

Statistical aspects of random fragmentations

Thierry Huillet*

*Laboratoire de Physique Théorique et Modélisation, CNRS-UMR 8089 et Université de Cergy-Pontoise,
2 Avenue Adolphe Chauvin, 95032 Cergy-Pontoise, France*

Received 11 May 2004; received in revised form 29 October 2004

Abstract

Three random fragmentation of an interval processes are investigated. For each of them, there is a splitting probability and a probability not to split at each step of the fragmentation process whose overall effect is to stabilize the global number of splitting events. Some of their statistical features are studied in each case among which fragments' size distribution, partition function, structure of the underlying random fragmentation tree, occurrence of a phase transition. In the first homogeneous model, splitting probability does not depend on fragments' size at each step. In the next two fragmentation models, splitting probability is fragments' length dependent. In the first such models, fragments further split with probability one if their sizes exceed some cutoff value only; in a second model considered, splitting probability of finite-size objects is assumed to increase algebraically with fragments' size at each step. The impact of these dependencies on statistical properties of the resulting random partitions are studied. Several examples are supplied.

© 2004 Elsevier B.V. All rights reserved.

MSC: primary 60G57; 62E17; secondary 60K99; 62E15; 62E20

Keywords: Fragmentation models; Random partitions

1. Introduction

Fragmentation processes describe numerous phenomena arising in various fields of application, such as Astrophysics, Crystallography, Geology and Fracture, Nuclear Physics, Polymer and Computer Sciences

* Tel.: +33 1 34 25 75 14; fax: +33 1 34 25 75 00.

E-mail address: Thierry.Huillet@ptm.u-cergy.fr (T. Huillet).

(to cite only a few, see [1,7–9] and references therein, for applications to Physics at large; see also [2,3,10] and references therein for applications more related to Computer Science and historical background).

In certain collision processes, an unstable particle splits into b daughter fragments with some probability p . If splitting occurs, the parental mass is shared between them at random, different laws of partition fitting different splitting processes. With complementary probability $1 - p$, splitting does not take place and stable parental fragments are left unchanged for ever. This splitting process is then iterated independently on first generation sub-fragments and so on until exhaustion of the process (see [6] for a similar construction). In this note, three random fragmentation models within this class are investigated and several examples discussed.

The Krapivsky, Ben-Naim and Gross model: here, splitting probability is independent of fragments' size at each step. This homogeneous model can be understood from Galton–Watson branching processes. Using this background, some statistical features of the induced limiting fragmentation are discussed: fragments' size distribution, partition function of mass, Rényi's average fragments' size, sizes of smallest and largest fragments. This model exhibits a phase transition between sub-critical and super-critical regimes which is studied in some detail. Although, in this part the probability to generate offspring does not depend on fragments' length, we shall consider an intermediate case where it depends on the fraction of parental mass transmitted to each descendant fragment at each step of the fragmentation process. In some instances, this causes the emergence of a Dirac delta mass at zero of fragments' size distribution which otherwise presents generically an algebraic divergence of its small-size tail.

Next, we focus on two inhomogeneous fragmentation models for which splitting probability is, by nature, fragments' length dependent.

The first is the Dean–Majumdar model. In this model whose origin is to be found in Computer Science, a fragment with initial size x_0 splits with probability one if its size exceeds some cutoff value $x_c = 1$. In subsequent steps, fragmentation proceeds for each sub-fragments whose sizes are bigger than x_c only. This process naturally terminates with probability one. Fragments' size distribution, partition function of mass and statistical properties of the Dean–Majumdar tree are discussed together with the Dean–Majumdar phase transition on the variance of its internal nodes in the large x_0 limit.

A second new model is finally investigated which is loosely related to the Brennan–Durrett construction: here the size x_0 of the initial fragment to be split is necessarily bounded by 1. Next, the splitting probability is assumed to increase algebraically with fragments' size at each step, as x_0^a for some $a > 0$. This process is shown to terminate with probability one. It is also shown that this model presents some exactly solvable statistical features. In particular, we derive a closed form expression of the fragments' size distribution, together with some information on the underlying tree structure, with simplifications when $x_0 \rightarrow 0$. Simple examples are supplied.

2. The homogeneous fragmentation model

In this section, we introduce the simplest homogeneous random fragmentation process where the splitting probability is independent of fragments' sizes. This model can be understood from the discrete time Galton–Watson branching process. An intermediate case where the splitting probability is dependent on the fraction of parental mass received by each descendant fragment is also considered.

2.1. The Krapivsky, Ben-Naim and Grosse fragmentation model

We start with an interval of length 1. At step one, there is a probability $p \in (0, 1)$ to split the interval and so, with probability $\bar{p} := 1 - p$ the initial unit fragment remains unchanged for ever. If it splits, it splits into $b > 1$ fragments with random sizes, say (U_1, \dots, U_b) , where (U_1, \dots, U_b) has exchangeable distribution throughout, implying in particular that each U_k , $k = 1, \dots, b$ all share the same distribution, say the one of U_1 . On each first-generation sub-fragment, the splitting process is then iterated, independently, a property which we refer to in the sequel as the *renewal* structure of the process. We also assume that U_1 has a density $\pi(u) > 0$ on $(0, 1)$ with total mass 1. Similar processes have been investigated by Krapivsky, Ben-Naim and Grosse in the binary case when the branching number is $b = 2$; the term homogeneous refers to the fact that in such models, the splitting probability is independent of fragments' sizes at each step. Before proceeding with the detailed study of this particular model, we first give some remarkable examples of first-generation partition laws. These will be common to the three fragmentation models investigated in this manuscript.

2.1.1. First-generation fragment size distribution: remarkable partition laws

(1) With $\theta > 0$, assume that (U_1, \dots, U_b) is distributed according to the (exchangeable) Dirichlet- $D_b(\theta)$ density function on the simplex that is to say

$$\pi(u_1, \dots, u_b) = \frac{\Gamma(b\theta)}{\Gamma(\theta)^b} \prod_{k=1}^b u_k^{\theta-1} \cdot \delta_{(\sum_{k=1}^b u_k - 1)}. \quad (1)$$

Alternatively, (U_1, \dots, U_b) is characterized by its joint moment function

$$\phi(q_1, \dots, q_b) := \mathbf{E} \left[\prod_{k=1}^b U_k^{q_k} \right] = \frac{\Gamma(b\theta)}{\Gamma(b\theta + \sum_{k=1}^b q_k)} \prod_{k=1}^b \frac{\Gamma(\theta + q_k)}{\Gamma(\theta)}.$$

In this case, $U_k \stackrel{d}{=} U_1$, $k = 2, \dots, b$ and the individual fractions are all identically distributed. Their common density on the interval $(0, 1)$ is given by

$$\pi(u) = \frac{\Gamma(b\theta)}{\Gamma(\theta)\Gamma((b-1)\theta)} u^{\theta-1} (1-u)^{(b-1)\theta-1}.$$

This is the one of a beta(θ , $(b-1)\theta$) random variable, with moment function

$$\phi(q) := \mathbf{E}(U_1^q) = \frac{\Gamma(b\theta)}{\Gamma(b\theta + q)} \frac{\Gamma(\theta + q)}{\Gamma(\theta)}, \quad q > -\theta. \quad (2)$$

In particular, the mean value is $\mathbf{E}(U_1) = 1/b$ and the variance $\sigma^2(U_1) = \frac{b-1}{b^2(b\theta+1)}$.

The case $\theta = 1$ corresponds to the uniform partition into b fragments for which

$$\phi(q) = \frac{(b-1)!}{(q+b-1)(q+b-2) \cdots (q+1)}, \quad q > -1.$$

This remarkable family of models is in the larger class of those for which $U_k = S_k / (S_1 + \cdots + S_b)$ where the S_k , $k = 1, \dots, b$ are independent and identically distributed (i.i.d) positive random variables. Indeed, assuming $S_1 \stackrel{d}{\sim} \text{gamma}(\theta)$, the joint distribution of (U_1, \dots, U_b) is Dirichlet- $D_b(\theta)$.

(2) A related model in higher dimension is as follows: suppose we start with a unit-cube in dimension $d \geq 2$. Splitting of the cube consists in generating $b = 2^d$ sub-cubes with i.i.d random volume given by $U_1 = V_1 \dots V_d$ where V_1, \dots, V_d are i.i.d uniform. In this case, for $k = 1, \dots, 2^d$, $U_k \stackrel{d}{\sim} \exp\{-\text{gamma}(d)\}$ are log-gamma distributed, with density and distribution functions, respectively, given by

$$\pi(u) = \frac{1}{(d-1)!} (-\log u)^{d-1}, \quad (3)$$

$$\Pi(u) := \int_0^u \pi(v) dv = u \sum_{l=0}^{d-1} \frac{1}{l!} (-\log u)^l, \quad u \in (0, 1). \quad (4)$$

The U_1 distribution on the interval $(0, 1)$ is also characterized by its moment function $\mathbf{E}(U_1^q) = (1+q)^{-d}$, $q > -1$, with mean value $\mathbf{E}(U_1) = 2^{-d}$. \square

Let us now consider the study of the KBG model, starting with fragments' size distribution.

2.2. Fragments' size distribution in the KBG model

Let X_h be the random size of any fragment among those available in the splitting process up to step (or height) $h \in \mathbb{N} := \{0, 1, 2, \dots\}$. Using the renewal structure of the process, we obtain the recursive identity in distribution

$$X_{h+1} \stackrel{d}{=} 1 \cdot \bar{B}_1 + B_1 U_1 X_h^{(1)}, \quad h \geq 0, \quad X_0 = 1.$$

Here B_1 is a $\{0, 1\}$ -valued Bernoulli random variable, taking the value 1 with probability p , $\bar{B}_1 := 1 - B_1$, $X_h^{(1)} \stackrel{d}{=} X_h$ is a statistical copy of X_h and $(B_1, U_1, X_h^{(1)})$ are mutually independent random variables.

As we shall use this argument several times in the sequel, let us briefly comment this identity: if there is no splitting at step one (the event $\bar{B}_1 = 1$ occurring with probability \bar{p}) the random size X_{h+1} at step $h+1$ is just the one of the initial unit interval that does not undergo further fragmentation. If splitting occurs at step one (the event $B_1 = 1$ of probability p), then X_{h+1} coincides in law with $X_h^{(1)}$ at step h , after scaling properly by U_1 .

