Statistical analysis of self-similar conservative fragmentation chains

Marc Hoffmann, Nathalie Krell

Statistical analysis of self-similar conservative fragmentation chains

Marc Hoffmann* and Nathalie Krell†

March 6, 2008

Abstract

We explore statistical inference in self-similar conservative fragmentation chains, when only (approximate) observations of the size of the fragments below a given threshold are available. This framework, introduced by Bertoin and Martinez, is motivated by mineral crushing in mining industry.

The underlying estimated object is the step distribution of the random walk associated to a randomly tagged fragment that evolves along the genealogical tree representation of the fragmentation process. We compute upper and lower rates of estimation in a parametric framework, and show that in the non-parametric case, the difficulty of the estimation is comparable to ill-posed linear inverse problems of order 1 in signal denoising.

Keywords: fragmentation chains; parametric, non-parametric estimation, key renewal theorem.
Mathematical Subject Classification: 60J80, 60J25,62G05, 62M05 .
e-mail: marc.hoffmann@univ-mlv.fr, nathalie.krell@courriel.upmc.fr

*Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées, CNRS-UMR 8050, 5, boulevard Descartes, 77454 Marne-la-Vallée Cedex 2, France.

†Université Paris 6, Laboratoire de Probabilités et Modèles Aléatoires, CNRS-UMR 7599, 16, rue Clisson, 75013 Paris Cedex, France.

1 Introduction

1.1 Motivation

Random fragmentation models, commonly used in a variety of physical models, lay their theoretical roots in the works of Kolmogorov [11] and Filippov [9] (see also [1, 5, 13, 14] and the references therein). Informally, we imagine an object that falls apart randomly as time passes. The resulting particles break independently of each other in an independent and self-similar way. A thorough account on random fragmentation processes and chains is given in the book by Bertoin [5], a key reference for the paper.

In this work, we take the perspective of statistical inference. We focus on the quite specific class of *self-similar fragmentation chains*. The law of the fragmentation is then entirely determined by its dislocation measure and its index of self-similarity, which govern the way and the rate at which the fragments split. If one is allowed to observe the whole fragmentation process up to some fixed time, then the statistical problem is somehow degenerate[1]. We postulate a more realistic observation scheme, motivated by mining industry, where the goal is to separate metal from non valued components in large mineral blocks by a series of blasting, crushing and grinding. In this setting, one rather observes approximatively the fragments arising from an initial block of size m only when they reach a size smaller than some screening threshold, say $\eta > 0$. Asymptotics are taken as the ratio $\varepsilon := \eta/m$ vanishes. See Bertoin and Martinez [7] and the references therein.

1.2 Organization and of the paper

In Section 2, we recall some well known facts about the construction of conservative fragmentation chains, following closely the book by Bertoin [5]. For statistical purposes, our main tool is the empirical measure $\mathcal{E}_\varepsilon$ of the size of fragments when they reach a size smaller than a threshold ε in the limit $\varepsilon \to 0$. We highlight the fact that $\mathcal{E}_\varepsilon$ captures information about the dislocation measure through the Lévy measure π of a randomly tagged fragment associated to the fragmentation process.

[1] in the sense that it can be mapped into relatively standard equivalent inference problems such as probability distribution estimation from independent observations, see Section 4.4

In Section 3, we give a rate of convergence for the empirical measure $\mathcal{E}_\varepsilon$ toward its limit in Theorem 3.1, extending former results (under more stringent assumptions) of Bertoin and Martinez [7]. The rate is of the form $\varepsilon^{1/2-\ell(\pi)}$, where $\ell(\pi) > 0$ can be made arbitrarily small under adequate exponential moment conditions for π. We add-up the more realistic framework of observations with limited accuracy, where each fragment is actually known up to a systematical stochastic error of order $\sigma \ll \varepsilon$.

In Section 3, we construct estimators related to functionals of π in the absolute continuous case. In the parametric case (Theorem 3.4), we establish that the best achievable rate is $\varepsilon^{1/2}$ in the particular case of binary fragmentations, where a particle splits into two blocks at each step exactly. We construct a convergent estimator in a general setting (Theorem 3.2) with an error of order $\varepsilon^{1/2-\ell'(\pi)}$, for another $\ell'(\pi) > 0$ that can be made arbitrarily small under appropriate assumptions on the density of π near 0 and $+\infty$. In the non-parametric case, we construct an estimator that achieves (Theorem 3.5) a rate of the form $(\varepsilon^{1-\ell''(\pi)})^{s/(2s+3)}$, where $s > 0$ is the local smoothness of the density of π, up to appropriate rescaling. Except for the factor $\ell''(\pi) > 0$, we obtain the same rate as for ill-posed linear inverse problems of degree 1. We suggest a simple interpretation of this result in terms of the asymptotic form of $\mathcal{E}_\varepsilon$ in the discussion Section 4, appended with further remarks about Theorems 3.1, 3.2 and 3.5.

An appendix (Section 6) recalls sharp results on the key renewal theorem from Sgibnev [16] that are used to derive Theorem 3.1.

2 Statistical model

2.1 Fragmentation chains

Let $X = \big(X(t), t \geq 0\big)$ be a fragmentation chain with state space

$$\mathcal{S}^{\downarrow} := \big\{\mathbf{s} = (s_1, s_2, \ldots),\ s_1 \geq s_2 \geq \ldots \geq 0,\ \sum_{i=1}^{\infty} s_i \leq 1\big\}.$$

We assume that X has self-similar parameter $\alpha \geq 0$. For well-definiteness, see *e.g.* Bertoin [5], the following mild assumptions on the dislocation measure $\nu(d\mathbf{s})$ of X are in force throughout the paper:

Assumption A. *We have* $\nu\big((1,0,\ldots)\big) = 0$ *and* $\nu\big(s_1 \in (0,1)\big) > 0$. *Moreover, for every* $\varepsilon > 0$*:* $\int_{\mathcal{S}^{\downarrow}} \sum_{i=1}^{\infty} 1_{\{s_i > \varepsilon\}} \nu(d\mathbf{s}) < \infty$.

We denote by $\mathbb{P}_m$ the law of X started from the initial configuration $(m, 0, \ldots)$ with $m \in (0,1]$. Under $\mathbb{P}_m$, X is a Markov process and its evolution can be described as follows: a fragment with size x lives for an exponential time with parameter $x^\alpha \nu(\mathcal{S}^\downarrow)$ and then splits and gives rize to a family of smaller fragments distributed as $x\xi$, where ξ is distributed according to $\nu(\bullet)/\nu(\mathcal{S}^\downarrow)$. Under $\mathbb{P}_m$, the law of X is entirely determined by α and $\nu(\bullet)$.

We will repeatedly use the representation of fragmentation chains as random infinite marked trees. Let

$$\mathcal{U} := \bigcup_{n=0}^{\infty} \mathbb{N}^n$$

denote the infinite genealogical tree (with $\mathbb{N}^0 := \{\emptyset\}$) associated to X as follows: to each node $u \in \mathcal{U}$, we set a mark

$$(\xi_u, a_u, \zeta_u),$$

where ξ_u is the size of the fragment labelled by u, a_u is its birthtime and ζ_u is its lifetime. We have the following identity between point measures on $(0, +\infty)$:

$$\sum_{i=1}^{\infty} 1_{\left\{X_i(t)>0\right\}} \delta_{X_i(t)} = \sum_{u \in \mathcal{U}} 1_{\left\{t \in [a_u, a_u+\zeta_u)\right\}} \delta_{\xi_u}, \quad t \geq 0,$$

with $X(t) = \big(X_1(t), X_2(t), \ldots\big)$, and where δ_x denotes the Dirac mass at x. Finally, X has the following branching property: for every fragment $\mathbf{s} = (s_1, \ldots) \in \mathcal{S}^\downarrow$ and every $t \geq 0$, the distribution of $X(t)$ given $X(0) = \mathbf{s}$ is the same as the decreasing rearrangement of the terms of independent random sequences $X^{(1)}(t)$, $X^{(2)}(t), \ldots$ where, for each i, $X^{(i)}(t)$ is distributed as $X(t)$ under $\mathbb{P}_{s_i}$.

2.2 Observation scheme

For simplicity, we assume from now on that $\nu(\mathcal{S}^\downarrow) = 1$. Keeping in mind the motivation of mineral crushing, we consider the fragmentation under[2] $\mathbb{P} := \mathbb{P}_1$, initiated with a unique block of size $m = 1$ and we observe the

[2] We often need to accomodate further random variables independent of X. Abusing notation slightly, we will still use the notation $\mathbb{P}$ without further notice, working tacitly on an appropriate enlargement of the original probability space.

process stopped at the time when all the fragments become smaller than some given threshold $\varepsilon > 0$, so we have data ξ_u, for every $u \in \mathcal{U}_\varepsilon$, with

$$\mathcal{U}_\varepsilon := \big\{u \in \mathcal{U},\ \ \xi_{u-} \geq \varepsilon,\ \xi_u < \varepsilon\big\},$$

where we denote by $u-$ the parent of the fragment labelled by u. We will further assume that the total mass of the fragments remains constant through time:

Assumption B. *(Conservative property). We have:* $\nu\big(\sum_{i=1}^{\infty} s_i = 1\big) = 1$.

We next consider the empirical measure

$$\mathcal{E}_\varepsilon(g) := \sum_{u \in \mathcal{U}_\varepsilon} \xi_u\, g(\xi_u/\varepsilon),$$

where $g(\bullet)$ is a test function. Indeed, under Assumption B, we have

$$\sum_{u \in \mathcal{U}_\varepsilon} \xi_u = 1 \ \ \mathbb{P}-\text{almost surely}, \tag{2.1}$$

so $\mathcal{E}_\varepsilon(g)$ appears as a weighted empirical version of $g(\bullet)$. Bertoin and Martinez show in [7] that under mild assumptions on $\nu(\bullet)$, the measure $\mathcal{E}_\varepsilon(g)$ converges to

$$\mathcal{E}(g) := \frac{1}{c(\nu)} \int_0^1 \frac{g(a)}{a} \int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i\, 1_{\{s_i < a\}} \nu(d\mathbf{s}) da$$

in probability, as $\varepsilon \to 0$, with $c(\nu) = -\int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i \log s_i\, \nu(d\mathbf{s})$, tacitly assumed to be well-defined. This suggests a strategy for recovering information about $\nu(\bullet)$ by picking suitable test functions $g(\bullet)$.

2.3 First estimates

From now on, we assume we have data

$$X_\varepsilon := \big(\xi_u,\ \ u \in \mathcal{U}_\varepsilon\big) \tag{2.2}$$

and we specialize in the estimation of $\nu(\bullet)$. Clearly, the data give no information about the self-similar parameter α that we consider as a nuisance parameter[3]. Assumptions A and B are in force. At this stage,

[3] See however Section 4.4 for auxiliary results about the inference on α.

we can relate $\mathcal{E}(g)$ to a more appropriate quantity by means of the so-called *tagged frament* approach.

The randomly tagged fragment. Let us first consider the homogenous case $\alpha = 0$. Assume we can "tag" a point at random –according to a uniform distribution– on the initial fragment and imagine we can follow the evolution of the fragment that contains this point.

Let us denote by $\big(\chi(t), t \geq 0\big)$ the process of the size of the fragment that contains the randomly chosen point. This fragment is a typical observation in our data set X_ε, and it appears at time

$$T_\varepsilon := \inf\big\{t \geq 0, \;\; \chi(t) < \varepsilon\big\}.$$

Bertoin [5] shows that the process $\zeta(t) := -\log \chi(t)$ is a subordinator, with Lévy measure:

$$\pi(dx) := e^{-x} \sum_{i=1}^{\infty} \nu(-\log s_i \in dx). \tag{2.3}$$

We can anticipate that the information we get from X_ε is actually information about the Lévy measure $\pi(dx)$ of $\zeta(t)$ throughout $\zeta(T_\varepsilon)$. The dislocation measure $\nu(d\mathbf{s})$ and $\pi(dx)$ are related by (2.3) which reads

$$\int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} s_i f(s_i) \nu(d\mathbf{s}) = \int_{(0,+\infty)} f(e^{-x}) \pi(dx), \tag{2.4}$$

for any suitable $f(\bullet) : [0,1] \to [0, +\infty)$. In particular, by Assumption B and the fact that $\nu(\mathcal{S}^\downarrow) = 1$, $\pi(dx)$ is a probability measure hence $\zeta(t)$ is a compound Poisson process. Informally, a typical observation takes the form $\zeta(T_\varepsilon)$, which is the value of a subordinator with Lévy measure $\pi(dx)$ at its first passage time strictly above $-\log \varepsilon$. The case $\alpha \neq 0$ is a bit more involved and reduces to the homogenous case by a time change.

