



Jean Bertoin

The asymptotic behavior of fragmentation processes

Received November 10, 2001 / final version received June 30, 2003

Published online September 26, 2003 – © Springer-Verlag & EMS 2003

Abstract. The fragmentation processes considered in this work are self-similar Markov processes which are meant to describe the evolution of a mass that falls apart randomly as time passes. We investigate their pathwise asymptotic behavior as $t \rightarrow \infty$. In the so-called homogeneous case, we first point at a law of large numbers and a central limit theorem for (a modified version of) the empirical distribution of the fragments at time t . These results are reminiscent of those of Asmussen and Kaplan [3] and Biggins [12] for branching random walks. Next, in the same vein as Biggins [10], we also investigate some natural martingales, which open the way to an almost sure large deviation principle by an application of the Gärtner-Ellis theorem. Finally, some asymptotic results in the general self-similar case are derived by time-change from the previous ones. Properties of size-biased picked fragments provide key tools for the study.

Key words. fragmentation – self-similar – central limit theorem – large deviations

1. Introduction

It seems that the first probabilistic works on random fragmentations were due to Kolmogorov [19] himself and one of his students, Filippov [18]. More precisely, Kolmogorov provided an explanation to the observation that the logarithms of the sizes of mineral grains are often normally distributed. In this direction, he considered the following Markov chain. At the initial time we have a mass, say $m_0 \in]0, \infty[$. At time 1, this mass is broken randomly, which produces two smaller masses, say $m_{1,1} = V_0 m_0$ and $m_{1,2} = (1 - V_0) m_0$, where V_0 is a random variable with values in $[1/2, 1[$ which has a given distribution. The next steps consist in independent iterations, in the sense that at time $n + 1$, each of the 2^n masses obtained at time n , say $m_{1,n}, \dots, m_{2^n,n}$, is again randomly broken into two pieces, independently of the past. More precisely, the split uses independent variables $V_{1,n}, \dots, V_{2^n,n}$ which are all distributed as V_0 . Roughly, Kolmogorov showed that there is a simple rescaling of the empirical distribution of $\frac{1}{n} \ln m_{1,n}, \dots, \frac{1}{n} \ln m_{2^n,n}$ which converges with probability one to the standard Gaussian distribution.

J. Bertoin: Laboratoire de Probabilités et Modèles Aléatoires et Institut universitaire de France, Université Pierre et Marie Curie, et C.N.R.S. UMR 7599, 175, rue du Chevaleret, F-75013 Paris, France, e-mail: jbe@ccr.jussieu.fr

Mathematics Subject Classification (2000): 60J25, 60G09

In the recent years, there has been a renewed and growing interest amongst probabilists in models of random splittings and the dual models of coalescence. We refer to Aldous' survey [1] for motivations, connections with other scientific areas, references, ... The present work is concerned with certain random fragmentation processes in continuous time that we now introduce.

The state space \mathcal{S}^\downarrow consists in decreasing numerical sequences $x = (x_1, \dots)$ with

$$x_1 \geq x_2 \geq \dots \geq 0 \quad \text{and} \quad \sum_1^\infty x_i \leq 1.$$

An element $x \in \mathcal{S}^\downarrow$ should be thought of as the ranked family of masses arising from the split of some object with unit mass. A portion of the initial mass may be lost during the split, which corresponds to the situation when $\sum_1^\infty x_i < 1$. In [6, 7], we considered a class of Markov processes $X = (X(t), t \geq 0)$ with values in \mathcal{S}^\downarrow , called *self-similar fragmentations*, which fulfill both the so-called *fragmentation* and the *scaling* properties. The first can be viewed as a version of the branching property. Specifically, let \mathbb{P}_r stand for the law of X started from $(r, 0, \dots)$, i.e. at the initial time, there is a single fragment with mass r . Then for every $s, t \geq 0$, conditionally on $X(t) = (x_1, \dots)$, $X(t+s)$ has the same law as the variable obtained by ranking in the decreasing order the terms of the random sequences $X^{(1)}(s), X^{(2)}(s), \dots$, where the latter are independent variables with values in \mathcal{S}^\downarrow , such that $X^{(i)}(s)$ has the same distribution as $X(s)$ under \mathbb{P}_{x_i} for each $i = 1, \dots$. The scaling property means that there is an index of self-similarity $\alpha \in \mathbb{R}$ such that for every $r \in [0, 1]$, the distribution of $(rX(r^\alpha t), t \geq 0)$ under \mathbb{P}_1 is \mathbb{P}_r . In the special case when $\alpha = 0$, we will call the fragmentation *homogeneous* rather than self-similar with index 0.

A simple example of such a self-similar fragmentation with index $\alpha > 0$ was studied by Brennan and Durrett [14, 15] as a model for polymer degradation. Specifically, they considered a particle system in which particles are specified by their sizes and the evolution of the system is described by the following dynamics. A particle with size $r > 0$ waits an exponential time with parameter r^α and then splits into two particles with respective sizes rV and $r(1-V)$, where V is a random variable with values in $[1/2, 1[$ which has a fixed distribution and is independent of the past of the system. This fragmentation is binary and discrete, in the sense that each split produces two new particles, and splitting times of the system can only accumulate at ∞ . More recently, motivated by the analysis of the so-called standard additive coalescent, Aldous and Pitman [2] logged the continuum random tree along its skeleton, and constructed this way a self-similar fragmentation with index $\alpha = 1/2$ which is binary but not discrete. We further refer to [5] for an alternative construction of the same fragmentation from the Brownian excursion, to [7] for another binary and non discrete self-similar fragmentation with negative index $\alpha = -1/2$, which is also naturally connected to the Brownian excursion, and to [23, 24] for other more recent related examples which are based on stable continuum random trees.

In short, it has been shown in [6, 7] that the dynamics of a self-similar fragmentation are entirely characterized by the index of self-similarity, a so-called

splitting measure ν which specifies the rate at which sudden dislocations occur, and an erosion rate $c \geq 0$ which accounts for the continuous component of the process. For the sake of simplicity, we shall focus in this work on the case when the erosion rate is zero, which induces little loss of generality. For instance, in the model of Brennan and Durrett described above, the index of self-similarity is α , the erosion rate $c = 0$ and the splitting measure ν given by the distribution of the sequence $(V, 1 - V, 0, \dots)$. Excluding implicitly the degenerate case when X remains constant, it can be seen that when t goes to infinity, $X(t) = (X_1(t), \dots)$ converges a.s. to $(0, \dots)$, say in the sense of the uniform distance on \mathcal{S}^\downarrow , and the purpose of this work is to investigate the asymptotic behavior of the fragmentation in terms of its characteristics.

In this direction, note that in the homogeneous case $\alpha = 0$, the discrete-time skeleton of the logarithm of the fragmentation, $(\log X_1(n), \dots)$, for $n = 0, 1, \dots$, is a special case of a branching random walk in the sense of e.g. Biggins [10–12]. A great deal of information on the asymptotic behavior of branching random walks is available in the literature. Typically, in the case when the number $Z(n)$ of fragments at time n is finite a.s., $Z(\cdot)$ is a Galton-Watson process, and it can be deduced from the work of Asmussen and Kaplan [3] that the following pathwise central limit theorem holds. Under suitable conditions there are two real numbers μ and $\sigma^2 > 0$, such that the distribution function of the rescaled empirical measure at time n ,

$$y \rightarrow \frac{1}{Z(n)} \text{Card} \{i : \log X_i(n) \in] - \infty, -\mu n + \sigma \sqrt{n}y]\}$$

converges a.s. to the normal distribution when $n \rightarrow \infty$, which can also be viewed as an extension of Kolmogorov [19] that we mentionned at the beginning of this introduction. Unfortunately, it is not clear how to extend such a result in our setting, because in the most interesting cases, the number of fragments is always infinite. Moreover the instants when splitting occur in fragmentation processes are in general everywhere dense a.s., and the discrete arguments used in the study of branching random walks cannot be applied.