Letting $h \uparrow \infty$, the limiting fragments size $X := X_\infty$, if this random variable exists, should thus satisfy the distributional equality

$$X \stackrel{d}{=} \bar{B}_1 + B_1 U_1 X^{(1)}.$$

Let $\bar{F}(x) = \mathbf{P}(X > x)$ be the complementary probability distribution of X . This renewal structure can be used to obtain its governing equation. With $\mathbf{I}(\cdot)$ the set indicator function, we obtain

$$\bar{F}(x) = \bar{p} \cdot \mathbf{I}(x \leq 1) + p \int_x^1 \bar{F}\left(\frac{x}{u}\right) \pi(u) du. \quad (5)$$

This functional equation can best be solved by employing the moment function technique. Let thus $\Phi(q) := \mathbf{E}(X^q)$ and $\phi(q) := \mathbf{E}(U_1^q)$ be the moment functions of X and U_1 . As is well-known, moment

functions of $[0, 1]$ -valued random variables are monotone decreasing (actually completely monotone) functions of $q > 0$. From the above identity in law, we easily obtain

$$\Phi(q) = \frac{\bar{p}}{1 - p\phi(q)}, \quad q \geq 0, \quad (6)$$

which can be checked to be the moment function of a nondegenerate $[0, 1]$ -valued random variable X with total mass $\Phi(0) = 1$. For example, its mean value is $\mathbf{E}(X) = \Phi(1) = \frac{\bar{p}b}{b-p} \in (0, 1)$ whereas its variance reads $\sigma^2(X) = \frac{\bar{p}}{1-p\phi(2)} - [\frac{\bar{p}b}{b-p}]^2$. As $\phi(\infty) = \mathbf{P}(U_1 = 1) = 0$, we note that $\Phi(\infty) = \mathbf{P}(X = 1) = \bar{p}$.

Let us now interpret statistically the random variable X . With $U_1^{(0)} := 1$, we clearly have from (6)

$$X \stackrel{d}{=} \prod_{h=0}^H U_1^{(h)},$$

where $(U_1^{(1)}, \dots, U_1^{(h)}, \dots)$ are i.i.d. random variables with law the one of U_1 and “height” H is an integral-valued random variable, independent of $(U_1^{(1)}, \dots, U_1^{(h)}, \dots)$, with geometric distribution $\mathbf{P}(H = h) = \bar{p}p^h$, $h \geq 0$. So, X naturally interprets as a geometric random product of U_1 ’s. Furthermore, the moment function of X may also be written as

$$\Phi(q) = \bar{p} + p \frac{\bar{p}\phi(q)}{1 - p\phi(q)} \quad (7)$$

showing, as required, that $X \in [0, 1]$ has an atom at $x = 1$, with mass \bar{p} (if initially there is no splitting). On the other hand, with probability p , X has a continuous component, say X_c , with moment function $\Phi_c(q) := \mathbf{E}(X_c^q) = \frac{\bar{p}\phi(q)}{1 - p\phi(q)}$ and density f_c . This continuous component corresponds to the geometrical product $X_c \stackrel{d}{=} \prod_{h=1}^{H+1} U_1^{(h)}$ and we notice that now $\Phi_c(\infty) = \mathbf{P}(X_c = 1) = 0$.

Finally, the definition domain of $\Phi(q)$ is (q_c, ∞) where $q_c = \inf(q : \phi(q) \leq 1/p) \leq 0$. Two main cases arise, depending on the convergence radius of $\phi(q)$ itself:

- (1) Suppose $q_* := \sup(q : \phi(q) = \infty) \in [-\infty, 0)$, the case $q_* = -\infty$ arising when $\phi(q)$ is an entire function. Then $\Phi(q)$ is defined for $q > q_c$ where $q_* < q_c < 0$. The number q_c is uniquely determined by $\phi(q_c) = 1/p > 1$. Note that $\phi(q_c) < \infty$, $\phi'(q_c) \in (-\infty, 0)$ and that $\Phi(q_c) = \infty$, $\Phi(\infty) = \bar{p}$.
- (2) Suppose $\phi(q)$ is only defined for $q \geq 0$, with $\phi(0) = 1$. Then $q_c = 0$ and $\Phi(q)$ is defined for $q \geq 0$, with $\Phi(0) = 1$.

Example 1. If we assume that the moment function of U_1 , say $\phi(q) := \mathbf{E}(U_1^q)$ is defined for $q > q_*$ with $q_* \in [-\infty, 0)$, then $q_c < 0$ is determined by $\phi(q_c) = 1/p$ and is a dominant simple pole for $\Phi(q)$: in a neighborhood of zero, we have an algebraic divergence of the fragment size density, the exponent of which being $-(q_c + 1)$. As q_c approaches 0 from the left, fragments with small masses occur with higher (diverging) probability.

In the two first following examples, we consider the Dirichlet partition example.

- (1) Suppose U_1 is uniformly distributed (as in the Dirichlet- $D_2(1)$ model), then $\phi(q) = 1/(1 + q)$ and $q_* = -1 < q_c = -\bar{p}$.

Then $\Phi(q) = \frac{\bar{p}(1+q)}{\bar{p}+q} = \bar{p} + p \frac{\bar{p}}{\bar{p}+q}$ where $\Phi_c(q) = \frac{\bar{p}}{\bar{p}+q}$ is the moment function of a $\text{beta}(\bar{p}, 1)$ distributed random variable with density $f_c(x) = \bar{p}x^{-\bar{p}}$, $x \in (0, 1)$ and distribution function $F_c(x) = x^{\bar{p}}$.

For the related random splitting of d -dimensional cubes, $\phi(q) = (1+q)^{-d}$ and so $q_c = -1 + p^{1/d}$. As $d \uparrow \infty$, $q_c \sim \frac{1}{d} \log p \rightarrow 0^-$. This constitutes an example of fragmentation with emergence of small masses.

(2) Suppose $U_1 \stackrel{d}{\sim} \text{beta}(1, b-1)$, $\theta > 0$ (Dirichlet- $D_b(1)$ model). Then $\phi(q) = \frac{(b-1)!}{(q+b-1)(q+b-2)\cdots(q+1)}$, $q > q_* = -1$ and so

$$\Phi(q) = \frac{\bar{p}(q+b-1)(q+b-2)\cdots(q+1)}{(q+b-1)(q+b-2)\cdots(q+1) - p(b-1)!}.$$

Tuning b to larger values tends to shift q_c to the right towards 0. In particular, if $b=3$, we find $-\bar{p} < q_c = (-3 + \sqrt{1+8\bar{p}})/2 < 0$.

Although the exact shape of the distribution of X is quite complex in general, when $q_c < 0$ exists, singularity analysis of $\Phi(q)$ shows that we always have the algebraic divergence of the fragment size density

$$f_c(x) \sim_{x \downarrow 0} Ax^{-(1+q_c)}, \quad (8)$$

where $A = \lim_{q \rightarrow q_c} (q - q_c)\Phi_c(q) = \frac{\bar{p}}{-p^2\phi'(q_c)} > 0$.

(3) Homogeneous fragmentation with small masses. Although generically this algebraic divergence holds, the following example shows that in some extreme cases one should be cautious with this.

Suppose that the tails of $-\log U_1 > 0$ are given by $\mathbf{P}(-\log U_1 > x) \sim_{x \rightarrow \infty} \frac{1}{\Gamma(1-\alpha)} x^{-\alpha} L(x)$, $\alpha \in (0, 1)$, where $L(x)$ is a slowly varying function at ∞ . Then $\phi(q) \sim_{q \downarrow 0} q^\alpha L(1/q)$, $q \geq 0$ and there is no $q_c < 0$: $\phi(q_c) = 1/p$. However

$$\Phi(q) = \frac{\bar{p}}{1 - p\phi(q)}$$

remains defined for $q \geq 0$ and so X is well-defined. Moreover, in a vicinity of $x = 0$, using Karamata's Tauberian theorem (see [4, p. 445])

$$F(x) \sim_{x \downarrow 0} \frac{p}{\bar{p}\Gamma(1-\alpha)} (-\log x)^{-\alpha} L(-\log x)$$

showing that the probability mass in a neighborhood of zero is much greater in this case than in the algebraic previous one. If $b=2$, this can for example be achieved while assuming $U_1 = S_1/(S_1 + S_2)$ where S_1 and S_2 are i.i.d., positive, with $\mathbf{P}(S_1 > s) \sim_{s \rightarrow \infty} \frac{1}{\Gamma(1-\alpha)} (\log s)^{-\alpha} L(\log s)$. \square

Remark. Suppose we start with an initial fragment with given length $x_0 > 0$. Let X_{x_0} be the size of a randomly chosen fragment in a KBG splitting process of an interval of length x_0 . Scaling arguments easily show the scaling property $X_{x_0} \stackrel{d}{=} x_0 X$ where X is obtained from the fragmentation of a unit interval.

Let $\Phi_{x_0}(q) := \mathbf{E}(X_{x_0}^q)$ be the moment function of X_{x_0} . Then, $\Phi_{x_0}(q) = x_0^q \Phi_1(q)$ where $\Phi_1(q) = \Phi(q)$ is given by (6). Note that if $x_0 \rightarrow X_0$, assuming initial length $X_0 > 0$ to be random and if the splitting process is independent of X_0 , then $\mathbf{E}\Phi_{X_0}(q) = \mathbf{E}(X_0^q)\Phi(q)$; the definition domain $\{q : \mathbf{E}(X_0^q) < \infty\}$ should also be taken into account to compute the one of $\mathbf{E}\Phi_{X_0}(q)$.

This scaling behavior is characteristic of situations where the splitting probability p is independent of the initial fragment size x_0 . As we shall see below, this is far from the case if not.

2.3. Partition function of the KPG model

Let N_h be the number of available fragments at step h of the fragmentation process, either stable or unstable. Let $X_{i,h}$ be the mass attached to fragment number i , $i = 1, \dots, N_h$ (its length). Then, under our assumptions, for each h , the $X_{i,h}$, $i = 1, \dots, N_h$ constitute an exchangeable partition of the unit interval, with, in particular $X_{i,h} \stackrel{d}{=} X_h$ defined above. This partition has random size N_h . In the sequel, we shall study the random partition function $Z_h(\beta) := \sum_{i=1}^{N_h} X_{i,h}^\beta$. Some of its basic statistical properties are then emphasized leading to some exact calculations.