In terms of the limit of the empirical measure $\mathcal{E}_\varepsilon(g)$, we equivalently have

$$\mathcal{E}(g) = \frac{1}{c(\pi)} \int_0^1 \frac{g(a)}{a} \pi(-\log a, +\infty)\, da = \frac{1}{c(\pi)} \int_0^{+\infty} g(e^{-x}) \pi(x, +\infty)\, dx,$$

with $c(\pi) = \int_{(0,+\infty)} x\, \pi(dx)$, the two representations being useful either way. This approach will prove technically convenient and will be detailed later on. Except in the binary case (a particular case of interest, see

Section 4.1), the knowledge of $\pi(\bullet)$ does not allow to recover $\nu(\bullet)$ in general.

Measurements with limited accuracy. It is unrealistic to assume that we can observe exactly the size ξ_u of the fragments. This becomes even more striking if the dislocation splits at a given time into infinitely many fragments of non zero size, a situation that we do not discard in principle. Therefore, we replace (2.2) by the more realistic observation scheme $X_{\varepsilon,\sigma} := \big(\xi_u^{(\sigma)},\ u \in \mathcal{U}_{\varepsilon,\sigma}\big)$ with

$$\mathcal{U}_{\varepsilon,\sigma} := \big\{u \in \mathcal{U},\ \xi_{u-}^{(\sigma)} \geq \varepsilon,\ \xi_u^{(\sigma)} < \varepsilon\big\},$$

and

$$\xi_u^{(\sigma)} := \xi_u + \sigma U_u. \tag{2.5}$$

The random variables $(U_u, u \in \mathcal{U})$ are identically distributed, and account for a systematic experimental noise in the measurement of X_ε, independent of X_ε. We assume furthermore that, for every $u \in \mathcal{U}$,

$$|U_u| \leq 1 \ \text{ and } \ \mathbb{E}[U_u] = 0.$$

The noise level $0 \leq \sigma = \sigma(\varepsilon) \ll \varepsilon$ is assumed to be known and represents the accuracy level of the statistician.

The observations $\xi_u + \sigma U_u$ are further discarded below a threshold $\sigma \leq t_\varepsilon \leq \varepsilon$ beyond which they become irrelevant, leading to the modified empirical measure

$$\mathcal{E}_{\varepsilon,\sigma}(g) := \sum_{u \in \mathcal{U}_{\varepsilon,\sigma}} 1_{\{\xi_u^{(\sigma)} \geq t_\varepsilon\}}\, \xi_u^{(\sigma)}\, g\big(\xi_u^{(\sigma)}/\varepsilon\big).$$

In the sequel, we take $t_\varepsilon = \gamma_0 \varepsilon$ for some (arbitrary) $0 < \gamma_0 < 1$.

3 Main results

We first exhibit explicit rates in the convergence $\mathcal{E}_\varepsilon(g) \to \mathcal{E}(g)$ as $\varepsilon \to 0$, extending Proposition 1.12 in Bertoin[4] [5]. We then turn to the estimation of $\pi(\bullet)$.

[4] See also Bertoin and Martinez [7].

3.1 A rate of convergence for the empirical measure

For $\kappa > 0$, we say that a spread-out[5] probability measure $\pi(dx)$ defined on $[0, +\infty)$ belongs to $\Pi(\kappa)$ if

$$\int_{[0,+\infty)} e^{\kappa x}\, \pi(dx) < +\infty,$$

appended with $\Pi(\infty) := \bigcap_{\kappa>0} \Pi(\kappa)$. For $m > 0$, define the class of continuous functions

$$\mathcal{C}(m) := \big\{ g(\bullet) : [0,1] \to \mathbb{R},\ \ \|g\|_\infty := \sup_x |g(x)| \leq m \big\},$$

and $\mathcal{C}'(m)$ the class of continuously differentiable functions $g(\bullet)$ such that $g' \in \mathcal{C}(m)$.

Theorem 3.1. *Grant Assumptions A and B. Let $0 < \kappa \leq \infty$ and assume that $\pi \in \Pi(\kappa)$.*

- *For every $m > 0$ and $0 < \mu < \kappa$, we have*

$$\sup_{g \in \mathcal{C}(m)} \mathbb{E}\left[\big(\mathcal{E}_\varepsilon(g) - \mathcal{E}(g) \big)^2 \right] = o\big(\varepsilon^{\mu/(\mu+1)}\big). \tag{3.1}$$

- *The convergence* (3.1) *remains valid if we replace $\mathcal{E}_\varepsilon(\bullet)$ by $\mathcal{E}_{\varepsilon,\sigma}(\bullet)$ and $\mathcal{C}(m)$ by $\mathcal{C}'(m)$, up to an additional error term:*

$$\sup_{g \in \mathcal{C}'(m)} \mathbb{E}\left[\big(\mathcal{E}_{\varepsilon,\sigma}(g) - \mathcal{E}_\varepsilon(g) \big)^2 \right] = \mathcal{O}(\sigma^2 \varepsilon^{-2}). \tag{3.2}$$

3.2 Statistical estimation

We study the estimation of $\pi(\bullet)$ by constructing estimators based on $\mathcal{E}_\varepsilon(\bullet)$ or rather $\mathcal{E}_{\varepsilon,\sigma}(\bullet)$. We need the following regularity assumption:

Assumption C. *The probability measure $\pi(dx)$ is absolutely continuous.*

We denote by $x \rightsquigarrow \pi(x)$ its density function. We distinguish two cases: the *parametric case*, where we estimate a linear functional of $\pi(\bullet)$ of the form

$$m_k(\pi) := \int_0^{+\infty} x^k \pi(x) dx, \ \ \text{for some } k \geq 1,$$

[5] We recall some properties on spread-out measures in the Appendix.

and the *non-parametric case*, where we estimate the function $x \rightsquigarrow \pi(x)$ pointwise. In that latter case, it will prove convenient to assess the local smoothness properties of $\pi(\bullet)$ on a logarithmic scale. Henceforth, we consider the mapping

$$a \rightsquigarrow \beta(a) := a^{-1}\pi(-\log a), \quad a \in (0,1). \tag{3.3}$$

In the non-parametric case, we estimate $\beta(a)$ for every $a \in (0,1)$.

3.3 The parametric case

For $k \geq 1$, we estimate

$$m_k(\pi) := \int_0^{+\infty} x^k\, \pi(x)dx = \int_0^1 \log(1/a)^k \beta(a)da$$

by the correspondence (3.3) and implicitly assumed to be well-defined. We first focus on the case $k = 1$. Pick a sufficiently smooth test function $f(\bullet) : [0,1] \to \mathbb{R}$ such that $f(1) = 0$ and let $g(a) := -af'(a)$. Plainly

$$\begin{aligned}
\mathcal{E}(g) &= \frac{1}{c(\pi)} \int_0^1 \frac{g(a)}{a} \pi(-\log a, +\infty)\, da \\
&= -\frac{1}{m_1(\pi)} \int_0^1 f'(a) \int_0^a \beta(u)du\, da = \frac{1}{m_1(\pi)} \int_0^1 f(a)\beta(a)da.
\end{aligned}$$

Formally, taking $f(\bullet) \equiv 1$ would identify $1/m_1(\pi)$ since $\beta(\bullet)$ integrates to one, but this choice is forbidden by the boundary condition $f(1) = 0$. We then consider instead the following approximation. Let $f_\gamma(\bullet) : [0,1] \to \mathbb{R}$ with $0 < \gamma < 1$ be a smooth function such that

- $f_\gamma(a) = 1$ for $a \leq 1-\gamma$ and $f_\gamma(1) = 0$.
- $\|f_\gamma\|_\infty = 1$ and $\|f'_\gamma\|_\infty \leq c\,\gamma^{-1}$, for some $c > 0$,

a choice which is obviously possible. For a parametrization $\gamma := \gamma_\varepsilon \to 0$, we set $g_{\gamma_\varepsilon}(a) := -af'_{\gamma_\varepsilon}(a)$ and define

$$\widehat{m}_{1,\varepsilon} := \frac{1}{\mathcal{E}_{\varepsilon,\sigma}\big(g_{\gamma_\varepsilon}\big)}. \tag{3.4}$$

More generally, for $k > 1$, we define successive moment estimators as follows. Set $h_{\gamma_\varepsilon}(a) := f_{\gamma_\varepsilon}(1-a)\log(1/a)^k$ and $\widetilde{g}_{\gamma_\varepsilon}(a) := -ah'_{\gamma_\varepsilon}(a)$. Let

$$\widehat{m}_{k,\varepsilon} := \frac{\mathcal{E}_{\varepsilon,\sigma}\big(\widetilde{g}_{\gamma_\varepsilon}\big)}{\mathcal{E}_{\varepsilon,\sigma}\big(g_{\gamma_\varepsilon}\big)}.$$

We can describe the performances of $\widehat{m}_{k,\varepsilon}$ under an additional decay condition of $\pi(\bullet)$ near the origin. For $\kappa > 0$, we say[6] that the probability $\pi(\bullet)$ belongs to the class $\mathcal{R}(\kappa)$ if

$$\limsup_{x\to 0} x^{-\kappa+1}\pi(x) < +\infty$$

appended with $\mathcal{R}(\infty) := \bigcap_{\kappa>0}\mathcal{R}(\kappa)$. We obtain the following upper bound:

Theorem 3.2. *Grant Assumptions A, B and C. Let $0 < \kappa_1, \kappa_2 \leq \infty$ with $\kappa_1 > \max\{1, \kappa_2\}$.*

For $1 \leq \mu < \kappa_1$, let $\widehat{m}_{k,\varepsilon}$ be specified by $\gamma_\varepsilon := \varepsilon^{\mu/(\mu+1)(2\kappa_2+1)}$. The family

$$\varepsilon^{-\mu\,\kappa_2/(\mu+1)(2\kappa_2+1)}\big(\widehat{m}_{k,\varepsilon} - m_k(\pi)\big)$$

is tight under $\mathbb{P}_1$ as soon as

$$\pi \in \Pi(\kappa_1) \cap \mathcal{R}(\kappa_2)$$

and $\sigma\varepsilon^{-3}$ remains bounded.

Some remarks: the convergence of $\widehat{m}_{k,\varepsilon}$ to $m_k(\pi)$ is of course no surprise by (3.1). However, the dependence in ε in the test function $g_\varepsilon(\bullet)$ (in particular $g_\varepsilon(\bullet)$ is unbounded as $\varepsilon \to 0$) requires a slight improvement of Theorem 3.1. This can be done thanks to Assumption C, see Proposition 5.2 in the proof Section 5.2. The requirement $\sigma\varepsilon^{-3} = \mathcal{O}(1)$ ensures that the additional term coming from the approximation of $\mathcal{E}_\varepsilon(\bullet)$ by $\mathcal{E}_{\sigma,\varepsilon}(\bullet)$ is negligible. This condition is probably not optimal, see Section 4.

Our next result shows that the exponent $\mu\kappa_2/(\mu+1)(2\kappa_2+1) \leq 1/2$ in the rate of convergence is nearly optimal, to within an arbitrarily small polynomial order.

Definition 3.3. *Let $\pi_0(\bullet)$ satisfy the assumptions of Theorem 3.2. The rate $0 < v_\varepsilon \to 0$ is a lower rate of convergence for estimating $m_k(\pi_0)$ if there exists a family $\pi_\varepsilon(\bullet)$ satisfying the assumptions of Theorem 3.2 and a constant $c > 0$ such that*

$$\liminf_{\varepsilon\to 0}\inf_{F_\varepsilon}\max_{\pi\in\{\pi_0,\pi_\varepsilon\}} \mathbb{P}\left[v_\varepsilon^{-1}|F_\varepsilon - m_k(\pi)| \geq c\right] > 0, \tag{3.5}$$

where the infimum is taken (for every ε) over all estimators constructed with $X_{\varepsilon,\sigma}$ at stage ε.

[6] In the notation, we identify the probability measure $\pi(dx)$ and its density function $x \rightsquigarrow \pi(x)$ when no confusion is possible.

Definition 3.3 expresses a kind of local min-max information bound: given $\pi_0(\bullet)$, one can find an opponent $\pi_\varepsilon(\bullet)$ such that no estimator can discriminate between $\pi_0(\bullet)$ and $\pi_\varepsilon(\bullet)$ at a rate faster than v_ε.