We shall first study the asymptotic behavior of fragmentation in the homogeneous case. In this direction, it is convenient to consider the following modification of the empirical distribution of the fragments:

$$\sum_{i=1}^{\infty} X_i(t) \delta_{X_i(t)}(dy),$$

where $\delta_r(dy)$ stands for the Dirac point mass at r . Roughly speaking, fragments decay exponentially fast as time passes, which incites us to introduce the image of the preceding random measure by the map $y \rightarrow t^{-1} \log y$, i.e. we set

$$\rho_t(dy) = \sum_{i=1}^{\infty} X_i(t) \delta_{t^{-1} \log X_i(t)}(dy).$$

By considering the evolution of the size of the fragment containing a randomly tagged point, we shall first easily obtain a law of large numbers and a (pathwise)

central limit theorem. Next, we turn our attention to the large deviations for ρ_t , let us sketch our approach. For every $p > 0$, we consider

$$\Sigma(t, p) := \sum_{i=1}^{\infty} X_i^{p+1}(t) = \int_{\mathbb{R}} e^{tpy} \rho_t(dy).$$

The key is that there is a real number $\Phi(p)$, which is given explicitly in terms of the splitting measure ν , such that $e^{t\Phi(p)} \Sigma(t, p)$ is a uniformly integrable martingale, at least for certain values of p . This can be viewed as an analog in continuous time of the well-known result of Biggins [10] for branching random walks. By the convergence theorem for martingales, this entails the asymptotic

$$\int_{\mathbb{R}} e^{tpy} \rho_t(dy) \asymp e^{-t\Phi(p)} \quad \text{as } t \rightarrow \infty, \text{ a.s.}$$

which in turn opens the way to the analysis of large deviations for ρ_t via an application of the Gärtner-Ellis theorem. In particular, this enables us to estimate for large times t the number of fragments with mass of order e^{at} for some fixed $a < 0$.

Next we shall turn our attention to the case of positive indices of self-similarity, that is $\alpha > 0$. Our main result in this setting extends a theorem due to Filippov [18], which in turn was re-discovered in a special case by Brennan and Durrett [15]. It states (under certain conditions) the convergence in probability as $t \rightarrow \infty$ of the random measures

$$\rho_t^{(\alpha)}(dy) = \sum_{i=1}^{\infty} X_i(t) \delta_{t^{1/\alpha} X_i(t)}(dy)$$

towards some deterministic measure on $]0, \infty[$ that is specified in terms of the characteristics of the fragmentation. Note that the rate of decay of fragments is now polynomial and not exponential as in the homogeneous case. Finally, the case of a negative index of self-similarity is studied via the homogeneous case, using an invertible transformation which changes a self-similar fragmentation into a homogeneous one. We shall show that the entire mass disappears in a finite time, in the sense that

$$\inf \{t \geq 0 : X(t) = (0, \dots)\} < \infty \quad \text{a.s.}$$

If further $\alpha < -1$, then we shall also point out that at each fixed time $t > 0$, there are only a finite number of fragments with positive size, a.s.

At the heart of this work lies the fact that many important statistical features of the fragmentation can be derived by the study of a few specific fragments (often the fragment that contained a point tagged at random, independently of the process). In particular, the analysis can be developed without keeping track simultaneously of the evolution of each fragment, which would clearly be much more difficult. The rest of this paper is organized as follows. The next section is devoted to the analysis of the homogeneous case $\alpha = 0$. The self-similar cases $\alpha > 0$ and $\alpha < 0$ are discussed in Sect. 3.

2. Homogeneous fragmentations

Throughout this section, we focus on homogeneous fragmentations, i.e. with index of self-similarity $\alpha = 0$. We will first develop some preliminaries, then establish a law of large numbers and a central limit theorem. Next we will investigate the so-called scattering rates via some remarkable martingales, and deduce large deviation estimates. We shall further assume that the fragmentation X starts from a single fragment with unit mass, i.e. we implicitly work under $\mathbb{P} := \mathbb{P}_1$.

2.1. Preliminaries

We first recall some features lifted from [4,6], and also refer to Pitman [25] and Schweinsberg [26] for related works. Roughly, homogeneous fragmentations result from the combination of two different phenomena: a continuous erosion and sudden dislocations. The erosion is a continuous deterministic mechanism, whereas the dislocations occur randomly and can be viewed as the jump-component of the process. More precisely, if X is a pure-jump homogeneous fragmentation and $c > 0$ an arbitrary real number, then $(e^{-ct} X(t), t \geq 0)$, is again a homogeneous fragmentation, but now with a continuous component corresponding to an erosion. Any homogeneous fragmentation can be obtained from a pure-jump one by this elementary transformation. For the sake of simplicity, and because adding an erosion term to a homogeneous fragmentation is straightforward, **we will only consider homogeneous fragmentations with no erosion (i.e. $c = 0$) in the sequel.**

Let us now focus on the sudden dislocations (i.e. the jumps). Just like in the celebrated Lévy-Itô decomposition for subordinators, their distribution can be described in terms of a certain intensity measure. More precisely, set $\mathcal{S}^* = \mathcal{S}^\downarrow \setminus \{(1, 0, \dots)\}$ for the space of decreasing numerical sequences $x = (x_1, \dots)$ with $\sum x_i \leq 1$ and $x_1 < 1$. The splitting measure (also called Lévy measure in [6]) is a measure ν on \mathcal{S}^* which fulfills the condition

$$\int_{\mathcal{S}^*} (1 - x_1) \nu(dx) < \infty. \tag{1}$$

One can construct a Poisson point process $((\Delta(t), k(t)), t \geq 0)$ with values in $\mathcal{S}^* \times \mathbb{N}$, with characteristic measure $\nu \otimes \#$, where $\#$ stands for the counting measure on $\mathbb{N} = \{1, 2, \dots\}$, such that the following holds. The process $X(\cdot)$ jumps only at times $t \geq 0$ at which a point $(\Delta(t), k(t))$ occurs, and then $X(t)$ is obtained from $X(t-)$ by replacing its $k(t)$ -th term $X_{k(t)}(t-)$ by the sequence $X_{k(t)}(t-)\Delta(t)$, and ranking all the terms in the decreasing order. Of course, it may happen that $X_{k(t)}(t-) = 0$, and in that case we have $X(t) = X(t-)$. In the sequel, we shall always implicitly exclude the trivial case when $\nu \equiv 0$. More significantly, **we will always assume that**

$$\sum_{i=1}^{\infty} x_i = 1 \quad \text{for } \nu\text{-almost every } x \in \mathcal{S}^*,$$

which means that no mass can be lost when a sudden dislocation occurs. This restriction is really needed only in some statements, but we prefer to make it once for all for the sake of simplicity.

Some information about the splitting law ν , and hence on the distribution of the homogeneous fragmentation X , is captured by the function

$$\Phi(q) := \int_{\mathcal{S}^*} \left(1 - \sum_{i=1}^{\infty} x_i^{q+1} \right) \nu(dx), \quad q > \underline{p}, \tag{2}$$

where

$$\underline{p} := \inf \left\{ p \in \mathbb{R} : \int_{\mathcal{S}^*} \sum_{i=2}^{\infty} x_i^{p+1} \nu(dx) < \infty \right\} \tag{3}$$

(recall the requirement (1) and note that the sum above starts from $i = 2$). Clearly we must have $-1 \leq \underline{p} \leq 0$. Using the convention $0 \times \infty = 0$, observe that we may also express (2) in the form

$$\Phi(q) = \int_{\mathcal{S}^*} \left(\sum_{i=1}^{\infty} x_i (1 - x_i^q) \right) \nu(dx).$$

The function Φ is positive on $]0, \infty[$ and negative on $] \underline{p}, 0[$. It has a crucial role in this work, which essentially stems from the following fact.