2.3.1. Partition function

With $Z_h(1) = 1$, $Z_h(0) = N_h$, consider then the random partition function at height h

$$Z_h(\beta) := \sum_{i=1}^{N_h} X_{i,h}^\beta. \quad (9)$$

The smallest (largest) term of the $X_{i,h}$, $i = 1, \dots, N_h$ is clearly smaller (larger) than $1/N_h$. As a result, the range of $Z_h(\beta)$ is $[N_h^{1-\beta}, 1]$ if $\beta \geq 1$ and $[1, N_h^{1-\beta}]$ if $\beta < 1$.

Let $Z_h^{(k)}(\beta)$, $k = 1, \dots, b$ be independent statistical copies of $Z_h(\beta)$. Using the renewal structure of the KBG fragmentation process, we have the recursive identity in law

$$Z_{h+1}(\beta) \stackrel{d}{=} \bar{B}_1 + B_1 \sum_{k=1}^b U_k^\beta Z_h^{(k)}(\beta), \quad h \geq 0, \quad Z_0(\beta) = 1.$$

If $Z(\beta) := Z_\infty(\beta)$ exists, it must solve

$$Z(\beta) \stackrel{d}{=} \bar{B}_1 + B_1 \sum_{k=1}^b U_k^\beta Z^{(k)}(\beta).$$

Let $z_1(\beta) := \mathbf{E}(Z(\beta))$ be the first moment of $Z(\beta)$.

Define next $\beta_c := \inf(\beta : \phi(\beta) \leq 1/(bp)) < 1$. Then

$$z_1(\beta) = \frac{\bar{p}}{1 - bp\phi(\beta)} \quad \text{for } \beta > \beta_c, \quad (10)$$

$$z_1(\beta) = \infty \quad \text{otherwise.} \quad (11)$$

When $\beta > \beta_c$, higher-order moments $z_m(\beta) := \mathbf{E}(Z(\beta)^m)$, $m \geq 2$ exist. They can be computed recursively, using the multinomial identity, giving

$$z_m(\beta) = \bar{p} + p \sum_{\substack{m_1, \dots, m_b \geq 0: \\ \sum_{k=1}^b m_k = m}} \binom{m}{m_1 \dots m_b} \mathbf{E} \left[\prod_{k=1}^b U_k^{\beta m_k} \right] \prod_{k=1}^b z_{m_k}(\beta).$$

For example, if $\beta > \beta_c$

$$z_2(\beta) = \left(\bar{p} + 2p \binom{b}{2} \mathbf{E}[U_1^\beta U_2^\beta] z_1(\beta)^2 \right) / (1 - bp\phi(2\beta)) \quad (12)$$

is the second moment expressed in terms of the first and the full hierarchy $(z_m(\beta))_{m \geq 1}$ could be obtained in principle. Note that closed form expressions of $\mathbf{E}[\prod_{k=1}^b U_k^{\beta m_k}]$ are needed to do so.

In fact, using similar ideas, the moment function of $Z(\beta)$, $\beta > \beta_c$, which is $\mathbf{E}[Z(\beta)^\lambda]$, $\lambda \geq 0$, could be computed when λ is not an integer. Indeed, writing $Z(\beta) = 1 + (Z(\beta) - 1)$, with $(\lambda)_m := \lambda(\lambda - 1) \cdots (\lambda - m + 1)$, it can be formally obtained as

$$\mathbf{E}[Z(\beta)^\lambda] = 1 + \sum_{m \geq 1} \frac{(\lambda)_m}{m!} \tilde{z}_m(\beta) \quad (13)$$

in terms of $(\tilde{z}_m(\beta) := \mathbf{E}[(Z(\beta) - 1)^m], m \geq 1)$. Now, it may be checked that

$$\tilde{z}_m(\beta) = p \sum_{\substack{m_1, \dots, m_{b+1} \geq 0: \\ \sum_{k=1}^{b+1} m_k = m}} \binom{m}{m_1 \cdots m_{b+1}} \mathbf{E} \left[\prod_{k=1}^b U_k^{\beta m_k} \left(\sum_{k=1}^b U_k^\beta - 1 \right)^{m_{b+1}} \right] \prod_{k=1}^b \tilde{z}_{m_k}(\beta).$$

From these relations, $\tilde{z}_m(\beta)$, $m \geq 2$ can be in principle computed recursively starting from $\tilde{z}_1(\beta) = z_1(\beta) - 1$; an expression of $\mathbf{E}[\prod_{k=1}^b U_k^{\beta m_k} (\sum_{k=1}^b U_k^\beta - 1)^{m_{b+1}}]$ is needed to fully achieve this task.

2.3.2. Rényi's average fragments' size

With $\sum_{i=1}^N X_i = 1$ a random partition of unity, define the Rényi β -average of fragments' sizes to be

$$\langle X \rangle_\beta := \left[\sum_{i=1}^N X_i^{1+\beta} \right]^{1/\beta}, \quad \beta \in \mathbb{R}. \quad (14)$$

The 2-average $\langle X \rangle_2$ is often considered, but $\langle X \rangle_0 := \lim_{|\beta| \downarrow 0} \langle X \rangle_\beta = \prod_{i=1}^N X_i^{X_i}$ and $\langle X \rangle_{-1} = 1/N$ are also of interest. The function $\beta \rightarrow \langle X \rangle_\beta$ is nondecreasing with β ; in particular, with $X_{(1)} := X_1 \vee \cdots \vee X_N$ the size of the largest fragment and $X_{(N)} := X_1 \wedge \cdots \wedge X_N$ the one of the smallest, $\langle X \rangle_\beta \rightarrow \beta \uparrow \infty X_{(1)}$ and $\langle X \rangle_\beta \rightarrow \beta \downarrow -\infty X_{(N)}$, almost surely.

When $\beta > \beta_c - 1$, we have in our case $\langle X \rangle_\beta = Z(\beta + 1)^{1/\beta}$; the range of the random variable $\langle X \rangle_\beta$ is now $[\frac{1}{N}, 1]$ when $\beta \in (-1, \infty)$ and $[0, \frac{1}{N}]$, when $\beta \in (-\infty, -1)$.

The moment function of $\langle X \rangle_\beta$ is $\mathbf{E}[\langle X \rangle_\beta^\lambda] = \mathbf{E}[Z(\beta + 1)^{\lambda/\beta}]$. Assuming $\beta > \beta_c - 1$, it can be read from the one of $Z(\beta)$ described in (13), while substituting $\beta \rightarrow \beta + 1$, $\lambda \rightarrow \lambda/\beta$. We obtain

$$\mathbf{E}[\langle X \rangle_\beta^\lambda] = 1 + \sum_{m \geq 1} \frac{(\lambda/\beta)_m}{m!} \tilde{z}_m(\beta + 1). \quad (15)$$

Note that $\mathbf{E}[\langle X \rangle_\beta^\lambda]$ is also $\mathbf{E}[e^{-\lambda H_\beta}]$ where $H_\beta = -\log \langle X \rangle_\beta$ is the random Rényi β -entropy of the partition (X_1, \dots, X_N) , with, in particular, $H_0 = -\log \langle X \rangle_0 = -\sum_{i=1}^N X_i \log X_i$ Shannon's entropy.

2.3.3. Number of internal and external nodes of KBG fragmentation trees

In the study of the mass partition function, some understanding of the limiting number $N := N_\infty$ of fragments (the external nodes or leaves of the branching tree) is clearly useful. First, the average number of leaves, which is $z_1(0) = \mathbf{E}(N)$, reads

$$\mathbf{E}(N) = \frac{\bar{p}}{1 - bp} \quad \text{if } p < p_c := 1/b, \quad (16)$$

$$\mathbf{E}(N) = \infty \quad \text{if } p \geq p_c. \quad (17)$$

This suggests a critical phenomenon when p (or b) is tuned over the critical value $p_c = 1/b$ (or $b_c = 1/p$). Indeed, the following phase transition occurs

- *Sub-critical*: when $p \leq p_c$, $N < \infty$ almost surely and the full distribution of N is characterized by its generating function $\xi(u) := \mathbf{E}(u^N)$. It satisfies the functional equation: $\xi(u) = \bar{p}u + p\xi(u)^b$. In this case, $\beta_c \leq 0$.

- *Critical*: when $p = p_c$, $N < \infty$ almost surely but has no finite order 1 moment, reflecting the critical nature of the corresponding branching process; here $\beta_c = 0$.

- *Super-critical*: When $p > p_c$, $N = \infty$ with probability $1 - \rho$ where $\rho \in (0, 1)$ is the smallest real root in $(0, 1)$ of $g(u) := \bar{p} + pu^b = u$. With extinction probability ρ , N is finite and given $N < \infty$, the generating function of N solves $\xi(u) = \frac{\bar{p}}{\rho}u + p\rho^{b-1}\xi(u)^b$. In this case, $\beta_c \in (0, 1)$.

To see this and to solve the functional equation for $\xi(u)$, we need to enter into more details. Assume \mathcal{N}_h is the cumulative number of the internal and external nodes of the GW tree at step h of the fragmentation process. Using the renewal structure of the process, we get

$$\mathcal{N}_{h+1} \stackrel{d}{=} \bar{B}_1 + B_1 \left[1 + \sum_{k=1}^b \mathcal{N}_h^{(k)} \right], \quad h \geq 0, \quad \mathcal{N}_0 = 1,$$

where $(\mathcal{N}_h^{(k)}, k = 1, \dots, b)$ are b copies of \mathcal{N}_h . If $\mathcal{N} := \mathcal{N}_\infty$ exists, it satisfies

$$\mathcal{N} \stackrel{d}{=} \bar{B}_1 + B_1 \left[1 + \sum_{k=1}^b \mathcal{N}^{(k)} \right]. \quad (18)$$

If \bar{N} is the number of internal nodes of the GW tree, then $\mathcal{N} = \bar{N} + N$, where N is the number of its leaves. For b -ary trees, the range of \mathcal{N} is $1 + b\mathbb{N}$, with $\mathbb{N} := \{0, 1, 2, \dots\}$. In addition, it holds that $N = (b-1)\bar{N} + 1$ and so $N = (1 - 1/b)\mathcal{N} + 1/b$ and $\bar{N} = (\mathcal{N} - 1)/b$. Thus, statistical informations on N and \bar{N} can be obtained from the known ones on \mathcal{N} and we shall proceed in this way.