We further restrict our attention to binary fragmentations, see Section 4.1. In that case, the dislocation measure satisfies $\nu(s_1+s_2 \neq 1) = 0$, and, because of the conservation Assumption B, can be represented as

$$\nu(d\mathbf{s}) = \rho(ds_1)\delta_{1-s_1}(ds_2), \tag{3.6}$$

where $\rho(\bullet)$ is a probability measure on $[1/2, 1]$.

Assumption D. *(Binary case.) The probability measure $\rho(\bullet)$ associated to $\pi(\bullet)$ is absolutely continuous and its density function is bounded away from zero.*

Theorem 3.4. *Assume that the fragmentation is binary and grant Assumption D. In the same setting as in Theorem 3.2, the rate $\varepsilon^{1/2}$ is a lower rate of convergence for estimating $m_k(\pi)$.*

The restriction to the binary case is made for technical reason and is inessential. Theorem 3.4 presumably holds in a more general setting.

3.4 The non-parametric case

Under local smoothness assumptions on the parameter $\beta(\bullet)$, we estimate $\beta(a)$ for every $a \in (0,1)$. Given $s > 0$, we say that $\beta(\bullet)$ belongs to the Hölder class $\Sigma(s)$ if there exists a constant $c > 0$ such that

$$\left|\beta^{(n)}(y) - \beta^{(n)}(x)\right| \leq c|y-x|^{\{s\}},$$

where $s = n + \{s\}$, with n a non-negative integer and $\{s\} \in (0,1]$. We also need to relate $\beta(\bullet)$ to the decay of its corresponding Lévy measure $\pi(\bullet)$. Abusing again notation, we identify $\Pi(\kappa)$ with the set of $\beta(\bullet)$ such that $e^x\beta(e^{-x})dx \in \Pi(\kappa)$, thanks to the inverse of (3.3). Likewise for $\mathcal{R}(\kappa)$.

We construct an estimator of $\beta(\bullet)$ as follows: for $a \in (0,1)$ and a normalizing factor $0 < \gamma_\varepsilon \to 0$, set

$$\varphi_{\gamma_\varepsilon, a}(\bullet) := \gamma_\varepsilon^{-1}\varphi\big((\bullet - a)/\gamma_\varepsilon\big),$$

where $\varphi(\bullet)$ is a smooth function with support in $(0,1)$ that satisfies the following oscillating property: for some integer $N \geq 1$,

$$\int_0^1 \varphi(a)da = 1, \quad \int_0^1 a^k\varphi(a)da = 0, \quad k = 1, \ldots, N. \tag{3.7}$$

Our estimator then takes the form

$$\widehat{\beta}_\varepsilon(a) := \widehat{m}_{1,\varepsilon}\, \mathcal{E}_{\varepsilon,\sigma}\big(-\bullet\varphi'_{\gamma_\varepsilon,a}(\bullet)\big) \ \ a \in (0,1),$$

where $\widehat{m}_{1,\varepsilon}$ is the estimator of $m_1(\pi)$ defined in (3.4). We then have the following

Theorem 3.5. *Grant Assumptions A, B and C. Let $1 < \kappa_1, \kappa_2 \leq \infty$.*

For $1 \leq \mu < \min\{1, \kappa_1/2\}$, let $\widehat{\beta}_\varepsilon(\bullet)$ be specified by $\gamma_\varepsilon := \varepsilon^{\mu/(\mu+1)(2s+3)}$. For every $a \in (0,1)$, the family

$$\varepsilon^{-\mu s/(\mu+1)(2s+3)}\big(\widehat{\beta}_\varepsilon(a) - \beta(a)\big)$$

is tight under $\mathbb{P}_1$, as soon as

$$\beta \in \Sigma(s) \cap \Pi(\kappa_1) \cap \mathcal{R}(\kappa_2)$$

for $s < \max\{N, \kappa_2 - 1\}$ and $\sigma\varepsilon^{-3}$ remains bounded.

A proof of the (near)-optimality in the sense of the lower bound Definition 3.3 and in the spirit of Theorem 3.4 is presumably a delicate problem that lies beyond the scope of the paper. More in Section 4.3.

4 Discussion

4.1 Binary fragmentations

The case of binary fragmentations is the simplest, yet an important model of random fragmentation, where a particle splits into two blocs at each step (see *e.g.* [9], [8]). By using representation (3.6), if we assume further that $\rho(ds_1) = \rho(s_1)ds_1$ is absolutely continuous, so is $\pi(dx) = \pi(x)dx$ and we have

$$\pi(x) = e^{-2x}\big(\rho(e^{-x})1_{[0,\log 2]}(x) + \rho(1-e^{-x})1_{(\log 2,+\infty)}(x)\big), \tag{4.1}$$

for $x \in [0, +\infty)$ and

$$\beta(a) = a\big(\rho(a)1_{[1/2,1]}(a) + \rho(1-a)1_{[0,1/2)}(a)\big), \ \ a \in [0,1].$$

In particular, the regularity properties of $\beta(\bullet)$ are readily obtained from the local smoothness of $\rho(\bullet)$ and its behaviour near $1/2$. For instance, if $\rho(a + 1/2) = \mathcal{O}(a^{\kappa-1})$ near the origin, for some $\kappa > 0$, then

$$\pi \in \Pi(\kappa) \cap \mathcal{R}(\kappa).$$

4.2 Concerning Theorem 3.1

Optimal rate of convergence. First, Theorem 3.1 readily extends to error measurements of the form $\mathbb{E}\left[\left|\mathcal{E}_\varepsilon(g) - \mathcal{E}(g)\right|^p\right]$ with $1 \leq p \leq 2$. The rate becomes $\varepsilon^{-\mu p/2(\mu+1)}$ in (3.1) and $\sigma^p \varepsilon^{-p}$ in (3.2) under the less stringent condition $\mu < \kappa/2p$.

Generally speaking, we obtain in (3.1) the (normalized) rate $\varepsilon^{\mu/2(\mu+1)}$, for any $\mu < \kappa$. Intuitively, we have a number of observations that should be of order ε^{-1}, so the expected rate would rather be $\varepsilon^{1/2}$. Why cannot we obtain the rate $\varepsilon^{1/2}$ or simply $\varepsilon^{\kappa/2(\kappa+1)}$? The proof in Section 5.1 shows that we loose quite much information when applying Sgibnev's result (see Proposition 6.1 in Appendix) on the kew renewal theorem for a random walk with step distribution $\pi(\bullet)$ in the limit $\log(1/\varepsilon) \to +\infty$.

Proposition 6.1 ensures that if $\pi(\bullet)$ has exponential moments up to order κ, then we can guarantee in the renewal theorem the rate $o(\varepsilon^\mu)$ for any $\mu < \kappa$ with some uniformity in the test function, a crucial point for the subsequent statistical applications. It is presumably possible to improve this rate to $\mathcal{O}(\varepsilon^\kappa)$ by accomodating Ney's result [15]. However, a careful glance at the proof of Theorem 3.1 shows that we would then loose an extra logarithmic term when replacing $\varepsilon^{\mu/2(\mu+1)}$ by $\varepsilon^{\kappa/(2\kappa+1)}$. More generally, exhibiting exact rates of convergence in Theorem 3.1 remains a delicate issue: the kew renewal theorem is sensitive to a modification of the distribution outside a neighbourhood of $+\infty$, see *e.g.* Asmussen [2], p.196.

Uniformity in $\pi(\bullet)$. A slightly annoying fact is that convergence (3.1) is not uniform in $\pi(\bullet)$, which can become a methodological issue for the statistical applications of the subsequent Theorems 3.2 and 3.5, in particular if min-max results are sought. An inspection of the proof in Section 5.1 shows that we loose information about the uniformity in $\pi(\bullet)$ when applying Proposition 6.1 again. A glance at the proof of Sgibnev's result suggest that uniform results in $\pi(\bullet)$ could presumably be obtained over classes of $\pi(\bullet)$ defined in terms of appropriate bounds on their Stone decomposition [17].

The non-conservative case. If Assumption B is dropped, we define $p_- = \inf\left\{p > 0, \quad \int_{\mathcal{S}^\downarrow} \sum_{i=1}^\infty s_i^p \nu(d\mathbf{s}) < +\infty\right\}$ and make the so-called Malthusian hypothesis: there exists a (unique) solution $p^\star \geq p_-$ to the equation

$$\int_{\mathcal{S}^{\downarrow}} (1 - \sum_{i=1}^{\infty} s_i^{p^\star})\nu(d\mathbf{s}) = 1.$$

The empirical measure now becomes

$$\mathcal{E}_\varepsilon^{(p^\star)}(g) := \sum_{u \in \mathcal{U}_\varepsilon} \xi_u^{p^\star}\, g(\xi_u/\varepsilon).$$

The choice of the weights $\xi_u^{p^\star}$ is motivated by the fact that the process $\big(\sum_{|u|=n} \xi_u^{p^\star}, n \geq 0\big)$ is a positive martingale. We denote by $\mathcal{M}_\infty$ its terminal value. Note that under Assumption B, we always have $p^\star = 1$ and $\mathcal{M}_\infty = 1$. Bertoin and Martinez [7] prove the convergence of $\mathcal{E}_\varepsilon^{(p^\star)}(g)$ to

$$\mathcal{E}^{(p^\star)}(g) := \frac{\mathcal{M}_\infty}{c(\nu)} \int_0^1 \frac{g(a)}{a} \int_{\mathcal{S}^{\downarrow}} \sum_{i=1}^{\infty} s_i^{p^\star}\, 1_{\{s_i < a\}} \nu(d\mathbf{s}) da$$

in probability, as $\varepsilon \to 0$, with now $c(\nu) := -\int_{\mathcal{S}^{\downarrow}} \sum_{i=1}^{\infty} s_i^{p^\star} \log s_i\, \nu(d\mathbf{s})$. In this setting, Theorem 3.1 becomes

Corollary 4.1. *Grant Assumptions A, C and the Malthusian hypothesis. Let $0 < \kappa \leq \infty$ and $m > 0$. For every $0 < \mu < \kappa$, we have*

$$\sup_{g \in \mathcal{C}(m)} \mathbb{E}\left[\big(\mathcal{E}_\varepsilon^{(p^\star)}(g) - \mathcal{E}^{(p^\star)}(g)\big)^2\right] = o\big(\varepsilon^{\mu/(\mu+p^\star)}\big).$$

4.3 Concerning Theorems 3.2 and 3.5

The parametric case. We obtain the rate

$$\big(\varepsilon^{\mu/(\mu+1)}\big)^{\kappa_2/(2\kappa_2+2)}, \quad \text{for all } \mu < \kappa_1$$

which can be made arbitrary close to the lower bound $\varepsilon^{1/2}$ by assuming κ_1 and κ_2 large enough. The factor $\mu/(\mu+1)$ comes from Theorem 3.1 whereas the factor $\kappa_2/(2\kappa_2+1)$ arises when using the technical assumption $\pi \in \mathcal{R}(\kappa_2)$. We do not know how to improve it.

Also, the condition $\sigma\varepsilon^{-3} = \mathcal{O}(1)$ is fairly restrictive, and can be readily improved by assuming that κ_2 is large. Indeed, if $\kappa_2 \geq 1$, which only amounts to require that $\pi(\bullet)$ is bounded near the origin, a glance at the error term (5.10) in the proof Section 5.2 shows that the condition drops to $\sigma\varepsilon^{-2} = \mathcal{O}(1)$. In the limit $\kappa_2 \to \infty$, we obain $\sigma\varepsilon^{-3/2} = \mathcal{O}(1)$.