There is no loss of generality in assuming that the fragmentation is built from some nested family of random open subsets of $]0, 1[$, see Sect. 3.2 in [7]. This means that we can construct for each $t \geq 0$ a random open set $G(t) \subseteq]0, 1[$ with $G(t') \subseteq G(t)$ for every $0 \leq t \leq t'$, and such that $X(t) = (X_1(t), \dots)$ coincides with the ordered sequence of the lengths of the component intervals of $G(t)$. Suppose now that at the initial time, a point is tagged at random uniformly in $]0, 1[$ and independently of the $G(t)$'s. Denote by $\chi(t)$ the size of the interval component of $G(t)$ that contains this tagged point at time $t \geq 0$. In particular, $\chi(t)$ is a size-biased pick from the sequence $X(t) = (X_1(t), \dots)$, that is we have the identity in law

$$\chi(t) \stackrel{\mathcal{L}}{=} X_K(t),$$

where K is an integer valued variable whose conditional distribution given $X(t)$ is

$$\mathbb{P}(K = k \mid X(t)) = X_k(t), \quad k = 1, \dots.$$

Then the process

$$\xi(t) = -\log \chi(t), \quad t \geq 0$$

is a subordinator, i.e. it is an increasing process with independent and stationary increments, and we have

$$\mathbb{E}(\chi(t)^q) = \mathbb{E}(\exp(-q\xi(t))) = \exp(-t\Phi(q)), \quad t \geq 0. \tag{4}$$

See Sect. 5 in [6], and in particular Theorem 3 and Lemma 3 there, where (4) is established for every $q \geq 0$ (of course, the extension to $q < 0$ follows by

analyticity). In the sequel, we shall often use the fact that, as $\Phi :]p, \infty[\rightarrow]-\infty, \infty[$ is the Laplace exponent of a subordinator, it is a concave increasing analytic function.

We stress that $\Phi(0+) = 0$, so the subordinator ξ has no killing, and $\chi(t) > 0$ a.s. for every $t \geq 0$. This implies that the homogeneous fragmentation preserves the total mass, i.e.

$$\mathbb{P} \left(\sum_{i=1}^{\infty} X_i(t) = 1 \right) = 1 \quad \text{for all } t \geq 0.$$

2.2. Law of large numbers and central limit theorem

Because total mass is preserved, it will be convenient here to represent the sequence $X(t)$ through the random probability measure

$$\rho_t(dy) := \sum_{i=1}^{\infty} X_i(t) \delta_{t^{-1} \log X_i(t)}(dy). \tag{5}$$

In words, ρ_t assigns a mass y at each point $t^{-1} \log y$ such that y is the size of a fragment of $X(t)$ (with the obvious convention in case of multiple fragments with identical size). Note also that ρ_t coincides with the conditional distribution of $t^{-1} \log \chi(t)$ given the fragmentation process, where $\chi(t)$ is a size-biased pick from $X(t)$.

We consider the ρ_t 's as random variables with values in the space of probability measures on \mathbb{R} , which is endowed with Prohorov's distance (so limits are taken in the sense of weak convergence of probability measures). The simplest results on the asymptotic behavior of homogeneous fragmentations are easy consequences of the classical law of large numbers and central limit theorem. In this direction, introduce the first and second right-derivatives of Φ at 0,

$$\mu := \Phi'(0+), \quad \sigma^2 := -\Phi''(0+);$$

so that in terms of the splitting measure

$$\mu = - \int_{\mathcal{S}^*} \left(\sum_{i=1}^{\infty} x_i \log x_i \right) \nu(dx), \quad \sigma^2 = \int_{\mathcal{S}^*} \left(\sum_{i=1}^{\infty} x_i (\log x_i)^2 \right) \nu(dx).$$

Theorem 1. (i) *Suppose that $\mu < \infty$. Then,*

$$\lim_{t \rightarrow \infty} \rho_t = \delta_{-\mu},$$

in probability.

(ii) *Suppose that $\sigma^2 < \infty$, and denote by $\tilde{\rho}_t$ the image of ρ_t by the map $x \rightarrow \sqrt{t}(x + \mu)/\sigma$ and by $\mathcal{N}(0, 1)$ the standard normal distribution, then*

$$\lim_{t \rightarrow \infty} \tilde{\rho}_t = \mathcal{N}(0, 1),$$

in probability.

Proof. (i) As it has been mentioned in the preceding section, we may assume that $X(t)$ is the sequence of the lengths of the interval components of an open set $G(t)$, where the family $(G(t), t \geq 0)$ is nested. Introduce two independent random variables, U and U' which are both uniformly distributed on $]0, 1[$ and independent of the fragmentation process. The lengths $\chi(t)$ and $\chi'(t)$ of the interval components of $G(t)$ that contain U and U' respectively, have both the same distribution and are not independent.

On the one hand, for an arbitrary bounded continuous function $f :]0, \infty[\rightarrow \mathbb{R}$, we have the first and second moments identities

$$\mathbb{E} \left(\int_{]0, \infty[} f(y) \rho_t(dy) \right) = \mathbb{E}(f(t^{-1} \log \chi(t)))$$

$$\mathbb{E} \left(\left(\int_{]0, \infty[} f(y) \rho_t(dy) \right)^2 \right) = \mathbb{E}(f(t^{-1} \log \chi(t)) f(t^{-1} \log \chi'(t))).$$

Recall that $-\log \chi(t)$ and $-\log \chi'(t)$ are both distributed as $\xi(t)$, where $\xi = (\xi(t), t \geq 0)$ is a subordinator with Laplace exponent Φ . The hypothesis that $\Phi'(0+) = \mu < \infty$ ensures the (weak) law of large numbers for ξ , i.e. $t^{-1}\xi(t)$ converges in probability as $t \rightarrow \infty$ to μ . We conclude that

$$\lim_{t \rightarrow \infty} \mathbb{E} \left(\int_{\mathbb{R}} f(y) \rho_t(dy) \right) = f(-\mu), \quad \lim_{t \rightarrow \infty} \mathbb{E} \left(\left(\int_{\mathbb{R}} f(y) \rho_t(dy) \right)^2 \right) = f(-\mu)^2,$$

and therefore

$$\lim_{t \rightarrow \infty} \int_{\mathbb{R}} f(y) \rho_t(dy) = \int_{\mathbb{R}} f(y) \delta_{-\mu}(dy) \quad \text{in } L^2(\mathbb{P}).$$

This readily entails our first claim.

(ii) The same argument as above yields the following identities for every bounded continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$

$$\mathbb{E} \left(\int_{]0, \infty[} f(y) \tilde{\rho}_t(dy) \right) = \mathbb{E} \left(f \left(\frac{\log \chi(t) + \mu t}{\sigma \sqrt{t}} \right) \right)$$

$$\mathbb{E} \left(\left(\int_{]0, \infty[} f(y) \tilde{\rho}_t(dy) \right)^2 \right) = \mathbb{E} \left(f \left(\frac{\log \chi(t) + \mu t}{\sigma \sqrt{t}} \right) f \left(\frac{\log \chi'(t) + \mu t}{\sigma \sqrt{t}} \right) \right).$$

Our assumption ensures the central limit theorem for the subordinator $\xi(t) = -\log \chi(t)$, viz.

$$\frac{\log \chi(t) + \mu t}{\sigma \sqrt{t}} \text{ converges in law as } t \rightarrow \infty \text{ towards } N,$$

where N has the $\mathcal{N}(0, 1)$ -distribution, and of course the same holds when we replace $\chi(t)$ by $\chi'(t)$. The key observation is that, although $\chi(t)$ and $\chi'(t)$ are not independent, the fragmentation property entails that the variables $t^{-1/2}(\log \chi(t) + \mu t)$ and $t^{-1/2}(\log \chi'(t) + \mu t)$ are asymptotically independent. Indeed, the first instant t when the tagged points U and U' are disjoint (i.e. belong to two different

intervals components of $G(t)$ is an a.s. finite randomized stopping time for which the strong Markov property thus applies. This readily yields the stated asymptotic independence.

Putting the pieces together, we see that

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{E} \left(\int_{\mathbb{R}} f(y) \tilde{\rho}_t(dy) \right) &= \mathbb{E}(f(N)), \\ \lim_{t \rightarrow \infty} \mathbb{E} \left(\left(\int_{\mathbb{R}} f(y) \tilde{\rho}_t(dy) \right)^2 \right) &= (\mathbb{E}(f(N)))^2, \end{aligned}$$

and we conclude that for every bounded continuous function f ,

$$\lim_{t \rightarrow \infty} \int_{\mathbb{R}} f(y) \tilde{\rho}_t(dy) = \mathbb{E}(f(N)) \quad \text{in } L^2(\mathbb{P}).$$

The convergence in probability $\tilde{\rho}_t$ to $\mathcal{N}(0, 1)$ now follows from standard arguments. □

To end this section, we point out that the same techniques apply as well to establish an extension of Theorem 1 in some cases when $\Phi'(0+) = -\infty$ (which impedes the application of the law of large numbers). In order to state such an extension, denote for every $\beta \in]0, 1[$ by $s_\beta(dy)$ the standard Stable(β)-distribution on $] - \infty, 0[$, i.e. with Laplace transform

$$\int_{-\infty}^0 e^{qy} s_\beta(dy) = \exp(-q^\beta), \quad q > 0.$$

Recall also that a function $f :]0, \infty[\rightarrow]0, \infty[$ varies regularly at $0+$ with index γ if for every $c > 0$ we have $\lim_{x \rightarrow 0+} f(cx)/f(x) = c^\gamma$.