From standard Galton–Watson (GW) branching processes theory (see [5]), the following phase transition takes place for \mathcal{N} :

- *Sub-critical*: when $p \leq p_c$, $\mathcal{N} < \infty$ almost surely and the full distribution of \mathcal{N} is characterized by its generating function $\zeta(u) := \mathbf{E}(u^{\mathcal{N}})$ satisfying the functional equation: $\zeta(u) = u[\bar{p} + p\zeta(u)^b]$. From Lagrange inversion theorem, with $[z^n]g(z)$ the z^n -coefficient of the Taylor series development at $z = 0$

of the analytic function g , it holds

$$\zeta(u) = \sum_{n \geq 1} \frac{u^n}{n} [z^{n-1}]g(z)^n = \sum_{m \geq 0} \frac{u^{mb+1}}{mb+1} \binom{mb+1}{m} \bar{p}^{m(b-1)+1} p^m, \quad (19)$$

giving $\mathbf{P}(\mathcal{N} = n) = [u^n]\zeta(u)$.

Recalling the relation $N = (1 - 1/b)\mathcal{N} + 1/b$ for b -ary trees, we obtain in this way

$$\xi(u) = u^{1/b} \zeta(u^{1-1/b}) = \sum_{m \geq 0} \frac{u^{m(b-1)+1}}{mb+1} \binom{mb+1}{m} \bar{p}^{m(b-1)+1} p^m. \quad (20)$$

The random variable \mathcal{N} is the total number of nodes in the corresponding sub-critical Galton–Watson tree whereas N stands for the number of its leaves (or external nodes). The range of N is thus $1 + (b-1)\mathbb{N}$.

Finally, recalling the relation $\bar{N} = (\mathcal{N} - 1)/b$ for b -ary trees, the generating function $\bar{\xi}(u) := \mathbf{E}[u^{\bar{N}}]$ of internal nodes reads

$$\bar{\xi}(u) = u^{-1/b} \zeta(u^{1/b}) = \sum_{m \geq 0} \frac{u^m}{mb+1} \binom{mb+1}{m} \bar{p}^{m(b-1)+1} p^m. \quad (21)$$

The range of \bar{N} is the whole set of natural numbers \mathbb{N} .

• *Critical*: when $p = p_c$, \mathcal{N} is nondegenerate but has no finite order 1 moment, reflecting the critical nature of the corresponding branching process.

• *Super-critical*: When $p > p_c$, $\mathcal{N} = \infty$ with probability $1 - \rho$, where $\rho \in (0, 1)$ is the smallest root of $g(u) = u$. With extinction probability ρ , \mathcal{N} is finite. Given $\mathcal{N} < \infty$, \mathcal{N} is sub-critical and its generating function solves $\zeta(u) = u\tilde{g}(\zeta(u))$ with $\tilde{g}(u) = g(\rho u)/\rho = \bar{p}/\rho + p\rho^{b-1}u^b$ and $\tilde{g}'(1) = b(1 - \bar{p}/\rho) < 1$.

In the (sub)-critical case, it is well-known from singularity analysis that for large n of the form $n = mb + 1$, $[u^n]\zeta(u) \sim Ca^{-n}n^{-3/2}$ where $a = 1/g'(\tau)$ with τ the unique positive root of $g(\tau) - \tau g'(\tau) = 0$. Namely,

$$a = \frac{1}{pb} \left(\frac{p(b-1)}{\bar{p}} \right)^{\frac{b-1}{b}} \geq 1.$$

Note that $a > 1$ in the sub-critical case and that $a = 1$ in the critical case. So $[u^n]\zeta(u) \sim Cn^{-3/2}$ has algebraic decay in the critical regime, whereas $[u^n]\zeta(u)$ tends to 0 exponentially fast with n in the sub-critical phase.

2.3.4. Partition function in the super-critical case

Assume $p > p_c$ as in the super-critical case, hence with $\beta_c \in (0, 1)$. Clearly, when $\beta < \beta_c$ $Z_h(\beta) \rightarrow \infty$ with some positive probability. Defining $\rho(\beta) := pb\phi(\beta) > 1$, let us consider the scaled quantity $\tilde{Z}_h(\beta) := \rho(\beta)^{-h} Z_h(\beta)$. Then, $\tilde{Z}_h(\beta)$ satisfies the modified recursive identity in law

$$\rho(\beta)\tilde{Z}_{h+1}(\beta) \stackrel{d}{=} \rho(\beta)^{-h}\bar{B}_1 + B_1 \sum_{k=1}^b U_k^\beta \tilde{Z}_h^{(k)}(\beta), \quad h \geq 0, \quad \tilde{Z}_0(\beta) = 1.$$

As $h \uparrow \infty$, $\tilde{Z}(\beta) := \tilde{Z}_\infty(\beta)$, if it exists, satisfies the identity in law

$$\tilde{Z}(\beta) \stackrel{d}{=} 0 \cdot \bar{B}_1 + B_1 \sum_{k=1}^b \tilde{U}_k(\beta) \tilde{Z}^{(k)}(\beta).$$

Here $\tilde{U}_k(\beta) := \rho(\beta)^{-1} U_k^\beta$, $k = 1, \dots, b$.

Define the structure function $\tau(\beta) = -\log_b \rho(\beta)$. It is concave on the interval $(0 > q_*, \beta_c)$ with $\tau(0) \in (-1, 0)$. Using martingale arguments, one can prove that $\tilde{Z}_h(\beta) \rightarrow_{h \uparrow \infty} \tilde{Z}(\beta)$ when $\beta \in (\beta_-, \beta_c)$. Here, $q_* < \beta_- < 0$ is defined by

$$\beta_- \tau'(\beta_-) = \tau(\beta_-).$$

Moreover, when $\beta \in (\beta_-, \beta_c)$, $\mathbf{E}(\tilde{Z}(\beta)) = 1$ and $\tilde{Z}(\beta) > 0$ (respectively, $\tilde{Z}(\beta) = 0$) with probability $1 - \rho$ (respectively, ρ). As a result, with nonextinction probability $1 - \rho$, we obtain

$$-\frac{1}{h} \log_b Z_h(\beta) \rightarrow_{h \uparrow \infty} \tau(\beta). \quad (22)$$

Using large deviation arguments, one can argue that, with $\alpha \in (\tau'(\beta_c), \tau'(\beta_-))$

$$\lim_{\varepsilon \downarrow 0} \lim_{h \uparrow \infty} \frac{1}{h} \log_b \# \left\{ i : -\frac{1}{h} \log_b X_{i,h} \in (\alpha - \varepsilon, \alpha + \varepsilon) \right\} = f(\alpha), \quad (23)$$

where $f(\alpha) = \inf_{\beta \in (\beta_-, \beta_c)} (\alpha \beta - \tau(\beta)) \in (0, -\tau(0))$ is the concave Legendre transform of $\tau(\beta)$. This is of course reminiscent of Mandelbrot's multifractal theory of multiplicative cascades.

2.3.5. Shortest and longest leaves' heights in GW trees

Let H_* be the height of a dangling leaf at shortest distance from root in a b -ary Galton–Watson tree. Using the renewal structure of the process, we have the identity in law

$$H_* \stackrel{d}{=} 0 \cdot \bar{B}_1 + B_1 \left[1 + \bigwedge_{k=1}^b H_*^{(k)} \right].$$

Here, $\bigwedge_{k=1}^b H_*^{(k)}$ is the smallest height to the root of the b independent first-generation sub-trees. If $\bar{F}_*(h) = \mathbf{P}(H_* > h)$, $h \geq 0$, we obtain from this

$$\bar{F}_*(h+1) = p \bar{F}_*(h)^b, \quad h \geq 1, \quad \bar{F}_*(0) = p. \quad (24)$$

So, $\bar{F}_*(h) = \prod_{l=0}^h p^{b^l} = p^{\sum_{l=0}^h b^l}$, $h \geq 0$ gives the distribution of H_* . Note that $H_* < \infty$ in any case.

Let H^* be the height of a leaf at longest distance from root in a b -ary Galton–Watson tree. Similarly, we have

$$H^* \stackrel{d}{=} 0 \cdot \bar{B}_1 + B_1 \left[1 + \bigvee_{k=1}^b H^{*(k)} \right].$$

If $F^*(h) = \mathbf{P}(H^* \leq h)$, $h \geq 0$,

$$F^*(h+1) = \bar{p} + pF^*(h)^b = g(F^*(h)), \quad h \geq 1, \quad F^*(0) = \bar{p} \quad (25)$$

gives the distribution function $F^*(h)$ by recurrence. Again, three cases arise

- *Sub-critical*: when $p \leq p_c$, $H^* < \infty$ almost surely.
- *Critical*: when $p = p_c$, $H^* < \infty$ almost surely but has no finite order 1 moment, reflecting the critical nature of the corresponding branching process. Indeed, one may check from the above recurrence that in the critical case, the sequence $\bar{F}^*(h) := \mathbf{P}(H^* > h)$, $h \geq 0$ is such that $\bar{F}^*(h+1)/\bar{F}^*(h) \rightarrow 1$ ($h \uparrow \infty$), with $\bar{F}^*(h) \sim Ch^{-1/2}$.

- *Super-critical*: When $p > p_c$, $H^* = \infty$ with probability $1 - \rho$, $\rho \in (0, 1)$ defined above. With extinction probability ρ , H^* is finite.

With $U_0^{(h)} = 1$, $(U_1^{(1)}, \dots, U_1^{(h)}, \dots)$ an i.i.d. sequence with law the one of U_1 , this suggests to introduce the random variables

$$X_* \stackrel{d}{=} \prod_{h=0}^{H_*} U_1^{(h)} \quad \text{and} \quad X^* \stackrel{d}{=} \prod_{h=0}^{H^*} U_1^{(h)},$$

where H_* and H^* are independent of $(U_1^{(1)}, \dots, U_1^{(h)}, \dots)$. The random variables (X_*, X^*) are the random masses carried by leaves at heights H_* and H^* , respectively.

We are led to the expressions $\mathbf{E}[X_*^q] = \sum_{h \geq 0} \phi(q)^h \mathbf{P}(H_* = h)$ and $\mathbf{E}[X^{*q}] = \sum_{h \geq 0} \phi(q)^h \mathbf{P}(H^* = h)$.

For example, if $q = 1$, $\mathbf{E}[X_*] = \bar{p} + \sum_{h \geq 1} b^{-h} p^{\sum_{l=0}^{h-1} b^l} (1 - p^{b^h})$ is the average mass carried by the leaf at shortest distance from the root.