The non-parametric case. The situation is a bit different than in the parametric case: we obtain now the rate

$$\big(\varepsilon^{\mu/(\mu+1)}\big)^{s/(2s+3)}, \quad \text{for all } \mu < \kappa_1$$

for the estimation of $\beta(a)$ for any $a \in (0,1)$. In the limit $\kappa_1 \to +\infty$ it becomes $\varepsilon^{s/(2s+3)}$, which can be related to more classical models: consider the apparently different problem of recovering a function $\beta(\bullet)$ in the integral white noise model

$$dY_a = K\beta(a)da + \varepsilon^{1/2}dW_a, \quad a \in [0,1], \tag{4.2}$$

from the observation of $(Y_a, a \in [0,1])$. Here, $(W_a, a \in [0,1])$ is a standard Brownian motion and $K\beta(a) := \int_0^a \beta(u)du$ is the integration operator. Model (4.2) serves as a toy representation for the problem of recovering a signal in white noise at level $\varepsilon^{1/2}$, when the observation is obtained from the action of a smoothing linear operator with unbounded inverse (here K). The difficulty of the problem is quantified by the degree of ill-posedness of the operator (equal to ν for ν-fold integration; here $\nu = 1$). The well-known optimal rate (see *e.g.* [12]) of pointwise recovery for a function $\beta \in \Sigma(s)$ is

$$\varepsilon^{s/(2s+2\nu+1)} = \varepsilon^{s/(2s+3)}.$$

The factor 2ν is a futher penalty in the rate of recovery quantifying the smoothing action of K. The same phenomenon seems to occur in the setting of fragmentation chain. Put $\sigma := 0$ here for simplicity. For a test function $g(\bullet)$, we can form the observation

$$\mathcal{E}_\varepsilon(g) \approx \mathcal{E}\big(g\big) = \frac{1}{m_1(\pi)}\int_0^1 g(u)K\beta(u)du$$

up to an error of (near)-order $\varepsilon^{1/2}$. If we discard the pre-factor $m_1(\pi)$ (which can be estimated at a fast rate when κ_2 is large) we obtain the same kind of statistics in Model (4.2) by considering

$$\int_0^1 g(a)dY_a = \int_0^1 g(u)K\beta(u)du + \varepsilon^{1/2}\mathcal{N}(g),$$

where $\mathcal{N}(g)$ is centred Gaussian with variance $\int_0^1 g(a)^2da$. Note that the order of the variance in the noise term $\mathcal{N}(g)$ is consistent with the improvement obtained in Proposition 5.2 in the proof Section 5.2: if $g \in \widetilde{\mathcal{C}}_{b_\varepsilon}(m)$, we have $\int_0^1 g(a)^2da \lesssim b_\varepsilon$.

This suggests the (near)-optimality of the result in the sense of Definition 3.3 but a complete proof lies beyond the scope of the paper.

4.4 Other statistical issues

Observation of the whole path of X. Suppose we observe continuously in time the sample path of X up to some fixed time $T > 0$. Asymptotics are taken as $T \to \infty$. Equivalently, we observe all triples (ξ_u, a_u, ζ_u) for every u in the random set

$$\mathcal{U}(T) := \{u \in \mathcal{U},\ \ a_u \leq T\}$$

with the restriction that ζ_u is set to $T - a_u$ when $a_u + \zeta_u > T$. We denote by $\mathcal{U}(T_-)$ the subset of $\mathcal{U}(T)$ such that $a_u + \zeta_u \leq T$. In this setting, statistical inference about the the self-similar parameter $\alpha \geq 0$ and the dislocation measure $\nu(d\mathbf{s})$ is relatively straightforward:

Estimation of α: conditional on $\big(\xi_u, u \in \mathcal{U}(T_-)\big)$ the sequence of random variables $\big(\zeta_u, u \in \mathcal{U}(T)\big)$ are independent and follow exponential distributions with parameters ξ_u^α. Conditional on $\mathrm{Card}\,\mathcal{U}(T) = n$ and since the law of the $(\xi_u, u \in \mathcal{U})$ does not depend on α we are in the setting of estimating a one-dimensional parameter from n independent random variables with explicit likelihood ratio

$$\alpha \rightsquigarrow \prod_{i=1}^{n} \xi_{u_i}^\alpha \exp\big(\xi_{u_i}^\alpha\, \zeta_{u_i}\big), \tag{4.3}$$

where the u_i range through $\mathcal{U}(T)$. The main difficulty remains that the law of $\mathrm{Card}\,\mathcal{U}(T)$ usually depends on α.

Estimation of $\nu(\bullet)$: for $u- \in \mathcal{U}(T_-)$, when the fragment of size ξ_u splits, conditional on $\xi_u = x$, it gives rize to the observation of the rescaled size of its offsprings $(x^{-1}\,\xi_{ui}, i \in \mathbb{N})$ which is a realization of the law $\nu(d\mathbf{s})$. As a consequence, conditional on $\mathrm{Card}\,\mathcal{U}(T) = n$, we observe a sequence of n independent and identically distributed random variables with law $\nu(d\mathbf{s})$. We are back to the classical problem of estimating a probability distribution from an n-sample.

More about estimating α. We cannot estimate the index of self-similarity α from the data X_ε. However, if we add the possibility to "tag" a point at random on the initial fragment[7] and if we can observe the random time T_ε when the tagged fragment becomes smaller then ε, then identifying α from the sole observation of T_ε is possible.

[7]in physical terms, we must be able to identify the mass or length of the fragment.

In the case $\alpha > 0$, if $\chi(t)$ denotes the size of the tagged fragment at time t, then

$$T_\varepsilon = \big\{t > 0, \;\; \chi(t) \leq \varepsilon\big\}.$$

Applying Proposition 3 of [6], the distribution of $\varepsilon^\alpha T_\varepsilon$ under $\mathbb{P}_1$ is tight as $\varepsilon \to 0$. Therefore, the rate $-1/\log\varepsilon$ is achievable for estimating α and it is attained by the estimator $\log(T_\varepsilon)/\log\varepsilon$. More precise results about limit laws can be obtained from [6].

5 Proofs

We will repeatedly use the convenient notation $a_\varepsilon \lesssim b_\varepsilon$ if $0 < a_\varepsilon \leq c\, b_\varepsilon$ for some constant $c > 0$ which may depend on $\pi(\bullet)$ and m only, any other dependence on other ancillary quantities being obvious from the context. A function $g \in \mathcal{C}(m)$ is tacitly defined on the whole real line by setting $g(a) = 0$ for $a \notin [0, 1]$.

5.1 Proof of Theorem 3.1

Step 1: A preliminary decomposition. We first use the fact that for $\eta > \varepsilon$, during the fragmentation process, the unobserved state X_η necessarily anticipates the state X_ε. The choice $\eta = \eta(\varepsilon)$ will follow later. This yields the following representation:

$$\mathcal{E}_\varepsilon(g) = \sum_{v \in \mathcal{U}_\eta} \xi_v \sum_{w \in \mathcal{U}} 1_{\left\{\xi_v \tilde{\xi}^{(v)}_{w-} \geq \varepsilon,\, \xi_v \tilde{\xi}^{(v)}_w < \varepsilon\right\}} \tilde{\xi}^{(v)}_w g\big(\xi_v \tilde{\xi}^{(v)}_w / \varepsilon\big),$$

where, for each label $v \in \mathcal{U}_\eta$ and conditional on X_η, a new independent fragmentation chain $(\tilde{\xi}^{(v)}_w, w \in \mathcal{U})$ is started thanks to the branching property, see Section 2.1. Define now

$$\lambda_\eta(v) := 1_{\{\xi_{v-} \geq \eta, \xi_v < \eta\}} \xi_v$$

and

$$Y_\varepsilon(v, g) := \sum_{w \in \mathcal{U}} 1_{\left\{\xi_v \tilde{\xi}^{(v)}_{w-} \geq \varepsilon,\, \xi_v \tilde{\xi}^{(v)}_w < \varepsilon\right\}} \tilde{\xi}^{(v)}_w g\big(\xi_v \tilde{\xi}^{(v)}_w / \varepsilon\big).$$

We obtain the decomposition of $\mathcal{E}_\varepsilon(g) - \mathcal{E}(g)$ as a sum of a centred and a bias term:

$$\mathcal{E}_\varepsilon(g) - \mathcal{E}(g) = M_{\varepsilon,\eta}(g) + B_{\varepsilon,\eta},$$

with
$$M_{\varepsilon,\eta}(g) := \sum_{v\in\mathcal{U}} \lambda_\eta(v)\big(Y_\varepsilon(v,g) - \mathbb{E}\left[Y_\varepsilon(v,g)\,|\,\lambda_\eta(v)\right]\big)$$
and
$$B_{\varepsilon,\eta}(g) := \sum_{v\in\mathcal{U}} \lambda_\eta(v)\big(\,\mathbb{E}\left[Y_\varepsilon(v,g)\,|\,\lambda_\eta(v)\right] - \mathcal{E}(g)\big),$$
where we used (2.1) in order to incoporate the limit term $\mathcal{E}(g)$ within the sum in v.

Step 2: The term $M_{\varepsilon,\eta}(g)$. Conditional on the sigma-field generated by $\big(1_{\{\xi_{u-}\geq\eta\}}\xi_u, u-\in\mathcal{U}\big)$, the variables $(Y_\varepsilon(v,g), v\in\mathcal{U})$ are independent. Therefore
$$\mathbb{E}\left[M_{\varepsilon,\eta}(g)^2\right] \leq \sum_{v\in\mathcal{U}} \mathbb{E}\left[\lambda_\eta(v)^2\,\mathbb{E}[Y_\varepsilon(v,g)^2\,|\,\lambda_\eta(v)]\right], \tag{5.1}$$
thus we first need to control the conditional variance of $Y_\varepsilon(v,g)^2$ given $\lambda_\eta(v) = u$, for $0\leq u\leq \eta$, since $\mathbb{P}$-almost surely, $\lambda_\eta(v)\leq\eta$. Moreover, we have $Y_\varepsilon(v,g) = 0$ on $\{\lambda_\eta(v) < \varepsilon\}$, hence we may assume $\varepsilon\leq u\leq\eta$.

To this end, we will use the following representation property:

Lemma 5.1. *Let $f(\bullet) : [0,+\infty)\to[0,+\infty)$. Then*
$$\mathbb{E}\Big[\sum_{v\in\mathcal{U}_\eta}\xi_v\, f(\xi_v)\Big] = \mathbb{E}^\star\left[f\big(\chi(T_\eta)\big)\right], \tag{5.2}$$
where $\chi(t) = \exp\big(-\zeta(t)\big)$ and $\big(\zeta(t), t\geq 0\big)$ is a subordinator with Lévy measure $\pi(\bullet)$ defined on an appropriate probability space $(\Omega^\star,\mathbb{P}^\star)$, and
$$T_\eta := \inf\big\{t\geq 0,\ \zeta(t) > -\log\eta\big\}.$$

The proof is given in Appendix 6.1. In order to remain self-contained, we elaborate in particular on the construction of the randomly tagged fragment following the presentation of Bertoin [5].

We now plan to bound the right-hand side of (5.1) thanks to Lemma 5.1. For $0<\varepsilon\leq u\leq\eta$, we have
$$\begin{aligned}\mathbb{E}\left[Y_\varepsilon(v,g)^2\,|\,\lambda_\eta(v)=u\right] &= \mathbb{E}\Big[\Big(\sum_{w\,\in\mathcal{U}_{\varepsilon/u}}\widetilde{\xi}_w^{(v)}g\big(\varepsilon u^{-1}\widetilde{\xi}_w^{(v)}\big)\Big)^2\,|\,\lambda_\eta(v)=u\Big]\\ &\leq \mathbb{E}\,\big[\sum_{w\in\mathcal{U}_{\varepsilon/u}}\widetilde{\xi}_w^{(v)}g\big(\varepsilon u^{-1}\widetilde{\xi}_w^{(v)}\big)^2\,|\,\lambda_\eta(v)=u\big]\end{aligned}$$

where we used Jensen's inequality combined with (2.1). Applying Lemma 5.1 we derive

$$\mathbb{E}\left[Y_\varepsilon(v,g)^2 \,|\, \lambda_\eta(v)=u\right] \le \mathbb{E}^\star\left[g\big(u\varepsilon^{-1}e^{-\zeta(T_{\varepsilon/u})}\big)^2\right]. \tag{5.3}$$

Let $U(\bullet)$ denote the renewal function associated with the subordinator $(\zeta(t), t \ge 0)$. By Proposition 2, Ch. III in [3], the right-hand side of (5.3) is equal to

$$\begin{aligned}
&\int_{\left[0,-\log(\varepsilon/u)\right)} dU(s) \int_{(-\log(\varepsilon/u)-s,+\infty)} g\big(u\,\varepsilon^{-1}e^{-x-s}\big)^2 \pi(dx),\\
=&\int_{\left[0,-\log(\varepsilon/u)\right)} dU(s) \int_{\mathcal{S}^\downarrow} \sum_{i=1}^\infty s_i\, 1_{\{s_i<\varepsilon u^{-1}e^s\}}\, g\big(s_i u\,\varepsilon^{-1}e^{-s}\big)^2 \nu(d\mathbf{s})\\
\lesssim&\frac{1}{c(\pi)}\|g\|_\infty^2 \log(u/\varepsilon),
\end{aligned}$$

where we successively used the representation (2.4) and the upper bound $U(s) \lesssim s/c(\pi)$, see for instance Proposition 1, Ch. I in [3]. Therefore, for $\varepsilon \le u \le \eta$,

$$\mathbb{E}\left[Y_\varepsilon(v,g)^2 \,|\, \lambda_\eta(v)=u\right] \lesssim \frac{1}{c(\pi)}\|g\|_\infty^2 \log(\eta/\varepsilon).$$

