamical 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 thermodynamical 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 $X_{\text{max}}$ is the unique summand with extensive value. In the thermodynamical limit, the fraction $X_{\text{max}}/L$ no longer fluctuates. Its value is given by the difference $\rho - \rho_c$.

A second scenario for the same class of models consists in taking the limit $L \to \infty$ keeping the number of sites (or summands) fixed. In such a situation, strictly speaking, there is no phase transition, since there is no tuning parameter allowing to move from one phase to another. Nevertheless, in this limit, the system condenses. There is again a single extensive summand $X_{\text{max}}$, but now the fraction $X_{\text{max}}/L$ 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 $\theta$ is smaller or larger than one. If $L$ is large but finite, the distribution of $X_{\text{max}}$ is peaked, with a width $L - \langle X_{\text{max}} \rangle$ scaling as $L^{1-\theta}$, if $\theta < 1$, and asymptotically equal to $(n-1)\langle X \rangle$, if $\theta > 1$.

This scenario is a good preparation for the study of condensation in free and tied-down renewal processes, 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 \cite{9, 13, 14}. 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, $\rho$, but a change in the value of the weight $w$ attached to each interval (or summand). In this respect it is worth noting the striking similarity between (2.20), (6.15) and (8.16) on one hand, and between (2.21), (6.16) and (8.17) on the other hand.

It turns out that, in the $L \to \infty$ limit, the distribution of the number of intervals, $N_L$, is completely universal, i.e., model independent, since it only depends on $w$ and not on the index $\theta$. 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 $\theta$ 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.

The phase transition occurring when $w$ passes through unity is second order for the density of intervals $\nu$ if $\theta < 1$, and first order if $\theta > 1$. On the other hand the correlation length diverges at the transition, see (4.4). The transition is therefore mixed order as was pointed out in \cite{11, 12} for the particular case of TDRP with Example 2, see (3.26). 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 $\pm 1$ in the condensed phase since condensation is total. More on this can be found in \cite{12}. If $\theta < 1$ the distribution of the magnetisation at criticality is broad and self-similar, both for free \cite{25, 21} and tied-down renewal processes \cite{24}. At criticality, for $\theta < 1$, the non stationary two-time (or two-space) correlation function is also self-similar, again for both processes \cite{21, 24}.

**Acknowledgments**

It is a pleasure to thank G Giacomin and J-M Luck for enlightening discussions.

**References**

[1] Bialas P, Burda Z and Johnston D 1997 *Nucl. Phys. B* **493** 505
[2] Drouffe JM, Godrèche C and Camia F 1998 *J. Phys. A* **31** L19
[3] Evans MR and Hanney T 2005 *J. Phys. A* **38** R195
[4] Godrèche C 2007 *Lect. Notes Phys.* Springer **716** 261
[5] Feller W 1968 1971 *An Introduction to Probability Theory and its Applications* volumes 1&2 (New York: Wiley)
[6] Cox DR 1962 *Renewal theory* (London: Methuen)
[7] Wendel JG 1964 *Math. Scand.* **14** 21
[8] Godrèche C 2017 *J. Phys. A* **50** 195003
[9] Fisher M 1984 *J. Stat. Phys.* **34** 667
[10] Poland D and Scheraga HA 1966 *J. Chem. Phys.* **45** 1456
[11] Bialas P, Burda Z and Johnston D 1999 *Nucl. Phys. B* **542** 413
Condensation for a fluctuating number of independent random variables

[12] Bar A and Mukamel D 2014 J. Stat. Mech. P11001
[13] Giacomin G 2007 Random polymer models (Imperial College Press: World Scientific)
[14] Giacomin G 2011 Disorder and critical phenomena through basic probability models Lect. Notes
  Math. Springer
[15] Godrèche C 2019 J. Stat. Mech. 063207
[16] Janson S 2012 Prob. Surveys 9 103
[17] Ferrari PA, Landim C and Sisko VV 2007 J. Stat. Phys. 128 1153
[18] Godrèche C and Luck JM 2001 Eur. Phys. J. B 23 473
[19] Bar A, Majumdar SN, Schehr G and Mukamel D 2016 Phys. Rev. 93 052130
[20] Barma M, Majumdar SN and Mukamel D 2019 J. Phys. A 52 254001
[21] Godrèche C and Luck JM 2001 J. Stat. Phys. 104 489
[22] Godrèche C Majumdar SN and Schehr G 2015 J. Stat. Mech. P03014
[23] Chistyakov V P 1964 Theor. Probab. Appl. 9 640
[24] Godrèche C 2017 J. Stat. Mech. P073205
[25] Lamperti J 1958 Trans. Amer. Math. Soc. 88 380