2.3.6. Largest and smallest piece in the limiting fragmentation process

This suggests to consider the following problem. Assume $p < p_c$. Let $X_{(1)} > \dots > X_{(N)}$ be obtained while ordering the fragments' lengths (X_1, \dots, X_N) . For example, the largest and smallest fragments' length are

$$X_{(1)} = \bigvee_{i=1}^N X_i \quad \text{and} \quad X_{(N)} = \bigwedge_{i=1}^N X_i,$$

where $X_i = \prod_{k=1}^{H_i} U_{1,i}^{(k)}$, $i = 1, \dots, N$ are N statistical copies of X with sum one $\sum_{i=1}^N X_i = 1$.

They are characterized by the identities in law

$$X_{(1)} \stackrel{d}{=} \bar{B}_1 + B_1 \bigvee_{k=1}^b U_k X_{(1)}^{(k)} \quad \text{and} \quad X_{(N)} \stackrel{d}{=} \bar{B}_1 + B_1 \bigwedge_{k=1}^b U_k X_{(N)}^{(k)}.$$

With $F_{(1)}(x) := \mathbf{P}(X_{(1)} \leq x)$ and $\bar{F}_{(N)}(x) := \mathbf{P}(X_{(N)} > x)$, this leads to the functional equations

$$F_{(1)}(x) = p \int \cdots \int_{[0,1]^b} \prod_{k=1}^b F_{(1)}\left(\frac{x}{u_k}\right) \pi(u_1, \dots, u_b) du_1 \dots du_b,$$

$$\bar{F}_{(N)}(x) = \bar{p} + p \int \cdots \int_{[x,1]^b} \prod_{k=1}^b \bar{F}_{(N)}\left(\frac{x}{u_k}\right) \pi(u_1, \dots, u_b) du_1 \dots du_b,$$

closed form solutions of which are currently out of reach, to the author's knowledge. There is, however, some reasonable hope that $X_{(1)}$ (respectively, $X_{(N)}$) is “close” to X_* (respectively, X^*). There are many open problems here, one of which could be: what are the height and index of the leaf carrying the largest (smallest) fragmentation mass?

2.4. A first extension: splitting probability depends on parental mass fraction

As we shall see in the sequel, there is some interest in considering situations where the splitting probability depends on lengths of fragments to be split at each step. Before we discuss two of these models, let us consider an intermediate step.

In this section, the probability p to generate offspring will not depend on fragments' length. However, we shall assume that it depends on the fraction of parental mass U_1 which is received by each descendant fragment at each step of the fragmentation process. Then

$$X_{h+1} \stackrel{d}{=} \bar{B}_1(U_1) + B_1(U_1)U_1X_h^{(1)}, \quad h \geq 0, \quad X_0 = 1$$

is the new recurrence to consider for fragment size distribution. Here, $B_1(U_1)$ is a Bernoulli random variable and $\mathbf{P}(B_1(U_1) = 1 \mid U_1 = u) = p(u) \in (0, 1)$ is the conditional probability that splitting occurs at first step given $U_1 = u$. We define $p := \int_0^1 p(u) du < 1$, $\bar{p} := 1 - p$ and the bounded probability density $q(u) := p(u)/p \in (0, 1/p)$. As $h \rightarrow \infty$, proceeding as above, we have to look for solutions to the distributional identity

$$X \stackrel{d}{=} \bar{B}_1(U_1) + B_1(U_1)U_1X^{(1)}.$$

Stated differently, using our notations, the functional equation on probability distributions to solve now reads

$$\bar{F}(x) = \bar{p} \cdot \mathbf{I}(x \leq 1) + p \int_x^1 \bar{F}\left(\frac{x}{u}\right) q(u) \pi(u) du. \quad (26)$$

In terms of moments, with $\Phi(q) = \mathbf{E}(X^q)$, this leads formally to

$$\Phi(q) = \frac{\bar{p}}{1 - p\psi(q)}. \quad (27)$$

Here, $\psi(q) = \int_0^1 u^q q(u) \pi(u) du$ is the moment transform of the function $u \rightarrow q(u) \pi(u)$ which, as a product of two probability densities is not in general a probability density. Indeed, its mass $\psi(0) = \int_0^1 q(u) \pi(u) du$

is different from 1 in general. However, as $0 < q(u) < 1/p$ is bounded, we have $0 < \psi(0) < 1/p$ and so $\Phi(q)$, $q \geq 0$, is well-defined there but with now $\Phi(0) = \bar{p}/1 - p\psi(0) > \bar{p}$. Two cases arise:

- If $\Phi(0) = \mathbf{P}(X > 0) < 1$ (or if $\psi(0) < 1$), then with probability $1 - \Phi(0) < p$, X will exhibit an atom at $x = 0$. The probability mass of X is defective as a result of mass being lost to a phase of zero-size particles.
- If $\Phi(0) \geq 1$ (or if $\psi(0) \geq 1$), then $\Phi(q)$ is not the moment function of a probability measure, but rather of some positive measure with excessive mass.

Note that $\Phi(0) = 1$ if and only if $\psi(0) = 1$; in this exceptional case, X , as a random variable on $[0, 1]$, has no atom at $x = 0$. One of these exceptional cases is when $b = 2$ and $\pi(u)$ is the uniform probability density.

Example 2. Consider a general model with branching number b and density $\pi(u)$ for U_1 .

Assume $p(u) = \Pi(u) := \int_0^u \pi(v) dv$. Then, $p = \int_0^1 \Pi(u) du = (b-1)/b$ and $q(u) = \frac{b}{b-1} \Pi(u)$. As a consequence, $\psi(0) = b/(b-1) \int_0^1 \Pi(u) \pi(u) du = b/[2(b-1)]$ and $\Phi(0) = 2/b \leq 1$ for any branching number $b \geq 2$.

Assume $p(u) = \bar{\Pi}(u) := 1 - \Pi(u)$. Then, $p = 1/b$. So $\psi(0) = b \int_0^1 \bar{\Pi}(u) \pi(u) du = b/2$ and $\Phi(0) = [2(b-1)]/b \geq 1$. In both cases, when $b = 2$, $\Phi(0) = 1$.

In the following two sections, we proceed with the study of strictly length dependent fragmentations. We shall consider two solvable fragmentation models for which the splitting probability depends on fragments' length. For the first one (the Dean–Majumdar model), we derive the fragment's length law. The Dean–Majumdar phase transition is also revisited. For the second one, which is reminiscent of the Brennan–Durrett model, we show that some computations are quite explicit.

3. The Dean–Majumdar model

Start with an interval with length $x_0 > 0$. Let $x_c := 1$ be some cutoff value. Suppose that the probability p to generate offspring at the first step depends on the lengths $(x_0, x_c := 1)$ according as $p = \mathbf{I}(x_0 > x_c)$. Thus fragmentation of the initial interval will only occur if its length x_0 is larger than 1. Otherwise, if $x_0 \leq 1$, the fragmentation stops at the first step.

If $x_0 > 1$, it splits into b daughter fragments with exchangeable random sizes, as usual. Iterate the fragmentation process and, at each step, assume that splitting will occur only if the lengths of sub-fragments to be eventually split are larger than $x_c = 1$. This model necessarily leads to a finite number of splits. It was recently reconsidered by Dean and Majumdar. We first derive the law of the length of a randomly chosen fragment according to this model. Then, the DM phase transition is revisited together with a preliminary study of the partition function.

3.1. Fragment length in the DM model

Let $X_h(x_0)$ be the random size of a fragment in the splitting process up to step (or height) $h \in \mathbb{N} := \{0, 1, 2, \dots\}$, starting from $x_0 > 1$. Using the renewal structure of the fragmentation process, we obtain

the recursive identity in distribution

$$X_{h+1}(x_0) \stackrel{d}{=} x_0 \cdot \mathbf{I}(U_1 x_0 \leq 1) + X_h^{(1)}(U_1 x_0) \cdot \mathbf{I}(U_1 x_0 > 1), \quad h \geq 0, \quad X_0(x_0) = x_0.$$

Here, $X_h^{(1)}(\cdot) \stackrel{d}{=} X_h(\cdot)$ is a statistical copy of $X_h(\cdot)$ and $(U_1, X_h^{(1)}(\cdot))$ are mutually independent random variables. Let us briefly comment this identity: as $x_0 > 1$, splitting takes place at step one. As a result, if $U_1 x_0 > 1$ (when splitting continues), $X_{h+1}(x_0)$ coincides in law with $X_h^{(1)}(U_1 x_0)$ at step h , after rescaling properly x_0 by U_1 ; otherwise (if $U_1 x_0 \leq 1$), $X_{h+1}(x_0)$ remains in the initial state x_0 as it does not undergo further fragmentation. Letting $h \uparrow \infty$, the limiting fragments size $X_{x_0} := X_\infty(x_0)$, if this random variable exists, should thus satisfy the distributional equality

$$X_{x_0} \stackrel{d}{=} x_0 \cdot \mathbf{I}(U_1 x_0 \leq 1) + X_{U_1 x_0}^{(1)} \cdot \mathbf{I}(U_1 x_0 > 1). \quad (28)$$

Note that $X_{x_0} \in (1, x_0]$ as fragmentation stops just before X_{x_0} crosses the critical value $x_c = 1$ from above. With $\bar{F}_{x_0}(x) := \mathbf{P}(X_{x_0} > x)$ and if $\Pi(1/x_0)$ is the probability mass of U_1 within the interval $(0, 1/x_0)$, the functional equation to solve now reads

$$\bar{F}_{x_0}(x) = \mathbf{I}(x < x_0) \Pi(1/x_0) + \int_{1/x_0}^1 \bar{F}_{x_0/u}(x) \pi(u) du. \quad (29)$$

In terms of moments, with $\Phi_{x_0}(q) := \mathbf{E}(X_{x_0}^{-q})$ the moment function of $X_{x_0}^{-1}$, we obtain from (28) that $\Phi_{x_0}(q)$ solves the functional equation

$$\Phi_{x_0}(q) = a_{x_0}(q) + \int_{1/x_0}^1 \Phi_{x_0 u}(q) \pi(u) du \quad (30)$$

with $a_{x_0}(q) = x_0^{-q} \Pi(1/x_0)$. To solve this equation, we introduce the z -transform

$$\widehat{\Phi}_z(q) := \int_1^\infty x_0^{-(z+1)} \Phi_{x_0}(q) dx_0, \quad z > 0.$$

We obtain

$$\widehat{\Phi}_z(q) = \frac{1 - \phi(z+q)}{(z+q)(1 - \phi(z))}. \quad (31)$$

To see this, let $\widehat{a}_z(q) := \int_1^\infty x_0^{-(z+1)} a_{x_0}(q) dx_0$. From (30), we get

$$\widehat{\Phi}_z(q) = \frac{\widehat{a}_z(q)}{1 - \phi(z)}.$$

Upon the change of variables $x_0 = e^t$, the functions $\widehat{\Phi}_z(q)$ and $\widehat{a}_z(q)$ are indeed nothing but the Laplace transforms of the functions $t > 0 \rightarrow \Phi_{e^t}(q)$ and $a_{e^t}(q)$, respectively. Using this remark and a subsequent integration by parts, we easily obtain $\widehat{a}_z(q) = [1 - \phi(z+q)]/(z+q)$.