Proposition 1. *Suppose that for some $\beta \in]0, 1[$, the function Φ varies regularly at $0+$ with index β , and denote by $\varphi :]0, \infty[\rightarrow]0, \infty[$ the inverse of the bijection Φ . Then, as t tends to ∞ , the random probability measures*

$$\sum_{i=1}^{\infty} X_i(t) \delta_{\varphi(1/t) \log X_i(t)}(dy)$$

converge to $s_\beta(dy)$, in probability.

Proof. The hypothesis enables us to apply Lévy’s stable limit theorem. More precisely, in the notation of the proof of Theorem 1, the variables $\varphi(1/t) \log \chi(t)$ converge in distribution to the stable law $s_\beta(dy)$. We can then follow the same argument as in the proof of Theorem 1. □

Similarly, the same method applies when $\Phi'(0+) = \mu < \infty$ and $p \rightarrow \Phi(p) - p\mu$ varies regularly at $0+$ with index $\beta \in]1, 2[$ (so in that case $\Phi''(0+) = \infty$, which impedes the application of the central limit theorem). One then obtains an extension of Theorem 1 involving the completely skewed stable distribution with index β .

2.3. Scattering rates and large deviations

For every real number $p > -1$ and $t \geq 0$, we introduce

$$\Sigma(t, p) := \sum_{i=1}^{\infty} X_i^{p+1}(t),$$

which can be thought of as a measurement of the scattering of the fragmentation at time t . An interesting question is thus to investigate its behavior when $t \rightarrow \infty$. In this direction, recall the notation introduced in Sect. 2.1.

We observe that if $(\mathcal{F}_t)_{t \geq 0}$ denotes the natural filtration of the fragmentation process, then the conditional moments of the size-biased picked fragment $\chi(t)$ are given by

$$\mathbb{E}(\chi^p(t) | \mathcal{F}_t) = \Sigma(t, p),$$

We may thus reformulate the identity (4) as

$$\mathbb{E}(\Sigma(t, p)) = \exp(-t\Phi(p)), \quad p > \underline{p}. \tag{6}$$

We then immediately deduce from the Markov, scaling and fragmentation properties that

$$M(t, p) := \exp(t\Phi(p)) \Sigma(t, p), \quad t \geq 0$$

is a positive martingale¹. This can be viewed as the analog in continuous time of a family of remarkable martingales associated to branching random walks, and our main result in the present setting will extend a well-known theorem due to Biggins [10]. In this direction, we need to introduce one more notation, which is the purpose of the next lemma.

Lemma 1. *The function $p \rightarrow \Phi(p)/(p + 1)$ reaches its maximum on $]\underline{p}, \infty[$ at a unique point, $\bar{p} > 0$, which is the unique solution to the equation*

$$(p + 1)\Phi'(p) = \Phi(p), \quad p > \underline{p}.$$

More precisely, the function $p \rightarrow \Phi(p)/(p + 1)$ increases on $]\underline{p}, \bar{p}[$ and decreases on $]\bar{p}, \infty[$, and this maximum $m := \max\{\Phi(p)/(p + 1), p > \underline{p}\}$ is given by

$$m = \Phi'(\bar{p}) = \Phi(\bar{p})/(\bar{p} + 1).$$

Proof. We first point out that

$$\text{the function } p \rightarrow (p + 1)\Phi'(p) - \Phi(p) \text{ decreases on }]\underline{p}, \infty[. \tag{7}$$

Indeed, this function has derivative $(p + 1)\Phi''(p)$, which is negative since Φ is concave. Recall that $\Phi(0) = 0$. On the other hand, it is well-known that the so-called drift coefficient is given by $\lim_{q \rightarrow \infty} \Phi(q)/q$, and we easily see from (2) that

¹ Alternatively, this can be derived by an optional projection on the natural filtration of the fragmentation process, from the well-known fact that if ξ is a subordinator with Laplace exponent Φ , then $\exp(-q\xi(t) + t\Phi(q))$ is a martingale. This observation lies at the heart of the recent work [9] in which the so-called conceptual approach of Lyons, Pemantle and Peres [22] is adapted to homogeneous fragmentations.

this quantity is zero. Hence the function $p \rightarrow \Phi(p)/(p + 1)$ has the same limit at 0 and at ∞ , so it reaches its overall maximum at a unique point $\bar{p} > \underline{p}$. In particular, we deduce from (7) that the derivative of $p \rightarrow \Phi(p)/(p + 1)$ is positive on $]\underline{p}, \bar{p}[$ and negative on $]\bar{p}, \infty[$. Finally, the derivative must be zero at \bar{p} , which entails that the overall maximum is given by $\Phi'(\bar{p}) = \Phi(\bar{p})/(\bar{p} + 1)$.

We may now state the following.

Theorem 2. *For every $p \in]\underline{p}, \bar{p}[$, the martingale $M(\cdot, p)$ is uniformly integrable and its terminal value, $M(\infty, p)$, is a strictly positive random variable a.s.*

Theorem 2 and its following corollary will be established in the next section. Recall (5) and observe the obvious identity

$$\int_{\mathbb{R}} e^{tpy} \rho_t(dy) = \Sigma(t, p).$$

Using the notation introduced in Lemma 1, define the convex decreasing function Λ on $]\underline{p}, \infty[$ by

$$\Lambda(p) = \begin{cases} -\Phi(p) & \text{if } \underline{p} < p < \bar{p}, \\ -(p + 1)m & \text{if } p \geq \bar{p}. \end{cases}$$

Corollary 1. *It holds a.s. that*

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \int_{\mathbb{R}} e^{tpy} \rho_t(dy) = \Lambda(p)$$

for every $p > \underline{p}$.

More precisely, it can also be shown that for $p < \underline{p}$, the above limit equals $+\infty$ a.s.; but as this will not be needed in the sequel, we shall not give the argument. Taking Corollary 1 for granted here, we derive large deviation estimates for the random measures ρ_t . Introduce the Fenchel-Legendre transform of Λ ,

$$\Lambda^*(a) = \sup_{p > \underline{p}} (ap - \Lambda(p))$$

Note that $\Lambda^*(a) = \infty$ for every $a > -m$ and that Λ^* is left-continuous at $-m$.

Corollary 2. *Suppose that $\underline{p} < 0$. The following holds a.s.*

(i) *For any closed set $F \subseteq \mathbb{R}$,*

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log \rho_t(F) \leq -\inf \{ \Lambda^*(p), p \in F \}.$$

(ii) *For any open set $G \subseteq \mathbb{R}$,*

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \log \rho_t(G) \geq -\inf \{ \Lambda^*(p), p > -\Phi'(\underline{p}+) \text{ and } p \in G \}.$$

(iii) If moreover $\Phi'(\underline{p}+) = \infty$, then (ρ_t) satisfy the LDP with the good convex rate function Λ^* .

Proof. We aim at applying the Gärtner-Ellis theorem (see Sect. 2.3 in Dembo and Zeitouni [17]). The fundamental condition on the behavior of the Laplace transform of ρ_t (see Assumption 2.3.2 in [17]), is the conclusion of Corollary 1. Note that the assumption $\underline{p} < 0$ ensures that 0 belongs to the interior of the domain of Λ . According to Lemma 2.3.9 in [17], every $x \in]-\Phi'(\underline{p}+), -\Phi'(\overline{p})[$ is a so-called exposed point of the Fenchel-Legendre transform Λ^* , and since $\Lambda^*(p) = \infty$ for every $p > -m$, the statements (i) and (ii) merely rephrase Theorem 2.3.6 in [17]. The last statement follows from the first two and Lemma 2.3.9 in [17]. \square

2.4. Proofs

We first establish Theorem 2.