Going back to (5.1), since $\lambda_\eta(v)^2 \le \eta\,\lambda_\eta(v)$ and using (2.1) again, we readily derive

$$\mathbb{E}\left[M_{\varepsilon,\eta}(g)^2\right] \lesssim \frac{1}{c(\pi)}\|g\|_\infty^2 \eta \log(\eta/\varepsilon) \lesssim \eta\log(\eta/\varepsilon). \tag{5.4}$$

Step 3: The bias term $B_{\varepsilon,\eta}(g)$. Note first that

$$\mathbb{E}\left[Y_\varepsilon(v,g) \,|\, \lambda_\eta(v)\right] = \xi_v^{-1}\, \mathbb{E}_{\xi_v}\left[\mathcal{E}_\varepsilon(g)\right],$$

$\mathbb{P}$-almost surely, henceforth

$$B_{\varepsilon,\eta}(g) = \sum_{v\in\mathcal{U}} \lambda_\eta(v)\Big(\xi_v^{-1}\, \mathbb{E}_{\xi_v}\left[\mathcal{E}_\varepsilon(g)\right] - \mathcal{E}(g)\Big). \tag{5.5}$$

Conditioning on the mark of the parent $u- = \omega$ of u and applying the branching property, we get that $\mathbb{E}_{\xi_v}\left[\mathcal{E}_\varepsilon(g)\right]$ can be written as

$$\mathbb{E}_{\xi_v}\Big[\sum_{\omega\in\mathcal{U}} 1_{\left\{\widehat{\xi}_\omega \ge \varepsilon\right\}} \widehat{\xi}_\omega \int_{\mathcal{S}^\downarrow} \sum_{i=1}^\infty 1_{\left\{\widehat{\xi}_\omega\, s_i<\varepsilon\right\}} s_i\, g\big(\widehat{\xi}_\omega\, s_i\,\varepsilon^{-1}\big)\Big].$$

where the $(\widehat{\xi}_w, w \in \mathcal{U})$ are the sizes of the marked fragments of a fragmentation chain with same dislocation measure $\nu(\bullet)$, independent of $(\xi_v, v \in \mathcal{U})$. Set

$$H_g(a) := \int_{\mathcal{S}^\downarrow} \sum_{i=1}^{\infty} 1_{\{s_i < e^{-a}\}} s_i g\big(s_i e^a\big) \nu(d\mathbf{s}), \quad a \geq 0.$$

It follows that $\mathbb{E}_{\xi_v}\big[\mathcal{E}_\varepsilon(g)\big]$ is equal to

$$\begin{aligned}
&\mathbb{E}_{\xi_v}\Big[\sum_{n=0}^{\infty} \sum_{|\omega|=n} 1_{\{\log \widehat{\xi}_\omega \geq \log \varepsilon\}} \widehat{\xi}_\omega H_g(\log \widehat{\xi}_\omega - \log \varepsilon)\Big] \\
&= \xi_v\, \mathbb{E}\Big[\sum_{n=0}^{\infty} \sum_{|\omega|=n} 1_{\{\log \widehat{\xi}_\omega \geq \log(\varepsilon/\rho)\}} H_g\big(\log \widehat{\xi}_\omega - \log(\varepsilon/\rho)\big)\Big]_{\rho=\xi_v}
\end{aligned}$$

by self-similarity, with the notation $|\omega| = n$ if $\omega = (\omega_1, \ldots, \omega_n) \in \mathcal{U}$. Using Lemma 1.4 in [5], we finally obtain

$$\mathbb{E}_{\xi_v}\big[\mathcal{E}_\varepsilon(g)\big] = \xi_v \sum_{n=0}^{\infty} \mathbb{E}\big[1_{\{S_n \leq \log(\rho/\varepsilon)\}} H_g\big(\log(\rho/\varepsilon) - S_n\big)\big]_{\rho=\xi_v},$$

where S_n is a random walk with step distribution $\pi(dx)$. We plan to apply a version of the renewal theorem with explicit rate of convergence as given in Sgibnev [16], see Proposition 6.1 in Appendix 6.2, with rate function $\varphi(a) := \exp(\mu a)$ for some arbitrary $\mu < \kappa/2$ and dominating function $r(a) := e^{-\kappa|a|}$. Indeed, for $a < 0$:

$$H_g(-a) = 1_{\{a \leq 0\}} \int_{(-a,+\infty)} g\big(e^{-x-a}\big)\pi(dx),$$

by (2.4). Since $g(\bullet)$ has support in $[0,1]$ and $\pi \in \Pi(\kappa)$,

$$\big|H_g(-a)\big| \leq \int_{(-a,+\infty)} \big|g\big(e^{-x-a}\big)\big|\pi(dx) \lesssim e^{\kappa a},$$

Therefore $|1_{\{a\leq 0\}} H_g(-a)| \lesssim r(a)$ for all $a \in \mathbb{R}$.

Since $\kappa > 2\mu$, Assumption F of Proposition 6.1 is readily checked. Let $A > 0$ (depending on κ, m and $\pi(\bullet)$ only) such that, if $\log(\xi_v/\varepsilon) \geq A$, then

$$\Big|\xi_v^{-1}\, \mathbb{E}_{\xi_v}\big[\mathcal{E}_\varepsilon(g)\big] - \frac{1}{\mathbb{E}^\star\big[S_1\big]} \int_0^{+\infty} H_g(a)da\Big| \leq \Big(\frac{\varepsilon}{\xi_v}\Big)^{\mu}. \qquad (5.6)$$

We next note that

$$\frac{1}{\mathbb{E}^\star\big[S_1\big]}\int_0^{+\infty} H_g(a)da = \mathcal{E}(g).$$

Introducing the family of events $\big\{\log(\xi_v/\varepsilon)\geq A\big\}$ in the sum (5.5), we obtain the following decomposition:

$$B_{\varepsilon,\eta}(g)^2 \lesssim I + II,$$

with

$$I := \sum_{v\in\mathcal{U}_\eta} \xi_v 1_{\big\{\log(\xi_v/\varepsilon)>A\big\}}\big(\xi_v^{-1}\,\mathbb{E}_{\xi_v}\big[\mathcal{E}_\varepsilon(g)\big]-\mathcal{E}(g)\big)^2,$$

and

$$II := \sum_{v\in\mathcal{U}_\eta} \xi_v 1_{\big\{\log(\xi_v/\varepsilon)\leq A\big\}}\big(\xi_v^{-1}\,\mathbb{E}_{\xi_v}\big[\mathcal{E}_\varepsilon(g)\big]-\mathcal{E}(g)\big)^2.$$

By (5.6), we have

$$I \leq \varepsilon^{2\mu}\sum_{v\in\mathcal{U}_\eta} 1_{\big\{-\log\xi_v<-A+\log(1/\varepsilon)\big\}}\xi_v\exp\big(2\mu(-\log\xi_v)\big).$$

Integrating w.r.t. $\mathbb{P}$ and applying Lemma 5.1 and in the same way as in Step 2, we have

$$\begin{aligned}
\mathbb{E}\big[I\big] &\leq \varepsilon^{2\mu}\,\mathbb{E}^\star\big[e^{2\mu\zeta(T_\eta)}\big] \\
&= \varepsilon^{2\mu}\int_{[0,-\log\eta)} dU(s)\int_{(-\log\eta-s,+\infty)} e^{2\mu(s+x)}\pi(dx) \\
&\leq \varepsilon^{2\mu}\int_{[0,-\log\eta)} e^{2\mu s}dU(s) \lesssim \big(\varepsilon\eta^{-1}\big)^{2\mu}\log(1/\eta)
\end{aligned}$$

for small enough ε and where we used $\pi\in\Pi(\kappa)$ with $2\mu<\kappa$. For the term II, we first notice that by (2.1) and self-similarity,

$$\mathbb{E}_{\xi_v}\big[\sum_{u\in\mathcal{U}_\varepsilon}\widehat{\xi}_u\big] = \xi_v,\ \ \mathbb{P}_{\xi_v}-\text{almost surely},$$

hence

$$\big(\xi_v^{-1}\,\mathbb{E}_{\xi_v}\big[\mathcal{E}_\varepsilon(g)\big]-\mathcal{E}(g)\big)^2 \leq 4\|g\|_\infty^2,\ \ \mathbb{P}_{\xi_v}-\text{almost surely}.$$

In the same way as for the term I, we derive

$$\begin{aligned}
\mathbb{E}\left[II\right] &\lesssim \mathbb{E}\Big[\sum_{v\in\mathcal{U}_\eta} \xi_v 1_{\{-\log\xi_v\geq -A+\log(1/\varepsilon)\}}\Big] \\
&= \mathbb{P}^\star\left[\zeta(T_\eta)\geq -A+\log(1/\varepsilon)\right] \\
&\leq \int_{[0,-\log\eta)} dU(s)\int_{(-A+\log(1/\varepsilon)-s,+\infty)}\pi(dx) \\
&\lesssim \varepsilon^\mu \log(1/\eta)
\end{aligned}$$

for small enough ε. Putting all the estimates together, we conclude

$$\mathbb{E}\left[B_{\varepsilon,\eta}(g)^2\right] \lesssim \left(\varepsilon^\mu + (\varepsilon\eta^{-1})^{2\mu}\right)\log(1/\eta). \tag{5.7}$$

Step 4: Proof of (3.1)**.** Putting the estimates (5.4) and (5.7), we have

$$\begin{aligned}
\mathbb{E}\left[\left(\mathcal{E}_\varepsilon(g)-\mathcal{E}(g)\right)^2\right] &\lesssim \mathbb{E}\left[M_{\varepsilon,\eta}(g)^2\right]+\mathbb{E}\left[B_{\varepsilon,\eta}(g)^2\right] \\
&\lesssim \eta\log(\eta/\varepsilon)+\left(\varepsilon\eta^{-1}\right)^{2\mu}\log(1/\eta).
\end{aligned}$$

The choice $\eta(\varepsilon) := \varepsilon^{2\mu/(2\mu+1)}$ yields the rate $\varepsilon^{2\mu/(2\mu+1)}\log(1/\varepsilon)$. Since $\mu<\kappa/2$ is arbitrary, the conclusion follows.

Step 5: Proof of (3.2)**.** We plan to use the following decomposition:

$$\mathcal{E}_{\varepsilon,\sigma}(g)-\mathcal{E}_\varepsilon(g) = I + II,$$

with

$$I := \sum_{u\in\mathcal{U}} \left(1_{\{\xi^{(\sigma)}_{u-}\geq\varepsilon,\,\xi^{(\sigma)}_u<\varepsilon\}} - 1_{\{\xi_{u-}\geq\varepsilon,\,\xi_u<\varepsilon\}}\right)\widetilde{\xi}^{(\sigma)}_u\, g\big(\xi^{(\sigma)}_u/\varepsilon\big),$$

and

$$II := \sum_{u\in\mathcal{U}_\varepsilon}\left(\widetilde{\xi}^{(\sigma)}_u\, g\big(\xi^{(\sigma)}_u/\varepsilon\big) - \xi_u\, g\big(\xi_u/\varepsilon\big)\right),$$

where we have set $\widetilde{\xi}^{(\sigma)}_u := \xi^{(\sigma)}_u 1_{\{\xi^{(\sigma)}_u\geq t_\varepsilon\}}$. Clearly,

$$\begin{aligned}
\left|1_{\{\xi^{(\sigma)}_{u-}\geq\varepsilon,\,\xi^{(\sigma)}_u<\varepsilon\}} - 1_{\{\xi_{u-}\geq\varepsilon,\,\xi_u<\varepsilon\}}\right| \leq& 1_{\{\xi^{(\sigma)}_{u-}\geq\varepsilon,\,\xi_{u-}<\varepsilon\}} + 1_{\{\xi^{(\sigma)}_u<\varepsilon,\,\xi_u\geq\varepsilon\}} \\
&+1_{\{\xi_{u-}\geq\varepsilon,\,\xi^{(\sigma)}_{u-}<\varepsilon\}} + 1_{\{\xi_u<\varepsilon,\,\xi^{(\sigma)}_u\geq\varepsilon\}}.
\end{aligned}$$