The large t (and then x_0) behavior of $\Phi_{x_0}(q)$ can be read from the singularity analysis of its z -transform $\widehat{\Phi}_z(q)$.

Example 3. Suppose U_1 is uniform (as in the Dirichlet- $D_2(1)$ model). Then $\Pi(1/x_0) = x_0^{-1}$ and $\widehat{a}_z(q) = \int_0^\infty e^{-zt} e^{-t(q+1)} dt = \frac{1}{z+(q+1)}$. So

$$\widehat{\Phi}_z(q) = \frac{1}{1+q} z^{-1} + \frac{q}{q+1} (z + (q+1))^{-1}.$$

As a result, $\Phi_{x_0}(q) = \frac{1}{1+q} + \frac{q}{1+q} x_0^{-(q+1)}$ and the moment function of X_{x_0} is explicit in this case. In particular, the mean value of X_{x_0} is $\Phi_{x_0}(-1) = 1 + \log x_0$, $x_0 > 1$, the second moment being $\Phi_{x_0}(-2) = 2x_0 - 1$. Note also that if $q > 0$, $\Phi_{x_0}(q) \rightarrow_{x_0 \uparrow \infty} \frac{1}{1+q}$, showing that $X_{x_0}^{-1}$ converges in distribution to a uniform random variable. Finally, the rescaled random variable $\widetilde{X}_{x_0} := X_{x_0}/x_0$ is $(1/x_0, 1)$ -valued. With $\widetilde{\Phi}_{x_0}(q) := \mathbf{E}[\widetilde{X}_{x_0}^q]$ its moment function, we get $\widetilde{\Phi}_{x_0}(q) = \frac{x_0^{-q}}{1-q} - \frac{q}{1-q} x_0^{-1}$. Note that $\widetilde{\Phi}_{x_0}(q)$ is still dependent on x_0 . This translates the fact that it is not true that for size-dependent fragmentations $X_{x_0} \stackrel{d}{=} x_0 X_1$, where X_1 is fragments size starting from a size-1 interval.

3.2. The Dean–Majumdar tree

The DM fragmentation model gives rise to a tree some statistical features of which are interesting. Incidentally, this model has an early history in Computer Science (for binary trees, see [10] and the References therein).

3.2.1. Internal nodes

First we start with the cumulative number of splits (internal nodes of the DM tree), starting with an interval of length x_0 , say $\overline{N}(x_0)$. From the renewal structure of the DM process, we have

$$\overline{N}(x_0) \stackrel{d}{=} 0 \cdot \overline{B}_1(x_0) + B_1(x_0) \left(1 + \sum_{k=1}^b \overline{N}^{(k)}(U_k x_0) \cdot \mathbf{I}(U_k x_0 > 1) \right).$$

Let us first consider the expected value $\mu(x_0) := \mathbf{E}(\overline{N}(x_0))$. If $x_0 > 1$, we obtain

$$\mu(x_0) = 1 + b \int_{1/x_0}^1 \mu(ux_0) \pi(u) du. \quad (32)$$

Defining $\widehat{\mu}(z) := \int_1^\infty x_0^{-(z+1)} \mu(x_0) dx_0 = \int_0^\infty e^{-zv} \mu(e^v) dv$, for values of $z > 1$ for which $\phi(z) < \frac{1}{b}$, we get

$$\widehat{\mu}(z) = \frac{1}{z} \frac{1}{1 - b\phi(z)}. \quad (33)$$

Thus, the function $x_0 \rightarrow \mu(x_0)$ is increasing, locally bounded with

$$\frac{1}{x_0} \mu(x_0) \rightarrow_{x_0 \uparrow \infty} - \frac{1}{b\phi'(1)}.$$

Example 4. If U_1 is uniform, $b = 2$, $\widehat{\mu}(z) = -z^{-1} + 2(z-1)^{-1}$ and so $\mu(x_0) = 2x_0 - 1$.

The function $\widehat{\mu}(z)$ has at least two simple poles at $z_0 = 0$ and $z_1 = 1$. Although it can have other (eventually complex) poles, $z_1 = 1$ is the pole with largest real part. Let (z_2, z_2^*) be the pair of complex conjugate pole with the next largest positive real part (i.e. with $0 < \delta := \operatorname{Re}(z_2) < 1$), if it exists. Then keeping only the leading corrections, the asymptotic large x_0 behavior of $\mu(x_0)$ reads

$$\mu(x_0) \sim A_1 x_0 + A_2 x_0^{z_2} + A_2^* x_0^{z_2^*},$$

where $A_k := -1/[bz_k \phi'(z_k)]$.

3.2.2. The DM phase transition: fluctuations of internal nodes

Consider the log-Laplace transform of $\overline{N}(x_0)$, namely $\psi_{x_0}(\lambda) := \log \mathbf{E}[e^{\lambda \overline{N}(x_0)}]$. If $x_0 > 1$, with $\pi(u_1, \dots, u_b)$ the joint density of (U_1, \dots, U_b) , we have

$$\psi_{x_0}(\lambda) = \lambda + \log \int \dots \int e^{\sum_{k=1}^b \psi_{u_k x_0}(\lambda) \mathbf{I}(u_k x_0 > 1)} \pi(u_1, \dots, u_b) du_1 \dots du_b.$$

Differentiating twice with respect to λ and putting $\lambda = 0$ gives the variance $\sigma^2(x_0)$ of $\overline{N}(x_0)$. It satisfies

$$\sigma^2(x_0) = h(x_0) + b \int_{1/x_0}^1 \sigma^2(u x_0) \pi(u) du, \quad (34)$$

where $h(x_0) = \mathbf{E}[(\sum_{k=1}^b \{\mu(U_k x_0) - \mathbf{E}[\mu(U_k x_0)]\})^2]$. If x_0 is large, using the large x_0 expansion of $\mu(x_0)$ gives

$$h(x_0) \sim B_2 x_0^{2z_2} + B_2^* x_0^{2z_2^*} + B_3 x_0^{(z_2 + z_2^*)},$$

where the B_k are constants.

Defining $\widehat{\sigma}^2(z) := \int_1^\infty x_0^{-(z+1)} \sigma^2(x_0) dx_0 = \int_0^\infty e^{-zv} \sigma^2(e^v) dv$, for values of $z > 1$ for which $\phi(z) < \frac{1}{b}$, we obtain

$$\widehat{\sigma}^2(z) = \frac{\widehat{h}(z)}{1 - b\phi(z)}. \quad (35)$$

Now, $\widehat{h}(z) := \int_1^\infty x_0^{-(z+1)} h(x_0) dx_0$ has poles at $z = 2z_2, 2z_2^*$ and $2\operatorname{Re}(z_2) = 2\delta$. So, if $\delta > \frac{1}{2}$, this pole will be the dominant one in the expression of $\widehat{\sigma}^2(z)$ and so, up to log-periodic oscillations

$$\begin{aligned} \sigma^2(x_0) &\approx x_0 & \text{if } \delta < \frac{1}{2}, \\ \sigma^2(x_0) &\approx x_0^{2\delta} & \text{if } \delta > \frac{1}{2}, \end{aligned}$$

which is the signature of a phase transition on the variance. As observed by Dean and Majumdar, large fluctuations are thus obtained when $\delta > \frac{1}{2}$.

Defining the scaled process $\widetilde{N}(x_0) := (\overline{N}(x_0) - \mu(x_0))/\sigma(x_0)$, one could also prove that, as $x_0 \rightarrow \infty$, the following convergence in distribution holds:

$$\begin{aligned} \widetilde{N}(x_0) &\xrightarrow{d} \mathcal{N}(0, 1) & \text{if } \delta < \frac{1}{2}, \\ \widetilde{N}(x_0) &\xrightarrow{d} \mathcal{S}(1/\delta) & \text{if } \delta > \frac{1}{2}. \end{aligned}$$

Here, $\mathcal{N}(0, 1)$ is the standard normal law and $\mathcal{S}(1/\delta)$ a symmetric-stable law on \mathbb{R} with tail index $1/\delta \in (0, 2)$.

Example 5. (1) Suppose U_1 is a beta($\theta, (b-1)\theta$) random variable, hence with moment function $\phi(q) = \frac{\Gamma(\theta+q)}{\Gamma(\theta)} \frac{\Gamma(b\theta)}{\Gamma(b\theta+q)}$. If $\theta = 1$, the equation $\phi(z) = 1/b$ takes the polynomial form

$$(z+b-1)(z+b-2)\cdots(z+1) = b!,$$

the second largest pair of zeroes (z_2, z_2^*) of which is searched for. The quantity $\text{Re}(z_2)$ increases monotonically with b and when $b \geq b_c = 27$, the real part of the second largest zero is larger than $\frac{1}{2}$ with $z_2 = 0.5170 + 2.1789i$ when $b = 27$.

(2) Suppose U_1 has log-gamma density $\pi(u) = \frac{1}{(d-1)!} (-\log u)^{d-1}$ arising in the random splitting of a d -dimensional cube into 2^d sub-cubes. Then $\phi(q) = (1+q)^{-d}$ and $b=2^d$. The equation $1-b\phi(z)=0$ reads $1-2^d(1+z)^{-d}=0$. Looking for the second largest zero of $(1+z)^d - 2^d = 0$, one gets $z_2 = -1 + 2e^{2i\pi/d}$, with $\delta = -1 + 2 \cos(2\pi/d)$. The value of d for which $2\delta > 1$ is $d_c = \lceil 8.693 \dots \rceil + 1 = 9$ and so $b_c = 2^{d_c} = 512$. Large fluctuations are expected to arise in 9-dimensional space and beyond.

Let us now study some aspects of the DM partition function.