Proof. Note that the martingale $M(\cdot, p)$ is given by the product of a monotone process and a monotone function, and hence is pure-jump. It is known that if for some $q \in]1, 2[$, the q -variation

$$V_q(p) := \sum_{t \geq 0} |M(t, p) - M(t-, p)|^q$$

of $M(\cdot, p)$ is an integrable variable, then $M(\cdot, p)$ is bounded in $L^q(\mathbb{P})$ and *a fortiori* uniformly integrable; see e.g. Lépyngle [21]. In this direction, observe that the jumps of $M(\cdot, p)$ can be expressed in terms of the points of $(\Delta(t), k(t))$ of the Poisson point process used in Sect. 2.1 to describe the sudden dislocations of the fragmentation:

$$|M(t, p) - M(t-, p)| = \exp(t\Phi(p)) X_{k(t)}^{p+1}(t-) \left| 1 - \sum_{i=1}^{\infty} \Delta_i^{p+1}(t) \right|,$$

where $\Delta_1(t) \geq \Delta_2(t) \geq \dots$ is the decreasing sequence of the terms of $\Delta(t)$. Since the characteristic measure of the Poisson point process is $\nu \otimes \#$, the compensation formula yields

$$\mathbb{E}(V_q(p)) = c(p, q) \int_0^\infty \exp(tq\Phi(p)) \mathbb{E}(\Sigma(t, q(p+1) - 1)) dt,$$

where

$$c(p, q) := \int_{S^*} \left| 1 - \sum_{i=1}^{\infty} x_i^{p+1} \right|^q \nu(dx).$$

Let us check that $c(p, q) < \infty$ whenever $q > 1$ is picked sufficiently close to 1. This is trivial when $p \geq 0$, since then

$$\left| 1 - \sum_{i=1}^{\infty} x_i^{p+1} \right|^q \leq 1 - x_1^{p+1} \leq (p+1)(1-x_1),$$

and the condition $\int_{\mathcal{S}^*} (1 - x_1) \nu(dx) < \infty$ ensures the finiteness of $c(p, q)$. Suppose now that $\underline{p} < p < 0$. We use the hypothesis $\sum_{i=1}^{\infty} x_i = 1$ for ν -almost every $s \in \mathcal{S}^*$, and Jensen's inequality to get that $\nu(dx)$ -a.e.

$$\left| 1 - \sum_{i=1}^{\infty} x_i^{p+1} \right|^q = \left(\sum_{i=1}^{\infty} x_i (x_i^p - 1) \right)^q \leq \sum_{i=1}^{\infty} x_i (x_i^p - 1)^q .$$

The right-hand side is bounded from above by

$$\sum_{i=1}^{\infty} x_i (x_i^{pq} - 1) = \sum_{i=1}^{\infty} x_i^{pq+1} - 1 .$$

Since $\underline{p} < p < 0$, we still have that $\underline{p} < pq < 0$ provided that $q > 1$ is chosen sufficiently close to 1, and then we see by the very definition of \underline{p} that

$$c(p, q) \leq \int_{\mathcal{S}^*} \left(\sum_{i=1}^{\infty} x_i^{pq+1} - 1 \right) \nu(dx) < \infty .$$

On the other hand, we know from (6) that $\mathbb{E}(\Sigma(t, q(p+1) - 1)) = \exp(-t\Phi(q(p+1) - 1))$, so we conclude that $M(\cdot, p)$ is uniformly integrable whenever we may find $q > 1$ arbitrarily close to 1 such that

$$q\Phi(p) < \Phi(q(p+1) - 1) . \tag{8}$$

Recall now the hypothesis $p < \bar{p}$, so we may pick $q > 1$ sufficiently close to 1 such that $p < q(p+1) - 1 < \bar{p}$. By the part (i), we know that

$$\frac{\Phi(p)}{p+1} < \frac{\Phi(q(p+1) - 1)}{q(p+1)} ,$$

so (8) holds and $M(\cdot, p)$ is uniformly integrable.

Thus all that we need is to check that its terminal value $M(\infty, p)$ is strictly positive a.s., which is straightforward. Indeed, we see from the scaling property that the probability $\mathbb{P}_r(M(\infty, p) = 0)$ does not depend on $r > 0$. On the other hand, the fragmentation property entails

$$\mathbb{P}(M(\infty, p) = 0 \mid X(t)) = \prod_{X_i(t) > 0} \mathbb{P}_{X_i(t)}(M(\infty, p) = 0) .$$

Since $\nu \not\equiv 0$ and $\nu(\sum_{i=1}^{\infty} x_i < 1) = 0$, the number of fragments at time $t > 0$, $\text{Card}\{i : X_i(t) > 0\}$, is always at least 1, and is greater than 1 with positive probability. Hence $\mathbb{P}(M(\infty, p) = 0)$ must be equal to 0 or 1, and the uniform integrability of $M(\cdot, p)$ rules out the case when $M(\infty, p) = 0$ a.s. \square

Next, we establish Corollary 1

Proof. Let us first prove the statement for a fixed $p > \underline{p}$. The case $\underline{p} < p < \bar{p}$ follows from Theorem 2, so suppose that $\bar{p} \leq p$. We first observe that

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log X_1(t) = -m \quad \text{a.s.} \tag{9}$$

Indeed, the trivial upper-bound $X_1(t) \leq \Sigma(t, \bar{p})^{1/(\bar{p}+1)}$ and the convergence of the martingale $M(t, \bar{p})$ yield

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log X_1(t) \leq -\Phi(\bar{p})/(\bar{p} + 1) = -m \quad \text{a.s.}$$

On the other hand, pick any $q > 0$ and $\varepsilon > 0$ such that $q + \varepsilon < \bar{p}$. As

$$\Sigma(t, q + \varepsilon) \leq X_1^\varepsilon(t) \Sigma(t, q),$$

we deduce from Theorem 2 that

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \log X_1(t) \geq \frac{\Phi(q) - \Phi(q + \varepsilon)}{\varepsilon}.$$

We let first $\varepsilon \rightarrow 0+$, and then $q \rightarrow \bar{p}$, and we obtain the lower-bound

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \log X_1(t) \geq -\Phi'(\bar{p}) = -m.$$

Next, using the bounds

$$X_1^{p+1}(t) \leq \Sigma(t, p) \leq X_1^{(p-\bar{p})}(t) \Sigma(t, \bar{p}),$$

we deduce by (9) and the convergence of the martingale $M(t, \bar{p})$ that

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \Sigma(t, p) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \int_{\mathbb{R}} e^{npy} \rho_t(dy) = -(p + 1)m = \Lambda(p) \quad \text{a.s.}$$

The limit above holds a.s. simultaneously for every rational number $p > \underline{p}$, and by an immediate monotonicity argument, the proof is complete. \square

2.5. Miscellaneous comments

Let us comment on the results presented in this section and make the connection with related results in the literature.

• **On Theorem 1:** It is likely that, at least under some more restrictive hypotheses, it should be possible to reinforce the convergence in probability into almost-sure convergence. Indeed, this is the case for branching random walks; see Asmussen and Kaplan [3] and Biggins [12], and also the comments below on Corollary 2. We point out that the approach of Theorem 1(ii) is inspired by the work of Brennan and Durrett [15], and is close to the so-called propagation of chaos (see Sznitman [27]). More precisely, for every integer k , if U_1, \dots, U_k are independent uniformly

distributed variables, then the argument of the proof shows that the k -tuple of the size of the interval components of $G(t)$ that contain respectively U_1, \dots, U_k converges in law as $t \rightarrow \infty$ after a suitable renormalization towards k independent normal variables.