Let $\delta>\sigma/\varepsilon$ and $\omega=u$ or $u-$. Since $|U_\omega|\leq 1$ for every ω, we readily check that

$$\left\{\xi^{(\sigma)}_\omega\geq\varepsilon,\,\xi_\omega<\varepsilon\right\}\subset\left\{(1-\delta)\varepsilon\leq\xi_\omega<\varepsilon\right\}$$

and

$$\big\{\xi_\omega \geq \varepsilon,\, \xi_\omega^{(\sigma)} < \varepsilon\big\} \subset \big\{\varepsilon \leq \xi_\omega < (1+\delta)\varepsilon\big\}.$$

It follows that $|I| \leq III + IV$, with

$$III := \sum_{u\in\mathcal{U}} 1_{\{(1-\delta)\varepsilon\leq\xi_{u-}\leq\varepsilon(1+\delta)\}} \big|\widetilde{\xi}_u^{(\sigma)}\, g\big(\xi_u^{(\sigma)}/\varepsilon\big)\big|$$

and

$$IV := \sum_{u\in\mathcal{U}} 1_{\{(1-\delta)\varepsilon\leq\xi_u\leq(1+\delta)\varepsilon\}} \big|\widetilde{\xi}_u^{(\sigma)}\, g\big(\xi_u^{(\sigma)}/\varepsilon\big)\big|.$$

By picking δ is small enough, we may (and will) assume that $\widetilde{\xi}_u^{(\sigma)} \lesssim \xi_u$. By (2.1), conditioning on the mark of the parent $u- = \omega$ of u and applying the branching property, $\mathbb{E}\left[III^2\right]$ is less than

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{v\in\mathcal{U}} 1_{\{(1-\delta)\varepsilon\leq\xi_v\leq\varepsilon(1+\delta)\}}\xi_v \int_{\mathcal{S}^\downarrow}\sum_{i=1}^{\infty} s_i\, g\big(\varepsilon^{-1}(\xi_v s_i + \sigma U_v)\big)^2\nu(d\mathbf{s})\Big]\\
=&\,\mathbb{E}\Big[\sum_{\omega\in\mathcal{U}} 1_{\{(1-\delta)\varepsilon\leq\xi_\omega\leq\varepsilon(1+\delta)\}}\xi_\omega G_1(\xi_\omega)\Big],
\end{aligned}$$

with

$$G_1(a) := \int_{\mathcal{S}^\downarrow}\sum_{i=1}^{\infty} s_i\,\mathbb{E}\left[g\big(\varepsilon^{-1}(as_i+\sigma U)\big)^2\right]\nu(d\mathbf{s})$$

and U distributed as the U_ω. Likewise,

$$\mathbb{E}\left[IV^2\right] \lesssim \mathbb{E}\Big[\sum_{u\in\mathcal{U}} 1_{\{(1-\delta)\varepsilon\leq\xi_u\leq\varepsilon(1+\delta)\}}\xi_u G_2(\xi_u)\Big],$$

with

$$G_2(a) := \mathbb{E}\left[g\big(\varepsilon^{-1}(a+\sigma U)\big)^2\right].$$

For $i = 1,2$, the crude bound $|G_i(a)| \leq \|g\|_\infty^2$ and the genealogical representation argument used in Step 3 enables to bound either $\mathbb{E}[III^2]$ or $\mathbb{E}[IV^2]$ by

$$\|g\|_\infty^2 \sum_{n=0}^{\infty}\mathbb{P}^\star\big[-\log(1+\delta) \leq S_n - \log(1/\varepsilon) \leq -\log(1-\delta)\big]$$

where S_n is a random walk with step distribution $\pi(\bullet)$. We proceed as in Step 3 and apply Proposition 6.1. The above term converges to

$$m_1(\pi)^{-1}\log\left(\frac{1+\delta}{1-\delta}\right) \lesssim \delta$$

uniformly in δ, as soon as δ is bounded, at rate ε^{μ} for any $0 < \mu < \kappa$, and thus is of order $\delta + \varepsilon^{\mu}$.

We now turn to the term II. We have $II := V + VI + VII$, with

$$\begin{aligned} V &:= \sum_{u \in \mathcal{U}_\varepsilon} \xi_u \Big(g\big(\xi_u^{(\sigma)}/\varepsilon\big) - g\big(\xi_u/\varepsilon\big) \Big), \\ VI &:= \sigma \sum_{u \in \mathcal{U}_\varepsilon} U_u 1_{\{\xi_u^{(\sigma)} \geq t_\varepsilon\}} g\big(\xi_u^{(\sigma)}/\varepsilon\big), \\ VII &:= - \sum_{u \in \mathcal{U}_\varepsilon} \xi_u 1_{\{\xi_u^{(\sigma)} < t_\varepsilon\}} g\big(\xi_u^{(\sigma)}/\varepsilon\big). \end{aligned}$$

From $g \in \mathcal{C}'(m)$, (2.1) and Jensen's inequality we derive

$$\mathbb{E}[V^2] \leq \|g'\|_\infty^2 \sigma^2 \varepsilon^{-2}.$$

From $|U_u| \leq 1$ and the inclusion $\{\xi_u^{(\sigma)} \geq t_\varepsilon\} \subset \{\xi_u \geq t_\varepsilon - \sigma\}$ we derive

$$\mathbb{E}\left[VI^2\right] \leq \|g\|_\infty^2 \frac{\sigma^2}{(t_\varepsilon - \sigma)^2} \mathbb{E}\Big[\big(\sum_{u \in \mathcal{U}_\varepsilon} \xi_u\big)^2\Big] \lesssim \frac{\sigma^2}{\varepsilon^2},$$

where we used that $t_\varepsilon = \gamma_0 \varepsilon$ with $0 < \gamma_0 < 1$. Likewise, the inclusion $\{\xi_u^{(\sigma)} < t_\varepsilon\} \subset \{\xi_u \leq t_\varepsilon + \sigma\}$ and Lemma 5.1 yield

$$\mathbb{E}\left[VII^2\right] \leq \|g\|_\infty^2 \, \mathbb{P}^\star\left[-\log \chi(T_\varepsilon) > -\log(t_\varepsilon + \sigma)\right] \lesssim \varepsilon^{\mu} \log(1/\varepsilon)$$

for any $\mu < \kappa$, in the same line as for the bound of the right-hand side of (5.3) in Step 2.

Putting all the estimates together with, for instance, $\delta := \sigma/2\varepsilon$ we obtain (3.1). The proof of Theorem 3.1 is complete.

5.2 Proof of Theorem 3.2

Preliminaries. Let $0 < b_\varepsilon \to 0$ as $\varepsilon \to 0$. For $m > 0$, define the class

$$\widetilde{\mathcal{C}}_{b_\varepsilon}(m) := \big\{ g \in \mathcal{C}(m), \ \big|\mathrm{supp}\big(g(\bullet)\big)\big| \leq m b_\varepsilon \big\}.$$

We have the following extension of Theorem 3.1.

Proposition 5.2. *Grant Assumptions A, B and C. In the same setting as Theorem 3.1, if in addition, we assume $\kappa > 1$, then, for every $\mu < \kappa$*

$$\sup_{g \in \widetilde{\mathcal{C}}_{b_\varepsilon}(m)} \mathbb{E}\left[\big(\mathcal{E}_\varepsilon(g) - \mathcal{E}(g)\big)^2\right] = o\big(\varepsilon^{\mu/(\mu+1)} b_\varepsilon\big).$$

Proof. We revisit carefully Steps 2 to 4 of the proof of Theorem 3.1 under the additional Assumption C, and we write $g(\bullet) = g_\varepsilon(\bullet)$ to emphasize that $g(\bullet)$ may now depend on the asymptotics.

In Step 2, the right-hand side of (5.3) is now bounded by the following chain of arguments:

$$\begin{aligned}
&\int_0^{-\log(\varepsilon/u)} dU(s) \int_{-\log(\varepsilon/u)-s}^{+\infty} g_\varepsilon\big(u\,\varepsilon^{-1}e^{-x-s}\big)^2 \pi(x)dx \\
=&\int_0^{-\log(\varepsilon/u)} dU(s) \int_0^{\varepsilon u^{-1}e^{s}} g_\varepsilon\big(xu\,\varepsilon^{-1}e^{-s}\big)^2 \beta(x)\,dx \\
\leq&\|\beta\|_\infty\, u^{-1}\varepsilon \int_{[0,-\log(\varepsilon/u))} e^s dU(s) \int_0^1 g_\varepsilon\big(x\big)^2 dx \lesssim b_\varepsilon \log(u/\varepsilon)
\end{aligned}$$

where we used that $|\operatorname{supp}(g_\varepsilon)| \lesssim b_\varepsilon$ and $U(s) \lesssim s/c(\pi)$ again. Note that $\|\beta\|_\infty \lesssim 1$ since $\kappa_1 > 1$ and $\kappa_2 > 1$. Therefore

$$\mathbb{E}\left[Y_\varepsilon(v,g)^2 \,|\, \lambda_\eta(v) = u\right] \lesssim b_\varepsilon,$$

Hence

$$\mathbb{E}\left[M_{\varepsilon,\eta}(g)^2\right] \lesssim b_\varepsilon \eta.$$

In Step 3, we replace $g(\bullet)$ by $g_\varepsilon(\bullet)$ in $\mathcal{E}_\varepsilon(g)$ and $\mathcal{E}(g)$. We have, for any $0 < \mu < \kappa$,

$$\begin{aligned}
\big|\mathcal{E}(g_\varepsilon)\big| &\leq \frac{1}{c(\pi)} \int_0^1 \frac{|g_\varepsilon(a)|}{a} \int_{\log(1/a)}^{+\infty} \pi(x)dx\,da \\
&\lesssim \int_0^1 |g_\varepsilon(a)| a^{\mu-1} da \lesssim b_\varepsilon
\end{aligned}$$

for $\mu \geq 1$ and since $\pi \in \Pi(\kappa)$ with $\kappa > 1$. By Cauchy-Schwarz, for $a < 0$,

$$\begin{aligned}
\big|H_{g_\varepsilon}(-a)\big| \leq& \Big(\int_{-a}^{+\infty} g_\varepsilon\big(e^{-x-a}\big)^2 \pi(x)dx\Big)^{1/2} \Big(\int_{-a}^{+\infty} \pi(x)dx\Big)^{1/2} \\
\lesssim& e^{a/2} \Big(\int_0^1 g_\varepsilon\big(y\big)^2 \beta(ye^a)dy\Big)^{1/2} e^{\kappa a/2} \\
\lesssim& b_\varepsilon^{1/2} e^{a(1+\kappa)/2},
\end{aligned}$$

using again that $\|\beta\|_\infty \lesssim 1$. Therefore $\xi_v^{-1}\,\mathbb{E}_{\xi_v}\left[\mathcal{E}_\varepsilon(g_\varepsilon)\right] \lesssim b_\varepsilon^{1/2}$, and we can apply Proposition 6.1 with rate function $\varphi(a) = \exp(\mu a)$ and dominating function $r(a) := e^{-(1+\kappa)|a|/2}$.

The terms I and II are bounded in the same way. We obtain

$$\mathbb{E}\left[B_{\varepsilon,\eta}(g)^2\right] \lesssim b_\varepsilon\big(\varepsilon\eta^{-1}\big)^{2\mu}\log(1/\eta)$$

for $\mu < \min\{\kappa/2, (1+\kappa)/2\} = \kappa/2$, uniformly over the class $\widetilde{\mathcal{C}}_{b_\varepsilon}(m)$. The trade-off between $M_{\varepsilon,\eta}(g_\varepsilon)$ and $B_{\varepsilon,\eta}(g_\varepsilon)$ yields the result. $\square$

Completion of proof of Theorem 3.2. We first write

$$\mathcal{E}(g_{\gamma_\varepsilon}) - m_1(\pi)^{-1} = \frac{1}{m_1(\pi)}\int_{1-\gamma_\varepsilon}^1 (f_{\gamma_\varepsilon}(a)-1)\beta(a)da.$$

We have

$$\Big|\int_{1-\gamma_\varepsilon}^1 \big(f_{\gamma_\varepsilon}(a)-1\big)\beta(a)da\Big| \leq 2\int_0^{-\log(1-\gamma_\varepsilon)} \pi(x)dx \lesssim \gamma_\varepsilon^{\kappa_2}$$

since $\pi \in \mathcal{R}(\kappa_2)$ and $-\log(1-x) \lesssim x$ for small enough $x \geq 0$. We derive

$$\big|\mathcal{E}(g_{\gamma_\varepsilon}) - m_1(\pi)^{-1}\big| \lesssim \gamma_\varepsilon^{\kappa_2}. \tag{5.8}$$

Next, by construction, $\gamma_\varepsilon^2 g_{\gamma_\varepsilon} \in \widetilde{\mathcal{C}}_{\gamma_\varepsilon}(1)$, hence for any $0 < \mu < \kappa_1$, Proposition 5.2 entails

$$\mathbb{E}\left[\big|\mathcal{E}_\varepsilon(g_{\gamma_\varepsilon}) - \mathcal{E}(g_{\gamma_\varepsilon})\big|\right] \lesssim \gamma_\varepsilon^{-1/2}\varepsilon^{\mu/(2\mu+2)}. \tag{5.9}$$

Moreover, since $\gamma_\varepsilon g_{\gamma_\varepsilon} \in \mathcal{C}'(1)$, we have, by (3.2) in Theorem 3.1

$$\mathbb{E}\left[\big|\mathcal{E}_\varepsilon(g_{\gamma_\varepsilon}) - \mathcal{E}_{\varepsilon,\sigma}(g_{\gamma_\varepsilon})\big|\right] \lesssim \gamma_\varepsilon^{-2}\sigma\varepsilon^{-1}. \tag{5.10}$$

The specification $\gamma_\varepsilon = \varepsilon^{\mu/(\mu+1)(2\kappa_2+1)}$ yields the correct rate for (5.8) and (5.9). The assumption that $\sigma\varepsilon^{-3}$ is bounded ensures that the right-hand side in (5.10) is asymptotically negligible. The conclusion readily follows for $\widehat{m}_{1,\varepsilon}$.