3.2.3. Leaves and partition function of the DM tree

First, consider the limiting number of available fragments (the number of leaves or external nodes of the DM tree), as those that never further split. Starting with an interval of size $x_0 > 1$, the renewal nature of the process gives

$$N(x_0) \stackrel{d}{=} \sum_{k=1}^b [N^{(k)}(U_k x_0) \cdot \mathbf{I}(U_k x_0 > 1) + \mathbf{I}(U_k x_0 \leq 1)].$$

Assume $x_0 > 1$. Then, if $m(x_0) := \mathbf{E}(N(x_0))$, we get $m(x_0) = b\Pi(1/x_0) + b \int_{1/x_0}^1 \mu(ux_0)\pi(u) du$.

Defining $\widehat{m}(z) := \int_1^\infty x_0^{-(z+1)} m(x_0) dx_0 = \int_0^\infty e^{-zv} m(e^v) dv$, for values of $z > 1$ for which $\phi(z) < \frac{1}{b}$, we get

$$\widehat{m}(z) = \frac{b(1 - \phi(z))}{z[1 - b\phi(z)]}. \quad (36)$$

There is no pole at $z=0$ and $z_1=1$ is a simple dominant pole; as a result, for large x_0 , $m(x_0) \sim \frac{-(b-1)}{b\phi'(1)} x_0$ with potential correcting terms if (z_2, z_2^*) exists.

Let now $Z(\beta, x_0)$ be the limiting partition function in the DM model. It satisfies

$$Z(\beta, x_0) \stackrel{d}{=} \sum_{k=1}^b [Z^{(k)}(\beta, U_k x_0) \cdot \mathbf{I}(U_k x_0 > 1) + (U_k x_0)^\beta \cdot \mathbf{I}(U_k x_0 \leq 1)]. \quad (37)$$

Looking at the average value of $Z(\beta, x_0)$, we get the functional equation

$$z_1(\beta, x_0) = b \int_{1/x_0}^1 z_1(\beta, ux_0)\pi(u) du + bx_0^\beta \mathbf{E}[U_1^\beta \cdot \mathbf{I}(U_1 x_0 \leq 1)].$$

Define $\widehat{z}_1(\beta, z) := \int_1^\infty x_0^{-(z+1)} z_1(\beta, x_0) dx_0$. To compute this quantity, observing that

$$\mathbf{E}[U_1^\beta \cdot \mathbf{I}(U_1 x_0 \leq 1)] = \int_0^{1/x_0} u^\beta \pi(u) du$$

we first need to evaluate the following integral:

$$I := \int_1^\infty x_0^{-(z+1-\beta)} dx_0 \int_0^{1/x_0} u^\beta \pi(u) du.$$

Integrating by parts, with $z > \beta$, we find $I = [\phi(\beta) - \phi(z)]/[z - \beta]$. Finally, with $z > \beta \vee 1$, we obtain

$$\widehat{z}_1(\beta, z) = \frac{b(\phi(\beta) - \phi(z))}{(z - \beta)[1 - b\phi(z)]}.$$

If $\beta < 1$, the dominant singularity of $\widehat{z}_1(\beta, z)$ is at $z = 1$, so for large x_0

$$z_1(\beta, x_0) \sim \frac{-(b\phi(\beta) - 1)}{b(1 - \beta)\phi'(1)} x_0.$$

If $\beta > 1$, $\widehat{z}_1(\beta, z)$ is defined for $z > \beta$ only. Defining $z' = z - \beta > 0$, we obtain

$$\widehat{z}_1(\beta, z' + \beta) = \frac{b(\phi(\beta) - \phi(z' + \beta))}{z'[1 - b\phi(z' + \beta)]},$$

where $\widehat{z}_1(\beta, z' + \beta) = \int_1^\infty x_0^{-(z'+1)} x_0^{-\beta} z_1(\beta, x_0) dx_0$ is the Laplace transform of $x_0 \rightarrow x_0^{-\beta} z_1(\beta, x_0)$. It has a pole at $z' = 1 - \beta$ and so for large x_0 : $x_0^{-\beta} z_1(\beta, x_0) \sim C x_0^{1-\beta}$, where $C = \frac{-(b\phi(\beta)-1)}{b(1-\beta)\phi'(1)}$.

To summarize, we obtain the limiting behavior

$$\frac{1}{x_0} z_1(\beta, x_0) \rightarrow_{x_0 \uparrow \infty} \frac{-(b\phi(\beta)-1)}{b(1-\beta)\phi'(1)} \quad (38)$$

for all values of β such that $\phi(\beta) < \infty$.

4. Brennan–Durrett-type model

We now present here a new fragmentation model for which the splitting probability is an increasing algebraic function of initial stick's size x_0 . In this model, x_0 is bound to belong to the interval $(0, 1)$ in a fragmentation process of finite-size items. This model is partly solvable explicitly. Although there is no explicit precise connection, this model was inspired from the Brennan–Durrett construction [2], in continuous time.

We first derive the fragments' size distribution before studying the induced tree structure.

4.1. Fragment size distribution

Suppose first that $x_0 \in (0, 1)$ and also that the splitting probability is $p(x_0) = x_0^a$, $a > 0$, increasing with initial fragment size x_0 . When it splits, each fragment splits into b daughter fragments with exchangeable

random sizes, as usual. The process is iterated independently on each sub-fragments, as usual. Under these assumptions, the moment function of the fragments' size now satisfies the integral equation

$$\Phi_{x_0}(q) = x_0^q(1 - x_0^a) + x_0^a \int_0^1 \Phi_{ux_0}(q) \pi(u) du. \quad (39)$$

In particular, $\mathbf{E}[X_{x_0}] := x(x_0)$ satisfies $x(x_0) = x_0(1 - x_0^a) + x_0^a \int_0^1 x(ux_0) \pi(u) du$. Upon iterating, $x(x_0)$ can be searched under the form

$$x(x_0) = x_0 \left(1 - \sum_{m \geq 1} a_m x_0^{ma} \right),$$

where the coefficients a_m are to be determined. Putting this expression into the functional equation which $x(x_0)$ satisfies, we obtain

$$\begin{aligned} x_0 \left(1 - \sum_{m \geq 1} a_m x_0^{ma} \right) &= x_0(1 - x_0^a) + x_0^{a+1} \left[\frac{1}{b} - \sum_{m \geq 1} a_m x_0^{ma} \phi(ma + 1) \right] \\ &= x_0 \left(1 - \left(1 - \frac{1}{b} \right) x_0^a - \sum_{m \geq 1} a_m x_0^{(m+1)a} \phi(ma + 1) \right). \end{aligned}$$

Identifying the terms, we get $a_1 = 1 - 1/b$, $a_m = a_{m-1} \phi(1 + (m-1)a)$, $m \geq 2$, leading to

$$a_m = a_1 \prod_{k=1}^{m-1} \phi(1 + ka), \quad m \geq 2.$$

Under our assumptions, the series with positive terms $\sum_{m \geq 1} a_m x_0^{ma}$ is always convergent because the ratio $\frac{a_{m+1} x_0^{(m+1)a}}{a_m x_0^{ma}} = \phi(1 + ma) x_0^a \rightarrow_{m \uparrow \infty} 0 < 1$. Note that for x_0 close to zero, $x(x_0)$ grows linearly with x_0 , whatever the value of a . Concerning $\Phi_{x_0}(q)$ itself, proceeding similarly, we find more generally

$$\Phi_{x_0}(q) = x_0^q \left(1 - \sum_{m \geq 1} a_m(q) x_0^{ma} \right), \quad (40)$$

where, with $q > q_*$

$$\begin{aligned} a_1(q) &= 1 - \phi(q), \\ a_m(q) &= a_1(q) \prod_{k=1}^{m-1} \phi(q + ka), \quad m \geq 2. \end{aligned}$$

When $0 > q_* > -\infty$ exists, $\Phi_{x_0}(q)$ has a pole there and so the fragments' size density f_{x_0} satisfies

$$f_{x_0}(x) \sim_{x \downarrow 0} A(x_0) x^{-(1+q_*)}.$$

The law of X_{x_0} has clearly an atom at $x = x_0$ with mass $1 - x_0^a$ when splitting does not take place. Note also that when $a \downarrow 0$, for $q > 0$, $\Phi_{x_0}(q) \rightarrow 0$; this means that $X_{x_0} \xrightarrow{\text{a.s.}} 0$: as splitting probability tends to

1, the fragments size tend to zero. When $a \uparrow \infty$, the splitting probability tends to 0 and one may check that $\Phi_{x_0}(q) \sim x_0^q$: the initial fragment is left unchanged for ever. Finally, the range of X_{x_0} is $[0, x_0]$. Defining $\tilde{X}_{x_0} := X_{x_0}/x_0$, the range of this scaled random variable is now $[0, 1]$; its moment function is $\tilde{\Phi}_{x_0}(q) := \mathbf{E}[\tilde{X}_{x_0}^q] = 1 - \sum_{m \geq 1} a_m(q) x_0^{ma}$. As for the DM model, it is not true that $X_{x_0} \stackrel{d}{=} x_0 X_1$ and so X_{x_0} is not scaling.

If $q = 0$, one may check that $\tilde{\Phi}_{x_0}(0) = 1$, whereas $\tilde{\Phi}_{x_0}(q) \rightarrow_{q \uparrow \infty} 1 - x_0^a$. The random variable \tilde{X}_{x_0} has an atom at $x = 1$ with mass $1 - x_0^a$ which is the probability not to split.

As a moment function of a $[0, 1]$ -valued random variable, function $q > 0 \rightarrow \tilde{\Phi}_{x_0}(q)$ takes values in $(1 - x_0^a, 1)$ for any $a > 0$ and $x_0 \in (0, 1)$. This entails that, for any $x := x_0^a \in (0, 1)$, $a > 0$, the series

$$h_q(x) := 1 + \sum_{m \geq 2} x^{m-1} \prod_{k=1}^{m-1} \phi(q + ka)$$

is convergent for any $q > 0$ and takes values in the interval $(0, (1 - \phi(q))^{-1})$. This fact will be used in the sequel.

Example 6. If $a = 1$, $b = 2$ and $\phi(q) = (1 + q)^{-1}$, we find

$$a_m(q) = a_1(q) \prod_{k=1}^{m-1} \frac{1}{q + k + 1} = \frac{q}{\prod_{k=1}^m (q + k)}$$

and so $a_m := a_m(1) = \frac{1}{(m+1)!}$, leading to the average fragments' length $x(x_0) = 1 + 2x_0 - e^{x_0} \in (0, x_0)$. The function $x_0 \rightarrow x(x_0)$ is not monotone. It is maximal when $x_0 = \log 2$ and $x(\log 2) = 2 \log 2 - 1 = 0.3863$. With $a_m(2) = \frac{4}{(m+2)!}$, the second moment is found to be $\Phi_{x_0}(2) = 3x_0^2 + 4(1 + x_0 - e^{x_0})$.