• **On Theorem 2:** It is interesting to observe that Markov’s inequality combined with the fact that $\mathbb{E}(M(t, p)) = 1$ for every $t \geq 0$ and $p \in]\underline{p}, \infty[$ entails the following concentration inequality: for every $\varepsilon, t > 0$, we have

$$\mathbb{P}(\Sigma(t, p) > \varepsilon) \leq \varepsilon^{-1} \exp(-t\Phi(p)).$$

In the same vein, as for $p > \bar{p}$, $M(\cdot, p)$ is a martingale with no positive jumps that converges to 0 at ∞ , the optional sampling theorem yields for every $a > 1$ the identity

$$\mathbb{P}(\Sigma(t, p) \geq a \exp(-t\Phi(p)) \text{ for some } t \geq 0) = 1/a.$$

• **On Corollary 1:** This result provides a consistent estimator of the function Φ evaluated on $]\underline{p}, \bar{p} - 1[$, which is constructed from a *single* sample of $X(t)$. Recall that Φ is analytic on $]\underline{p}, \infty[$, and thus is entirely determined by its values at such points. On the other hand, if the fragmentation is binary, then we can recover its splitting measure from Φ . In conclusion, if we know that a homogeneous fragmentation is binary and mass-conservative, then we can construct a consistent estimator of the law of the process from a single sample of the fragmentation observed at a large time.

• **On Corollary 2:** Of course, this LDP agrees with the law of large numbers of Theorem 1. More precisely, the assumption that $\underline{p} < 0$ implies that Φ has a finite derivative at 0 and that $\Lambda^*(-\Phi'(0+)) = 0$, as expected from Theorem 1(i). In this direction, we stress that the condition that $\underline{p} < 0$ is more demanding than the requirement of finiteness for $\Phi'(0+)$; and that the LDP yields a *strong* LLN (i.e. the convergence in Theorem 1(i) holds almost-surely and not merely in probability).

It may be also interesting to translate this LDP in terms of a multi-scale asymptotic analysis for numbers of fragments. More precisely, define for every $a < 0$

$$C(a) := \lim_{\varepsilon \rightarrow 0+} \lim_{t \rightarrow \infty} \frac{1}{t} \log (\text{Card}\{i \in \mathbb{N} : e^{(a-\varepsilon)t} \leq X_i(t) \leq e^{(a+\varepsilon)t}\})$$

whenever the limits above exist, so that informally, the number of fragments of size approximately e^{at} at time t is approximately $e^{tC(a)}$ when t is large. It is immediately seen that when the hypothesis of Corollary 1 is fulfilled, then we have for all $a < -\Phi'(\underline{p}+)$ that

$$C(a) = -(\Lambda^*(a) + a).$$

In this direction, note that elementary calculations yield for every $p > \underline{p}$

$$\Lambda^*(-\Phi'(p)) = \Phi(p) - p\Phi'(p).$$

On the other hand, recall from Lemma 1 that $\Phi'(\bar{p}) = \Phi(\bar{p})/(\bar{p} + 1)p = m$, and we conclude that then $C(-m) = 0$. Loosely speaking, this means that there

exists fragments of size $\approx \exp\{-tm\}$; however the number of such fragments is subexponential, i.e. smaller than $e^{\eta t}$ for every $\eta > 0$. On the other hand, for every $a > -m$ we have $C(a) = -\infty$, which implies that a.s. all the fragments have size smaller than e^{at} whenever t is sufficiently large, which is in agreement with (9). In the same vein, if moreover $\Phi'(p+) = \infty$, then $C(a) > 0$ for every sufficiently negative a . This means that there exists a.s. extremely small fragments (i.e. of size $\approx \exp\{at\}$, no matter how small a is).

Finally, we mention that in the setting of branching random walks, Biggins [13] also used the remarkable martingales analogous to those in Theorem 2 to establish precise large deviation estimates for the empirical measure of a branching random walk (see also [11]). Recently, Corollary 2 has been strengthened into precise large deviation estimates, see [9].

3. Self-similar fragmentations

We next turn our attention to general self-similar fragmentations, i.e. the index of self-similarity α is now an arbitrary real number. As connections between the homogeneous and the self-similar cases will have an important role in our analysis, and in order to avoid a possible confusion, it is convenient from now on to denote self-similar fragmentations with index α by $X^{(\alpha)}(t) = (X_1^{(\alpha)}(t), X_2^{(\alpha)}(t), \dots)$, and to keep the notation $X = X^{(0)}$ in the homogeneous case. Again, we shall implicitly assume that the fragmentation starts from a single fragment with unit mass. Let us now recall some results taken from [7].

3.1. Preliminaries

There is a simple transformation that changes a homogeneous fragmentation X into a self-similar one, $X^{(\alpha)}$.

Recall the interval-representation of a fragmentation sketched in Sect. 2.1. So $(G(t), t \geq 0)$ is a nested family of open sets in the unit interval, and $X(t)$ is the ranked sequence of the lengths of the intervals components of $G(t)$. For every $y \in]0, 1[$, let $I_y(t)$ denote the interval component of $G(t)$ that contains y if $y \in G(t)$, and $I_y(t) = \emptyset$ otherwise. We write $|I|$ for the length of an interval $I \subseteq]0, 1[$, and for every $y \in]0, 1[$ we consider the time-substitution

$$T_y^{(\alpha)}(t) := \inf \left\{ u \geq 0 : \int_0^u |I_y(v)|^{-\alpha} dv > t \right\}.$$

Because the open sets $G(t)$ are nested, we see that for every $y, z \in]0, 1[$, the intervals $I_y(T_y^{(\alpha)}(t))$ and $I_z(T_z^{(\alpha)}(t))$ are either identical or disjoint, so the family $\{I_y(T_y^{(\alpha)}(t)), 0 < y < 1\}$ can be viewed as the interval components of an open set $G^{(\alpha)}(t)$. It is straightforward that the family $(G^{(\alpha)}(t), t \geq 0)$ is nested. More precisely, if we write $X^{(\alpha)}(t)$ for the ordered sequence of the lengths of the interval components of $G^{(\alpha)}(t)$, then $(X^{(\alpha)}(t), t \geq 0)$ is a self-similar fragmentation with index α . See Theorem 2 in [7].

Any self-similar fragmentation $X^{(\alpha)}$ can be constructed from some homogeneous one X as above, and this construction can be inverted. In particular, the distribution of $X^{(\alpha)}$ is entirely determined by the index of self-similarity α , and the erosion coefficient $c \geq 0$ and the splitting measure ν of the homogeneous fragmentation X . In the sequel, the notation Φ will refer to the function defined in (2).

3.2. Case when the self-similarity index is positive

For the purpose of this section, it is convenient to describe the fragmentation $X^{(\alpha)}(t)$ via the following modified version of the empirical measure of its components,

$$\rho_t^{(\alpha)} := \sum_{i=1}^{\infty} X_i^{(\alpha)}(t) \delta_{t^{1/\alpha} X_i^{(\alpha)}(t)}.$$

Note that in comparison with (5), to each fragment, say y , corresponds a mass at $t^{1/\alpha} y$ instead of $t^{-1} \log y$. Our main result of this setting is the following theorem which can be viewed as an extension of Theorem 1 to the self-similar case.

Theorem 3. *Suppose that $\alpha > 0$, and assume that*

(i) *the erosion coefficient is $c = 0$ and $\nu(\sum_{i=1}^{\infty} x_i < 1) = 0$,*

(ii)
$$\Phi'(0+) = \mu = \sum_{i=1}^{\infty} \int_{S^*} x_i \log(1/x_i) \nu(dx) < \infty,$$

(iii) *the fragmentation is not geometric, that is there exists no $r > 0$ such that the size of every fragment at time $t > 0$ lies in the set $\{e^{-kr} : k = 0, 1, \dots\}$.*

Then the random measures $\rho_t^{(\alpha)}$ are probability measures a.s., and when $t \rightarrow \infty$, they converge in probability, in the sense of weak convergence of measures. The limit $\rho_{\infty}^{(\alpha)}$ is deterministic; it is determined by the moments

$$\int_{]0, \infty[} y^{\alpha k} \rho_{\infty}^{(\alpha)}(dy) = \frac{(k-1)!}{\alpha \mu \Phi(\alpha) \cdots \Phi((k-1)\alpha)} \quad \text{for } k = 1, \dots,$$

(with the usual convention that the right-hand side above equals $1/(\alpha\mu)$ for $k = 1$).