We now turn to higher moment estimators. Thanks to the proof for the case $k=1$, it suffices to show that

$$m_1(\pi)\mathcal{E}_\varepsilon\big(\widetilde{g}_{\gamma_\varepsilon}\big) \to \int_0^1 \log(1/a)^k\beta(a)da$$

in probability with the correct rate as $\varepsilon \to 0$. Note first that

$$\widetilde{g}_{\gamma_\varepsilon}(a) = -a\log(1/a)^k f'_{\gamma_\varepsilon}(1-a) + kf_{\gamma_\varepsilon}(1-a)\log(1/a)^{k-1}$$

is a sum of a function in $\widetilde{\mathcal{C}}_{\gamma_\varepsilon}(c_1)$ and a function in $\mathcal{C}'(c_2)$, for some positive c_1 and c_2, hence we may apply Proposition 5.2 and Theorem 3.1 to each term respectively. Next

$$\begin{aligned} c(\pi)\mathcal{E}\big(\widetilde{g}_{\gamma_\varepsilon}\big) &= -\int_0^1 \big(f_{\gamma_\varepsilon}(1-a)\log(1/a)^k\big)' \int_0^a \beta(u)du\,da \\ &= \int_0^1 f_{\gamma_\varepsilon}(1-a)\log(1/a)^k\beta(a)da, \end{aligned}$$

since $k > 1$ and $\pi \in \Pi(\kappa_1)$ so that the boundary terms vanish when integrating by part. We conclude by noticing that by Hölder's inequality, for any $\tau > 0$, we have

$$\begin{aligned} &\big|\int_0^1 \big(1 - f_{\gamma_\varepsilon}(1-a)\big)\log(1/a)^k\beta(a)da\big| \\ \leq& 2\Big(\int_{-\log\gamma_\varepsilon}^{+\infty}\pi(x)dx\Big)^{1-\tau}\Big(\int_0^{+\infty}x^{k/\tau}\pi(x)dx\Big)^{\tau} \lesssim \gamma_\varepsilon^{\kappa_1(1-\tau)}. \end{aligned}$$

This term also has the right order since the choice of τ is free and $\kappa_1 > \kappa_2$ by assumption. The proof of Theorem 3.2 is complete.

5.3 Proof of Theorem 3.4

With no loss of generality, we consider the homogeneous case with $\alpha = 0$. We may also assume that $\sigma = 0$, since adding experimental noise to the observation of the fragments only increases the error bounds.

Step 1: An augmented experiment. In the binary case, the dislocation measure $\nu(d\mathbf{s})$ is equivalently mapped by a probability measure on $[1/2, 1]$ with density function $a \rightsquigarrow \rho(a)$, see (3.6).

We prove a lower bound in the augmented experiment where one can observe all the sizes $\widetilde{X}_\varepsilon$ of the fragments until they become smaller than ε, namely

$$\widetilde{X}_\varepsilon := \big\{\xi_u,\, \xi_{u-} \geq \varepsilon\big\} \cup \big\{\xi_u,\, u \in \mathcal{U}_\varepsilon\big\}$$

Clearly, taking the infimum over all estimators based on $\widetilde{X}_\varepsilon$ instead of $X_\varepsilon = X_{\varepsilon,0}$ only reduces the lower bound.

For every $u \in \mathcal{U}_\varepsilon$, we have $\xi_{u-} \geq \varepsilon$. By the conservative Assumption B, there are at most ε^{-1} such ξ_{u-} so $\mathrm{Card}\,\mathcal{U}_\varepsilon \leq 2\varepsilon^{-1}$. For every node $u \in \mathcal{U}$, the fragmentation process gives rize to two offsprings with size $\xi_u U$ and $\xi_u(1-U)$, where U is a random variable independent of ξ_u with density function $\rho(\bullet)$. Therefore, the process of the sizes of the fragments in the enlarged experiment can be realized by less than

$$2\varepsilon^{-1}\big(1+\tfrac{1}{2}+\ldots+\tfrac{1}{2^{k(\varepsilon)}}\big) \leq \lfloor 4\varepsilon^{-1} \rfloor + 1 =: n(\varepsilon)$$

independent realizations of the law $\rho(\bullet)$, where $k(\varepsilon) := \log_2(2/\varepsilon)$, assumed to be integer with no loss of generality.

In turn, Theorem 3.4 reduces to proving that $\varepsilon^{1/2}$ is a lower rate of convergence for estimating $m_k(\pi)$ based on the observation of a $n(\varepsilon)$-sample of the law $\rho(\bullet)$. The one-to-one correspondence between $\rho(\bullet)$ and $\pi(\bullet)$ is given in (4.1).

Step 2: Construction of π_ε. We write $\rho_\pi(\bullet)$ to emphasize the dependence upon $\pi(\bullet)$. Let

$$\varphi_k(a) := a(\log(1/a)^k + (1-a)\log\big(1/(1-a)\big)^k, \;\; a \in [1/2, 1].$$

From (4.1), we have

$$m_k(\pi_0) = \int_{1/2}^1 \varphi_k(a)\rho_{\pi_0}(a)da.$$

Let $0 < \tau < 1$. Pick a function $\psi_k(\bullet) : [1/2, 1] \to \mathbb{R}$ such that

$$\|\psi_k\|_\infty \leq \tau \inf_a \rho_{\pi_0}(a), \;\; \int_{1/2}^1 \psi_k(a)da = 0, \;\; r(k) := \int_{1/2}^1 \varphi_k(a)\psi_k(a)da \neq 0,$$

a choice which is obviously possible thanks to Assumption D. For $\varepsilon > 0$, define

$$\rho_{\pi_\varepsilon}(a) := \rho_{\pi_0}(a) + \varepsilon^{1/2}\psi_k(a), \;\; a \in [1/2, 1].$$

(And so (4.1) defines $\pi_\varepsilon(\bullet)$ unambiguously.) By construction, $\rho_{\pi_\varepsilon}(\bullet)$ is a density function on $[1/2, 1]$ and has a corresponding binary fragmentation with Lévy measure given by $\pi_\varepsilon(\bullet)$. Moreover,

$$m_k(\pi_\varepsilon) = m_k(\pi_0) + r(k)\varepsilon^{1/2}.$$

Step 3: A two-point lower bound. The following chain of arguments is fairly classical. We denote by $\widetilde{\mathbb{P}}_\pi$ the law of the independent random variables $\big(U_i, i = 1, \ldots, n(\varepsilon)\big)$ with common density $\rho_\pi(\bullet)$ that we use to realize the augmented experiment.

Let F_ε be an arbitrary estimator based on $\widetilde{X}_\varepsilon$. Put $c := |r(k)|/2$. We have

$$\max_{\pi\in\{\pi_0,\pi_\varepsilon\}} \widetilde{\mathbb{P}}_\pi\big[\varepsilon^{-1/2}|F_\varepsilon - m_k(\pi)| \geq c\big]$$
$$\geq \tfrac{1}{2}\Big(\widetilde{\mathbb{P}}_{\pi_0}\big[\varepsilon^{-1/2}|F_\varepsilon - m_k(\pi_0)| \geq c\big] + \widetilde{\mathbb{P}}_{\pi_\varepsilon}\big[\varepsilon^{-1/2}|F_\varepsilon - m_k(\pi_\varepsilon)| \geq c\big]\Big)$$
$$\geq \tfrac{1}{2}\,\widetilde{\mathbb{E}}_{\pi_0}\Big[1_{\{\varepsilon^{-1/2}|F_\varepsilon - m_k(\pi_0)|\geq c\}} + 1_{\{\varepsilon^{-1/2}|F_\varepsilon - m_k(\pi_\varepsilon)|\geq c\}}\Big] - \tfrac{1}{2}\|\widetilde{\mathbb{P}}_{\pi_0} - \widetilde{\mathbb{P}}_{\pi_\varepsilon}\|_{TV},$$

where $\|\bullet\|_{TV}$ denotes the total variation distance between probability measures. By the triangle inequality, we have

$$\varepsilon^{-1/2}\big(|F_\varepsilon - m_k(\pi_0)| + |F_\varepsilon - m_k(\pi_\varepsilon)|\big) \geq |r(k)| = 2c,$$

so one of the two indicators within the expectation above must be equal to one with full $\widetilde{\mathbb{P}}_{\pi_0}$-probability. Therefore

$$\max_{\pi\in\{\pi_0,\pi_\varepsilon\}} \widetilde{\mathbb{P}}_\pi\big[\varepsilon^{-1/2}|F_\varepsilon - m_k(\pi)| \geq c\big] \geq \tfrac{1}{2}(1 - \|\widetilde{\mathbb{P}}_{\pi_0} - \widetilde{\mathbb{P}}_{\pi_\varepsilon}\|_{TV}),$$

and Theorem 3.4 is proved if

$$\limsup_{\varepsilon\to 0} \|\widetilde{\mathbb{P}}_{\pi_0} - \widetilde{\mathbb{P}}_{\pi_\varepsilon}\|_{TV} < 1. \tag{5.11}$$

By Pinsker's inequality

$$\|\widetilde{\mathbb{P}}_{\pi_0} - \widetilde{\mathbb{P}}_{\pi_\varepsilon}\|_{TV} \leq \frac{\sqrt{2}}{2}\Big(\widetilde{\mathbb{E}}_{\pi_0}\Big[\log\frac{d\,\widetilde{\mathbb{P}}_{\pi_0}}{d\,\widetilde{\mathbb{P}}_{\pi_\varepsilon}}\Big]\Big)^{1/2},$$

and

$$\widetilde{\mathbb{E}}_{\pi_0}\Big[\log\frac{d\,\widetilde{\mathbb{P}}_{\pi_0}}{d\,\widetilde{\mathbb{P}}_{\pi_\varepsilon}}\Big] = -\sum_{i=1}^{n(\varepsilon)} \widetilde{\mathbb{E}}_{\pi_0}\left[\log\frac{\rho_{\pi_\varepsilon}(U_i)}{\rho_{\pi_0}(U_i)}\right]$$
$$= -\sum_{i=1}^{n(\varepsilon)} \widetilde{\mathbb{E}}_{\pi_0}\left[\log\big(1+\varepsilon^{1/2}\psi_k(U_i)\rho_{\pi_0}(U_i)^{-1}\big) - \varepsilon^{1/2}\psi_k(U_i)\rho_{\pi_0}(U_i)^{-1}\right],$$

where we used

$$\widetilde{\mathbb{E}}_{\pi_0}\left[\psi_k(U_i)\rho_{\pi_0}(U_i)^{-1}\right] = \int_{1/2}^{1}\psi_k(a)da = 0.$$

We also have $\varepsilon^{1/2}|\psi_k(U_i)\rho_{\pi_0}(U_i)^{-1}| \leq \tau\varepsilon^{1/2}$ hence for small enough τ,

$$\big|-\log\big(1+\varepsilon^{1/2}\psi_k(U_i)\rho_{\pi_0}(U_i)^{-1}\big) + \varepsilon^{1/2}\psi_k(U_i)\rho_{\pi_0}(U_i)^{-1}\big| \leq \tau^2\varepsilon.$$

Therefore $\|\widetilde{\mathbb{P}}_{\pi_0} - \widetilde{\mathbb{P}}_{\pi_\varepsilon}\|_{TV} \leq \frac{\sqrt{2}}{2}\tau\varepsilon^{1/2}n(\varepsilon)^{1/2} < 1$ by picking τ small enough, and (5.11) follows. The proof of Theorem 3.4 is complete.