Another interesting example arises from the random splitting of the cube in dimension $d = 3$. Here, $b = 2^3$, $\phi(q) = (1 + q)^{-3}$ and assuming $a = \frac{2}{3}$, the splitting probability $x_0^{2/3}$ grows like the area of a face of the cube. This model could be relevant to polymer degradation.

4.2. Number of fragments, number of splitting events in the BD tree

First we start with the limiting number of leaves fragments in the process, starting with an interval of length x_0 , say $N(x_0)$. It satisfies

$$N(x_0) \stackrel{d}{=} \bar{B}_1(x_0) + B_1(x_0) \sum_{k=1}^b N^{(k)}(U_k x_0).$$

Let us first consider the expected value $m(x_0) := \mathbf{E}(N(x_0))$. If $x_0 \in (0, 1)$, we have

$$m(x_0) = 1 - x_0^a + b x_0^a \int_0^1 m(u x_0) \pi(u) du. \quad (41)$$

Searching for solutions under the form

$$m(x_0) = 1 + \sum_{m \geq 1} b_m x_0^{ma} \quad (42)$$

for some unknown sequence b_m , $m \geq 1$, we obtain

$$b_1 = b - 1,$$

$$b_m = b_1 b^{m-1} \prod_{k=1}^{m-1} \phi(ka), \quad m \geq 2.$$

For each $a > 0$, $x_0 \in (0, 1)$, the ratio $b_{m+1}x_0^{(m+1)a}/b_mx_0^{ma} = \phi(ma)bx_0 \rightarrow_{m \uparrow \infty} 0 < 1$. By Cauchy–d’Alembert rule, the sequence $m_n(x_0) := 1 + \sum_{m=1}^n b_mx_0^{ma}$ is therefore convergent, with $m(x_0)$ as the limit. At last, for small x_0 , we have $m(x_0) \sim 1 + (b-1)x_0^a$.

As $a \downarrow 0$, clearly $m(x_0) \rightarrow \infty$, whereas $m(x_0) \rightarrow 1$ as $a \uparrow \infty$. As $m(x_0)$ is a continuous and monotone function of a , one could have suspected the existence of a critical value a_c of a defined by $a_c = \inf(a > 0 : m(x_0) < \infty)$ so that the above formal series expansion of $m(x_0)$ is convergent for $a > a_c$. The above argument shows that $a_c = 0$, for any $x_0 \in (0, 1)$. Thus, in the BD fragmentation model, the expected number of leaves is finite in any case and there is no critical or super-critical phase as in the KPG model.

Let us now derive some bounds on $m(x_0)$. Recall that the series $h_q(x) = 1 + \sum_{m \geq 2} x^{m-1} \prod_{k=1}^{m-1} \phi(q+ka)$ is convergent for any $a, q > 0$, $x \in (0, 1)$. Furthermore, assuming $q = a$, one may check that in particular, if $bx < 1$

$$h_a(bx) = 1 + \frac{1}{b(b-1)\phi(a)x^2} \sum_{m \geq 3} b_mx^m$$

$$= \frac{1}{b(b-1)\phi(a)x^2} [m(x^{1/a}) - 1 - (b-1)x]$$

takes values in $(0, 1/(1 - \phi(a)))$.

We conclude the following: fix $a > 0$, then for any $x_0 < b^{-1/a} < 1$, we have the bounds

$$1 + (b-1)x_0^a < m(x_0) < 1 + (b-1)x_0^a \left(1 + bx_0^a \frac{\phi(a)}{1 - \phi(a)} \right). \quad (43)$$

If $a \rightarrow 0$, the range of values of x_0 for which $m(x_0)$ is bounded above in this way shrinks to 0.

Example 7. Let $b = 2$ and $\phi(q) = (1+q)^{-1}$. Assuming $a = 1$, we find

$$b_m = 2^{m-1} \prod_{k=1}^{m-1} \frac{1}{k+1} = \frac{2^{m-1}}{m!}$$

and so the expected number of leaves is $m(x_0) = 1 + \frac{1}{2}(e^{2x_0} - 1) < \infty$. When $x_0 < 2^{-1}$, we indeed have $1 + x_0 < m(x_0) < 1 + x_0(1 + 2x_0)$.

Consider now the variance $\sigma^2(x_0)$ of $N(x_0)$. Proceeding as for the Dean–Majumdar model, we get

$$\sigma^2(x_0) = x_0^a \left(h(x_0) + b \int_0^1 \sigma^2(ux_0) \pi(u) du \right), \quad (44)$$

where $h(x_0) = \mathbf{E}[(\sum_{k=1}^b \{m(U_k x_0) - \mathbf{E}[m(U_k x_0)]\})^2]$ is similar to the one in the Dean–Majumdar model.

When x_0 is small, to the two leading terms $m(x_0) \sim_{x \downarrow 0} 1 + (b-1)x_0^a$. As a result, if $a \neq 1$, $h(x_0) \sim Cx_0^{2a}$ and $\sigma^2(x_0) \sim bCx_0^{3a}$ for some constant $C > 0$.

Note that if $a = 1$, the approximation $m(x_0) \sim 1 + b_1x_0 + b_2x_0^2$ to the three leading terms is needed. From this, we get $h(x_0) \sim Cx_0^4$ and so $\sigma^2(x_0) \sim bCx_0^5$, with very small fluctuations. Thus, at $a = 1$, this model exhibits a singular behavior for external nodes' variance.

For the total number of nodes (both external and internal), say $\mathcal{N}(x_0)$, proceeding similarly, we find

$$\mathbf{E}(\mathcal{N}(x_0)) = 1 + bx_0^a \int_0^1 \mathbf{E}(\mathcal{N}(ux_0))\pi(u) du.$$

So, with $c_1 = b$, $c_m = b^m \prod_{k=1}^{m-1} \phi(ka)$, $m \geq 2$, with $a > a_c = 0$, we get

$$\mathbf{E}(\mathcal{N}(x_0)) = 1 + \sum_{m \geq 1} c_m x_0^{ma} < \infty.$$

In the simplest example discussed throughout, it is $\mathbf{E}[\mathcal{N}(x_0)] = e^{2x_0}$.

Recalling that in b -ary trees, the number of internal nodes is $\bar{N}(x_0) = \mathcal{N}(x_0) - 1/b$, the average number of splitting events follows. It is

$$\mu(x_0) = \mathbf{E}[\bar{N}(x_0)] = \sum_{m \geq 1} b^{m-1} \prod_{k=1}^{m-1} \phi(ka) x_0^{ma} = \frac{1}{b-1} (\mathbf{E}[\mathcal{N}(x_0)] - 1). \quad (45)$$

In the simplest example discussed throughout, we find $\mathbf{E}[\bar{N}(x_0)] = \frac{1}{2}(e^{2x_0} - 1)$.

5. Concluding remarks

Three random fragmentation models involving stable and unstable fragments have been considered. Statistical properties of the limiting random partitions were shown to be highly sensitive on the dependency structure of the splitting probability on fragments sizes.

More precisely, a homogeneous model for which splitting probability is independent of fragments' size at each step was first investigated. Statistical features of the induced limiting fragmentation have been discussed: fragments' size distribution, partition function of mass, Rényi's average fragments' size, sizes of smallest and largest fragments. Fragments' size distribution presents generically an algebraic divergence of its small-size tail and scaling properties. This model exhibits a phase transition between sub-critical and super-critical regimes which was outlined. A first extension where splitting probability depends on parental mass fraction transmitted to daughter fragments was next studied. This possibly provokes loss of mass to a phase of zero-size particles.

Next, two inhomogeneous fragmentation models with fragments length-dependent splitting probability were considered. In the first model, a fragment with initial size x_0 splits with probability one if its size exceeds some cutoff value $x_c = 1$, fragmentation further proceeding for each sub-fragments whose sizes are bigger than x_c only. This process naturally terminates. Fragments' size distribution, partition function of mass and statistical properties of the associated finite fragmentation tree were discussed together with Dean–Majumdar phase transition on the variance of its internal nodes (as $x_0 \rightarrow \infty$). A second new model was finally addressed in which the size x_0 of the initial fragment to be split is bounded by 1. The splitting

probability was assumed to increase algebraically with fragments' size at each step. This process was shown to terminate with probability 1 and to present some solvable statistical features, especially when $x_0 \rightarrow 0$. Fragments' size distribution, together with information on the underlying tree structure, were obtained. In both inhomogeneous models, fragment size distribution fails to scale.

References

- [1] D. Beysens, X. Campi, E. Pefferkorn (Eds.), *Fragmentation Phenomena*, Les Houches Series, World Scientific Publishing Co., Singapore, 1995.
- [2] M.D. Brennan, R. Durrett, Splitting intervals, II: limit laws for lengths, *Probab. Theory Related Fields* 75 (1) (1987) 109–127.
- [3] D.S. Dean, S.N. Majumdar, Phase transition in a random fragmentation problem with applications to computer science, *J. Phys. A: Math. Gen.* 35 (32) (2002) L501–L507.
- [4] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 2, Wiley, New York, 1971.
- [5] T.E. Harris, *The theory of branching processes*, *Die Grundlehren der Mathematischen Wissenschaften*, Bd. 119, Springer, Berlin; Prentice-Hall, Inc., Englewood Cliffs, NJ, 1963.
- [6] T. Huillet, B. Jeannet, Finitely generated multifractals can display phase transitions, *J. Phys. I* 6 (1996) 245–256.
- [7] P.L. Krapivsky, E. Ben-Naim, I. Grosse, Stable distributions in stochastic fragmentation, *J. Phys. A: Math. Gen.* 37 (8) (2004) 2863–2880.
- [8] P.L. Krapivsky, I. Grosse, E. Ben-Naim, Scale invariance and lack of self-averaging in fragmentation, *Phys. Rev. E* 61 (2000) R993–R996.
- [9] S. Redner, in: H.J. Hermann, S. Roux (Eds.), *Statistical Models for the Fracture of Disordered Media*, Elsevier Science, New York, 1990.
- [10] M. Sibuya, Y. Itoh, Random sequential bisection and its associated binary tree, *Ann. Inst. Statist. Math.* 39 (Part A) (1987) 69–84.