Proof. The proof is similar to that of Theorem 1. The key difference is that the size $\chi^{(\alpha)}(t)$ of the tagged fragment at time t (i.e. the length of the interval that contains a point which has been tagged at random according to the uniform distribution, and independently of the fragmentation process) has a more complicated expression. Corollary 2 of [7] states that the distribution of the process $\chi^{(\alpha)}(\cdot)$ can be described in terms of the function Φ as follows. Let $\xi = (\xi(t), t \geq 0)$ be a subordinator with Laplace exponent Φ , and define implicitly $\tau(t)$ for every $t \geq 0$ by the identity

$$\int_0^{\tau(t)} \exp(\alpha \xi_s) ds = t.$$

Then the processes $(\exp(-\xi_{\tau(t)}), t \geq 0)$ and $(\chi^{(\alpha)}(t), t \geq 0)$ have the same distribution. Roughly, this is an easy consequence of the transformation $X \rightarrow X^{(\alpha)}$ described above. In the terminology introduced by Lamperti [20], this means that $1/|\chi^{(\alpha)}(\cdot)|$ is an increasing self-similar Markov process with index $1/\alpha$, which is associated with the subordinator with Laplace exponent Φ .

The hypotheses (i–ii) are necessary and sufficient conditions for $\mathbb{E}(\xi_1) = \Phi'(0+) < \infty$, whereas (iii) means that the subordinator ξ is not arithmetic. In this case, it has been shown in [8] (extending Brennan and Durrett [15]) that $t^{1/\alpha} \exp(-\xi_{\tau(t)})$ has a limiting distribution as $t \rightarrow \infty$, which is given by the probability measure $\rho_\infty^{(\alpha)}$ which appears in the statement. By the scaling property, we thus have that for every $r > 0$, the distribution of $t^{1/\alpha} \chi^{(\alpha)}(t)$ under \mathbb{P}_r (i.e. when at the fragmentation starts from a single fragment of size r) converges weakly towards $\rho_\infty^{(\alpha)}$ as $t \rightarrow \infty$.

The rest of the proof follows exactly the same route as for Theorem 1(ii). One expresses the first and second moments of $\int f(y)\rho_t^\alpha(dy)$ in terms of $t^{1/\alpha} \chi^{(\alpha)}(t)$, and check the asymptotic independence of the size of two independently tagged fragments. □

Let us now comment on Theorem 3. A version of Theorem 3 in the case when the dislocation measure ν has a finite total mass was first proved by Filippov [18]. Then a special case of Filippov’s result was recovered by Brennan and Durrett [14, 15]. Roughly, these authors treated the case of discrete binary fragmentations (i.e. the Lévy measure ν has finite mass and is supported by the set of sequences $x = (x_1, \dots) \in \mathcal{S}^*$ with $x_1 + x_2 = 1$). Then it is plain that at any time $t > 0$, the number $N(t)$ of fragments with positive size is finite, so one can investigate directly the true empirical measure

$$\frac{1}{N(t)} \sum_{i=1}^{N(t)} \delta_{t^{1/\alpha} X_i(t)}.$$

It should also be stressed that Brennan and Durrett established almost-sure convergence and not merely convergence in probability as we did.

Motivated by the study of the so-called standard additive coalescent, Aldous and Pitman [2] introduced a self-similar fragmentation with index $\alpha = 1/2$ and for which

$$\Phi(q) = \sqrt{2} \Gamma(q + 1/2) / \Gamma(q),$$

see the identity (12) in [7]. They were able to describe explicitly the distribution of $X(t)$ at any time $t > 0$ in terms of the atoms of a certain conditioned Poisson measure on $]0, \infty[$, and then to investigate its asymptotic behavior for large times. More precisely, not only they obtained Theorem 3 in this case (then the limit distribution $\rho_\infty^{(\alpha)}$ has density $(2\pi y)^{-1/2} e^{-y/2}$, i.e. it is the law of the square of a normal $\mathcal{N}(0, 1)$ variable), but they also established Gaussian fluctuations, which does not seem easy to get in the general setting.

Just as in the homogeneous case (recall the comment about Corollary 1 in Sect. 2.4), Theorem 3 can be used to construct an consistent estimator of the

function Φ from a single sample of $X^{(\alpha)}(t)$. Indeed Theorem 3 yields an estimator of Φ evaluated at points $\alpha, 2\alpha, \dots$, and this estimator is consistent as we know that it converges to the exact values of Φ at such points when $t \rightarrow \infty$. On the other hand, because Φ is the Laplace exponent of some subordinator, it is easy to check that Φ is entirely determined by its values at such points.

Finally, as an easy application of Theorem 3, one can specify the asymptotic behavior of the scattering rates of $X^{(\alpha)}$. More precisely, if we set

$$\Sigma^{(\alpha)}(t, p) := \sum_{i=1}^{\infty} (X_i^{(\alpha)}(t))^p,$$

then we have the following estimate.

Corollary 3. *Under the assumptions of Theorem 3, we have for every $p \geq 1$ that*

$$\lim_{t \rightarrow \infty} t^{(p-1)/\alpha} \Sigma^{(\alpha)}(t, p) = \int_{]0, \infty[} y^{p-1} \rho_{\infty}^{(\alpha)}(dy) \in]0, \infty[,$$

where the convergence holds in probability.

Proof. We have the identity

$$t^{(p-1)/\alpha} \Sigma^{(\alpha)}(t, p) = \int_{]0, \infty[} y^{p-1} \rho_t^{(\alpha)}(dy);$$

however as the function $y \rightarrow y^{p-1}$ is unbounded, we cannot directly apply Theorem 3. Nonetheless, if we simply try to repeat the argument used there, we see immediately that all that is needed is to check that the family $(t^{1/\alpha} \chi^{(\alpha)}(t), t \geq 0)$ is bounded in $L^q(\mathbb{P})$ for every $q \geq 1$. In this direction, it is enough to exhibit a decreasing function $f :]0, \infty[\rightarrow]0, \infty[$ with $f(r) = o(r^{-q})$ at ∞ for every $q > 0$, such that

$$\mathbb{P}(t^{1/\alpha} \chi^{(\alpha)}(t) > r) \leq f(r), \quad \text{for all } t, r \geq 0.$$

For this, recall that $1/\chi^{(\alpha)}(t)$ is an increasing semi-stable Markov process corresponding to some subordinator ξ via Lamperti's transformation. If we set $\tau_b = \inf \{t \geq 0 : 1/\chi^{(\alpha)}(t) > b\}$, then Corollary 5 in [8] entails that for every $b > 1$, $b^{-\alpha} \tau_b$ is stochastically dominated by the variable $\int_0^{\infty} \exp(-\alpha \xi_s) ds$. We deduce that

$$\mathbb{P}(t^{1/\alpha} \chi^{(\alpha)}(t) > r) = \mathbb{P}(\tau_{t^{1/\alpha}/r} > t) \leq \mathbb{P}\left(\int_0^{\infty} \exp(-\alpha \xi_s) ds > r^{\alpha}\right).$$

As the variable $\int_0^{\infty} \exp(-\alpha \xi_s) ds$ has finite moments of arbitrary order (these are computed e.g. in [16]), this completes the proof. \square

We conclude this section with the following simple observation: when the conditions of Theorem 3 are fulfilled, the size of the largest fragment decays like $t^{-1/\alpha}$ as time goes to infinity, in the sense that

$$\log X_1^{(\alpha)}(t) \sim -\frac{1}{\alpha} \log t \quad \text{as } t \rightarrow \infty.$$

More precisely, the upper bound follows from Corollary 3, whereas the lower bound is obtained by considering a size-biased picked fragment.

3.3. Case when the self-similarity index is negative

Results in the case $\alpha < 0$ are simple applications of the transformation $X \rightarrow X^{(\alpha)}$ described in Sect. 3.1. An informal guideline is that small masses are then subject to intense fragmentation, and this makes them vanish entirely quickly. Here is a formal statement.

Proposition 2. (i) For $\alpha < 0$, it holds with probability one that

$$\zeta := \inf \{t \geq 0 : X^{(\alpha)}(t) = (0, \dots)\} < \infty.$$

(ii) For every $t > 0$ and $\alpha < -1$, it holds with probability one that

$$\text{Card}\{j \in \mathbb{N} : X_j^{(\alpha)}(t) > 0\} < \infty.$$

We stress that Proposition 2 holds in complete generality, in the sense that it is not required to suppose that the associated homogeneous fragmentation X has no erosion nor that its splitting law induces no loss of mass (but of course the case when X is constant is always excluded).