5.4 Proof of Theorem 3.5

We plan to use the following decomposition

$$\widehat{\beta}(a) - \beta(a) = \widehat{m}_{1,\varepsilon}\,\mathcal{E}_{\varepsilon,\sigma}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big) - \beta(a) = I + II + III + IV,$$

with

$$\begin{aligned}
I &:= \widehat{m}_{1,\varepsilon}\Big(\mathcal{E}_{\varepsilon,\sigma}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big) - \mathcal{E}_{\varepsilon}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big)\Big)\\
II &:= \widehat{m}_{1,\varepsilon}\Big(\mathcal{E}_{\varepsilon}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big) - \mathcal{E}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big)\Big),\\
III &:= \big(\widehat{m}_{1,\varepsilon} - m_1(\pi)\big)\,\mathcal{E}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big),\\
IV &:= m_1(\pi)\,\mathcal{E}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big) - \beta(a).
\end{aligned}$$

Considering I and II, the term $\widehat{m}_{1,\varepsilon}$ is bounded in probability by Theorem 3.2. By (3.2) in Theorem 3.1 together with the fact that $\gamma_\varepsilon^3\varphi'_{\gamma_\varepsilon} \in \mathcal{C}'(\|\varphi'\|_\infty)$, we have

$$\mathbb{E}\left[\big|\mathcal{E}_{\varepsilon}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big) - \mathcal{E}_{\varepsilon,\sigma}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big)\big|\right] \lesssim \gamma_\varepsilon^{-3}\sigma\varepsilon^{-1}.$$

By construction, we have $\gamma_\varepsilon^2\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet) \in \widetilde{\mathcal{C}}_{\gamma_\varepsilon}(\|\varphi\|_\infty)$, therefore, by Proposition 5.2

$$\mathbb{E}\left[\Big(\mathcal{E}_{\varepsilon}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big) - \mathcal{E}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big)\Big)^2\right] \lesssim \gamma_\varepsilon^{-3}\varepsilon^{\mu/(\mu+1)}. \tag{5.12}$$

Considering III, using (5.8), we have $\big|\mathcal{E}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big)\big| \lesssim \gamma_\varepsilon^{-1}$. By Theorem 3.2, we conclude that III^2 has order

$$\gamma_\varepsilon^{-2}\varepsilon^{2\mu\kappa_2/(\mu+1)(2\kappa_2+1)} \tag{5.13}$$

in probability. For IV, we first notice that

$$m_1(\pi)\,\mathcal{E}\big(-\bullet\,\varphi'_{\gamma_\varepsilon,a}(\bullet)\big) = \int_0^1 \varphi_{\gamma_\varepsilon,a}(u)\beta(u)du,$$

hence

$$IV^2 = \Big(\int_0^1 \varphi_{\gamma_\varepsilon,a}(u)\beta(u)du - \beta(a)\Big)^2 \lesssim \gamma_\varepsilon^{2s} \tag{5.14}$$

by a Taylor expansion and using that the terms up to order $s-1$ vanish by the cancellation property (3.7) of $\varphi(\bullet)$ since $s < N$.

Putting together (5.12), (5.13) and (5.14), we see that the specification $\gamma_\varepsilon = \varepsilon^{\mu/(\mu+1)(2s+3)}$ yields the correct rate for II and IV, as well as for III as soon as $\kappa_2 \geq s+1$. Finally, the term I proves asymptotically negligible in the same way as (5.10) thanks to the assumption that $\sigma\varepsilon^{-3}$ is bounded. The proof of Theorem 3.5 is complete.

6 Appendix

6.1 Proof of Lemma 5.1

First, we enrich the structure of the genealogical tree representation of Section 2 by adding a random mark $M : \mathcal{U} \to [0,1] \times [0,+\infty) \times [0,+\infty)$ on the tree together with a random branch[8] β of $\mathcal{U}$ and define unambiguously the law $\mathbb{P}^\star$ of (M, β) by setting

$$\mathbb{E}^\star \left[\Phi_n(M,\beta) \right] := \mathbb{E}\left[\sum_{|u|=n} \Phi_n(M,u)\xi_u \right], \quad n \geq 1,$$

where Φ_n is a bounded functional which depends on the mark M and the branch β up to the n-th generation only. If β_n is the node of the random branch at the n-th generation, we set $\chi_n := \xi_{\beta_n}$ and $\chi(t)$ for the size of the tagged particle at time t:

$$\chi(t) := \begin{cases} \chi_n & \text{if} \quad a_{\beta_n} \leq t < a_{\beta_n} + \zeta_{\beta_n}, \\ 0 & \text{if} \quad t \geq \lim_{n\to\infty} a_{\beta_n}, \end{cases}$$

where a_{β_n} and ζ_{β_n} denote respectively the birth-time and lifetime of the particle labeled by the tagged node β_n. We then have

$$\begin{aligned} \mathbb{E}\left[\sum_{v\in\mathcal{U}_\varepsilon} \xi_v f(\xi_v) \right] &= \mathbb{E}\left[\sum_{n=0}^{\infty} \sum_{|v|=n} 1_{\left\{\xi_{v-}\geq\eta,\, \xi_v<\eta\right\}} \xi_v f(\xi_v) \right] \\ &= \sum_{n=0}^{\infty} \mathbb{E}^\star \left[1_{\left\{\chi_{n-1}\geq\eta,\, \chi_n<\eta\right\}} f(\chi_n) \right]. \end{aligned}$$

By Proposition 1.6 in [5], $-\log \chi_n$ is a random walk under $\mathbb{P}^\star$ with step distribution $\pi(dx)$. In particular, the last term above is independent of α.

Consider now a homogeneous fragmentation process with same dislocation measure $\nu(\bullet)$ living on the same (possibly enlarged) probability space for simplicity. Applying the same construction above, we obtain a process $\left(\chi^{(0)}(t), t \geq 0\right)$ that can be expressed in the form $\chi^{(0)}(t) = \exp\left(-\zeta^{(0)}(t)\right)$ where $\left(\zeta^{(0)}(t), t \geq 0\right)$ is a compound Poisson process with jump intensity 1 and jump distribution $\pi(\bullet)$. By construction, we have $\chi^{(0)}\left(T_\eta^{(0)}\right) = \chi_n^{(0)}$ on the event $\{\chi_{n-1}^{(0)} \geq \eta,\, \chi_n^{(0)} < \eta\}$ therefore

$$\mathbb{E}^\star \left[f\left(\chi^{(0)}(T_\eta^{(0)})\right) \right] = \sum_{n=0}^{\infty} \mathbb{E}^\star \left[1_{\left\{\chi_{n-1}^{(0)}\geq\eta,\, \chi_n^{(0)}<\eta\right\}} f\left(\chi_n^{(0)}\right) \right].$$

[8] A branch is an infinite sequence of positive integers which we can think of as the line of ancestors of some leaf of the tree.

The conclusion follows by identifying the right-hand sides of the last two equalities.

6.2 Rates of convergence in the key renewal theorem

We give a version of Sgibnev's result [16] on uniform rates of convergence in the key renewal theorem, which is proved in a more general setting.

Let $F(dx)$ be a probability distribution with positive mean m and renewal function $H = \sum_{n=0}^{\infty} F^{n\star}$, with $F^{0\star} := \delta_0$, $F^{1\star} := F$ and $F^{(n+1)\star} := F \star F^{n\star}$, $n \geq 0$.

We assume that F is *spread-out*, that is, for some $n \geq 1$, F^{*n} has a non-zero absolutely continuous component. Stone [17] shows that then there exists a decomposition $H = H_1 + H_2$, where H_2 is finite measure and H_1 is absolutely continuous with bounded continuous density function $h(\bullet)$ such that $\lim_{x\to+\infty} h(x) = m^{-1}$ and $\lim_{x\to-\infty} h(x) = 0$.

We denote by $T(F)$ the σ-finite measure with density function

$$\int_{(x,+\infty)} F(du)1_{[0,+\infty)}(x) - \int_{(-\infty,x]} F(du)1_{(-\infty,0)}(x).$$

and $T^2(F) := T\big(T(F)\big)$. Let $\varphi(\bullet) : \mathbb{R} \to [0, +\infty)$ be a submultiplicative function, *i.e.* such that $\varphi(0) = 1$, $\varphi(x + y) \leq \varphi(x)\varphi(y)$. Then we have (see for instance [10], Section 6)

$$-\infty < r_1 := \lim_{x\to-\infty} \frac{\log \varphi(x)}{x} \leq \lim_{x\to+\infty} \frac{\log \varphi(x)}{x} =: r_2 < +\infty,$$

Assumption E. *We have $r_1 \leq 0 \leq r_2$ and there exists $r(\bullet) : \mathbb{R} \to \mathbb{R}$ an integrable function and such that the following conditions are fulfilled:*

$$\sup_x |r(x)|\varphi(x) < +\infty \ , \quad \lim_{|x|\to\infty} r(x)\varphi(x) = 0,$$

$$\lim_{x\to+\infty} \varphi(x) \int_{[x,+\infty)} r(u)du = \lim_{x\to-\infty} \varphi(x) \int_{(-\infty,x]} r(u)du = 0,$$

and

$$\int_{\mathbb{R}} \varphi(x)T^2(F)(dx) < \infty.$$

Sgibnev's result takes the form:

Proposition 6.1. *(Theorem 5.1 in [16]). Grant Assumption E. Then*

$$\lim_{|t|\to\infty} \varphi(t) \sup_{g,|g(x)|\le|r(x)|} \left| g \star H(t) - m^{-1} \int_{\mathbb{R}} g(x)dx \right| = 0.$$

We call $\varphi(\bullet)$ a rate function and $r(\bullet)$ a dominating function.

References

[1] Aldous, D.J. (1999). Deterministic and stochastic models for coalescence (aggregation, coagulation): a review of the mean-field theory for probabilists. *Bernoulli* **5**, 3-48.

[2] Asmussen, S. (2003). *Applied probability and queues*, 2nd ed., Springer-Verlag, New York, 2003.

[3] Bertoin, J. (1996). *Lévy processes.* Cambridge University Press.

[4] Bertoin, J. (2002). Self similar fragmentations. *Ann. Inst. Henri poincaré. Probabilités et Statistiques* **38**, 319–340.

[5] Bertoin, J. (2006). *Random fragmentation and Coagulation Processes.* Cambridge University Press.

[6] Bertoin, J. and Caballero, M-C. (2002). Entrance from 0+ for increasing semi-stable Markov processes. *Bernoulli* **8**, no. 2, 195–205.

[7] Bertoin, J. and Martinez, S. (2005). Fragmentation energy. *adv. Appl. Probab.* **37**, 553–570.

[8] Brennan, M.D. and Durett, R. (1987). Splitting intervals. II. Limit laws for lengths. *Probab. Theory Related Fields* **75**, 109–127.

[9] Filippov, A.F. (1961). On the distribution of the sizes of particles which undergo splitting. *Th. Probab. Appl.* **6**, 275–293.

[10] Hille, E. and Phillips, R.S. (1957). *Functional Analysis and Semigroups.* Amer. Math. Soc. Colloquium Publications, vol. 31, Providence.

[11] Kolmogorov, A.N. (1941). Über das logaritmisch normale Verteinlungsgesetz der Dimensionen des Teilchen bei Zerstückelung. *C.R. (Doklady) Acad. Sci. USSR* **31**, 99–101.

[12] Korostelev, A.P. and Tsybakov, A.B. (1993). *Minimax theory of image reconstruction*. Lecture notes in Statistics, vol. 82, Springer, Berlin.

[13] Krapivsky, P.L. and Ben-Naim, E. (1994). Scaling and multiscaling in models of fragmentation. *Phys. Rev. E* **50**, 3502–3507.

[14] Krapivsky, P.L., Ben-Naim, E. and Grosse, I. (2004). Stable distributions in stochastic fragmentation. *J. Phys. A* **37**, 2863–2880.

[15] Ney, P. (1981). A refinement of the coupling method in renewal theory. *Stoch. Processes Appl.*, **11**, 11–26.

[16] Sgibnev, M.S. (2002). Stone's decomposition of the renewal measure via Banach-algebraic techniques. *Proc. of the American Math. Soc.*, **130**, 2425–2430.

[17] Stone, C. (1966). On absolutey continuous components and renewal theory. *Ann. Math. Statist.* **37**, 271–275.