Of course, it would be interesting to determine the law of the lifetime ζ , but we have been able to tackle this question only in a couple of special cases. It is not even clear whether or not this distribution can be characterized just in terms of the exponent Φ . On the other hand, we stress that in general, no matter what the value of α is, there may exist random instants t at which

$$\text{Card}\{j \in \mathbb{N} : X_j^{(\alpha)}(t) > 0\} = \infty.$$

For instance in the case when the Lévy measure fulfills

$$\nu(x_j > 0 \text{ for all } j \in \mathbb{N}) = \infty,$$

then with probability one, there occur infinitely many sudden dislocations in the fragmentation process $X^{(\alpha)}$, each of which produces infinitely many masses. This does not induce any contradiction with Proposition 2 (ii) when $\alpha < -1$, because informally, as the index of self-similarity is negative, we know that small masses vanish quickly.

Proof. (i) Recall that with probability one

$$\limsup_{t \rightarrow \infty} t^{-1} \log X_1(t) < 0$$

(this holds for any non-trivial homogeneous fragmentation, as can be seen by the theorem of convergence of nonnegative martingales and the trivial bound $e^{t\Phi(p-1)} X_1(t) \leq M(t, p)$). In the notation used in Sect. 3.1, we may thus find $\varepsilon > 0$ and an a.s. finite random variable C such that

$$|I_y(r)| \leq C e^{-\varepsilon r} \quad \text{for all } 0 < y < 1 \text{ and } r \geq 0.$$

As $\alpha < 0$, we deduce that if we set $C' = C/(-\alpha\varepsilon) > 0$, then

$$\int_0^\infty |I_y(r)|^{-\alpha} dr \leq C',$$

and hence $T_y^{(\alpha)}(C') = \infty$ for all $0 < y < 1$. This shows that $X^{(\alpha)}(t) = (0, \dots)$ for every $t \geq C'$.

(ii) Recall the notation

$$\Sigma^{(\alpha)}(t, p) := \sum_{i=1}^{\infty} (X_i^{(\alpha)}(t))^p.$$

We first point out that for every $p > -\alpha$, we have the identity

$$\int_0^{\infty} \Sigma^{(\alpha)}(t, p + \alpha) dt = \int_0^{\infty} \Sigma(t, p) dt. \tag{10}$$

Indeed, in the notation of Sect. 3.1, we have for every $t \geq 0$

$$\Sigma^{(\alpha)}(t, p + \alpha) = \int_0^1 |I_y(T_y^{(\alpha)}(t))|^{p+\alpha-1} dy.$$

By the change of variables $T_y^{(\alpha)}(t) = s$, $dt = |I_y(s)|^{-\alpha} ds$, in the second identity below, we get

$$\begin{aligned} \int_0^{\infty} \Sigma^{(\alpha)}(t, p + \alpha) dt &= \int_0^1 dy \int_0^{\infty} dt |I_y(T_y^{(\alpha)}(t))|^{p+\alpha-1} \\ &= \int_0^1 dy \int_0^{\infty} ds |I_y(s)|^{p-1} \\ &= \int_0^{\infty} \Sigma(s, p) ds. \end{aligned}$$

As a consequence, if moreover $p > 1$, then by (6), the random variable in (10) has mean $1/\Phi(p - 1)$. Then let p decreases to $-\alpha$ (recall that the assumption $-\alpha > 1$), so $\Sigma^{(\alpha)}(t, p + \alpha)$ increases to

$$\Sigma^{(\alpha)}(t, 0) = \text{Card}\{j \in \mathbb{N} : X_j^{(\alpha)}(t) > 0\},$$

the number of fragments at time t . We get by monotone convergence

$$\int_0^{\infty} \mathbb{E}(\Sigma^{(\alpha)}(t, 0)) dt = \frac{1}{\Phi(-\alpha - 1)} < \infty,$$

which establishes our claim. □

Acknowledgements. I should like to thank Amir Dembo for stimulating discussions about self-similar fragmentations and suggesting an approach to tackle the study of fluctuations for the empirical measure. Thanks are also due to Gregory Miermont and Jason Schweinsberg for pointing at an error in an earlier draft.

References

1. Aldous, D.J.: Deterministic and stochastic models for coalescence (aggregation, coagulation): a review of the mean-field theory for probabilists. *Bernoulli* **5**, 3–48 (1999)
2. Aldous, D.J., Pitman, J.: The standard additive coalescent. *Ann. Probab.* **26**, 1703–1726 (1998)
3. Asmussen, S., Kaplan, N.: Branching random walks. I–II. *Stochastic Processes Appl.* **4**, 1–13 and 15–31 (1976)
4. Berestycki, J.: Ranked fragmentations. *ESAIM, Probab. Stat.* **6**, 157–176 (2002). Available via <http://www.edpsciences.org/ps/OnlinePSbis.html>
5. Bertoin, J.: A fragmentation process connected to Brownian motion. *Probab. Theory Relat. Fields* **117**, 289–301 (2000)
6. Bertoin, J.: Homogeneous fragmentation processes. *Probab. Theory Relat. Fields* **121**, 301–318 (2001)
7. Bertoin, J.: Self-similar fragmentations. *Ann. Inst. Henri Poincaré* **38**, 319–340 (2002)
8. Bertoin, J., Caballero, M.-E.: Entrance from $0+$ for increasing semi-stable Markov processes. *Bernoulli* **8**, 195–205 (2002)
9. Bertoin, J., Rouault, A.: Additive martingales and probability tilting for homogeneous fragmentations. Preprint (2003)
10. Biggins, J.D.: Martingale convergence in the branching random walk. *J. Appl. Probab.* **14**, 25–37 (1977)
11. Biggins, J.D.: Chernoff’s theorem in the branching random walk. *J. Appl. Probab.* **14**, 630–636 (1977)
12. Biggins, J.D.: The central limit theorem for the supercritical branching random walk, and related results. *Stochastic Processes Appl.* **34**, 255–274 (1990)
13. Biggins, J.D.: Uniform convergence of martingales in the branching random walk. *Ann. Probab.* **20**, 137–151 (1992)
14. Brennan, M.D., Durrett, R.: Splitting intervals. *Ann. Probab.* **14**, 1024–1036 (1986)
15. Brennan, M.D., Durrett, R.: Splitting intervals. II. Limit laws for lengths. *Probab. Theory Relat. Fields* **75**, 109–127 (1987)
16. Carmona, P., Petit, F., Yor, M.: On the distribution and asymptotic results for exponential functionals of Levy processes. In: Yor, M. (ed.), *Exponential functionals and principal values related to Brownian motion*. Biblioteca de la Revista Matemática Iberoamericana 1997
17. Dembo, A., Zeitouni, O.: *Large deviations techniques and applications*. Second edition. Berlin: Springer 1998
18. Filippov, A.F.: On the distribution of the sizes of particles which undergo splitting. *Theory Probab. Appl.* **6**, 275–293 (1961)
19. Kolmogoroff, A.N.: Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. *C. R. (Doklady) Acad. Sci. URSS* **31**, 99–101 (1941)
20. Lamperti, J.: Semi-stable Markov processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **22**, 205–225 (1972)
21. Lépingle, D.: La variation d’ordre p des semi-martingales. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **36**, 295–316 (1976)
22. Lyons, R., Pemantle, R., Peres, Y.: Conceptual proofs of $L \log L$ criteria for mean behaviour of branching processes. *Ann. Probab.* **23**, 1125–1138 (1995)
23. Miermont, G.: Self-similar fragmentations derived from the stable tree I: splitting at heights. Preprint (2003). To appear in *Probab. Theory Relat. Fields*
24. Miermont, G.: Self-similar fragmentations derived from the stable tree II: splitting at hubs. Preprint (2003)
25. Pitman, J.: Coalescent with multiple collisions. *Ann. Probab.* **27**, 1870–1902 (1999)
26. Schweinsberg, J.: Coalescent with simultaneous multiple collisions. *Electron J. Probab.* **5–12**, 1–50 (2000). <http://www.math.washington.edu/~ejpecp/ejp5contents.html>
27. Sznitman, A.-S.: *Topics in propagation of chaos*. Ecole d’été de Probabilités de St-Flour XIX, Lect. Notes Math. **1464**. Springer 1991