The Annals of Probability
2022, Vol. 50, No. 3, 1173–1203
https://doi.org/10.1214/21-AOP1556
© Institute of Mathematical Statistics, 2022

SHARP CONCENTRATION FOR THE LARGEST AND SMALLEST FRAGMENT IN A k-REGULAR SELF-SIMILAR FRAGMENTATION

BY PIOTR DYSZEWSKI[1,a], NINA GANTERT[1,b], SAMUEL G. G. JOHNSTON[2,d], JOSCHA PROCHNO[3,e] AND DOMINIK SCHMID[1,c]

[1]*Technical University of Munich,* [a]*piotr.dyszewski@tum.de,* [b]*gantert@ma.tum.de,* [c]*dominik.schmid@tum.de*
[2]*University of Bath,* [d]*sgj22@bath.ac.uk*
[3]*University of Passau,* [e]*joscha.prochno@uni-passau.de*

We study the asymptotics of the k-regular self-similar fragmentation process. For $\alpha > 0$ and an integer $k \geq 2$, this is the Markov process $(I_t)_{t\geq 0}$ in which each I_t is a union of open subsets of $[0, 1)$, and independently each subinterval of I_t of size u breaks into k equally sized pieces at rate u^α. Let k^{-m_t} and k^{-M_t} be the respective sizes of the largest and smallest fragments in I_t. By relating $(I_t)_{t\geq 0}$ to a branching random walk, we find that there exist explicit deterministic functions $g(t)$ and $h(t)$ such that $|m_t - g(t)| \leq 1$ and $|M_t - h(t)| \leq 1$ for all sufficiently large t. Furthermore, for each n, we study the final time at which fragments of size k^{-n} exist. In particular, by relating our branching random walk to a certain point process, we show that, after suitable rescaling, the laws of these times converge to a Gumbel distribution as $n \to \infty$.

1. Introduction. Eighty years ago, Kolmogorov [17] initiated the study of fragmentation processes, stochastic processes modelling an object of unit mass that breaks apart as time passes. While research in fragmentation processes continued into the latter half of the 20th century [2, 7, 8, 13], it was not until pathbreaking work by Bertoin [4, 5] and Berestycki [3] in the early 2000s that fragmentation processes were conceived in a unifying framework. This framework formulates a fragmentation process in terms of a stochastic process $(Y_t)_{t\geq 0}$ taking values in the set

$$\mathcal{S} := \left\{ (s_1, s_2, s_3, \dots) : s_1 \geq s_2 \geq \cdots \geq 0, \sum_{i=1}^{\infty} s_i \leq 1 \right\}, \tag{1}$$

whose law is governed by a *dislocation measure* ν on the set $\mathcal{S}$. For $Y_t = (y_1(t), y_2(t), \dots)$, the components $y_1(t) \geq y_2(t) \geq \dots$ of Y_t correspond to the sizes of the fragments in the process at time t listed in decreasing order.

In the setting where ν is finite, the homogeneous fragmentation processes first introduced by Bertoin [4] have a simple description in terms of the dislocation measure: each fragment of size u has an exponentially distributed lifetime with rate $\nu(\mathcal{S})$, and upon death is replaced by a random collection of fragments of sizes $us_1 \geq us_2 \geq \cdots$, where the sequence $(s_1, s_2, \dots)$ is distributed according to $\nu(\cdot)/\nu(\mathcal{S})$. In this context, *homogeneous* refers to the fact that the rate at which each fragment breaks is independent of its size, and that the lifetimes and dislocations of individual fragments are independent of the remainder of the system; see also [18]. We remark that in general the measure ν need not be finite; indeed, infinite dislocation measures may be used to describe the continuous 'crumbling' of fragments [4].

In the following, we will be interested in *self-similar fragmentation processes*, in which fragments behave independently but the rate at which a fragment of size u breaks apart is

Received February 2021; revised October 2021.
MSC2020 subject classifications. 60J27, 60J80, 60G55.
Key words and phrases. Fragmentation, branching random walk, point process.

proportional to u^α for some α in $\mathbb{R}$. Self-similar fragmentations were introduced by Filippov [13], with their rigorous formulation in terms of general dislocation measures first appearing in [5]. The real parameter α is a called *the index of self-similarity*, with $\alpha > 0$ entailing that larger fragments in the process break more quickly than smaller ones, and $\alpha < 0$ entailing the opposite.

Brennan and Durrett [7, 8] study the self-similar fragmentation process where, upon death, a fragment of mass of size u splits into exactly two fragments of sizes Vu and $(1-V)u$, where V is uniformly distributed on $[0, 1]$. They show that at large times t the total number of intervals in the process grows in the order $t^{1/\alpha}$ for $0 < \alpha < \infty$. Goldschmidt and Haas [14, 15] look at the explosive case $\alpha < 0$, in which after a finite amount of time the entire process consists of *dust* so that there are no intervals of positive size. A work of particular relevance is the article [6] of Bertoin, where it is shown that if $y_1(t)$ is the size of the largest fragment in a self-similar fragmentation with $\alpha > 0$, then

$$\lim_{t\to\infty} \frac{\log y_1(t)}{\log t} = -\frac{1}{\alpha} \quad \text{almost surely.} \tag{2}$$

See also the recent work of Dadoun [9] for growth-fragmentation processes. While a panoply of exotic dislocation mechanisms fall into the general apparatus of self-similar fragmentation processes, in the present article we will concentrate our attention on the simplest possible fragmentation mechanism:

DEFINITION 1.1. Fix an integer $k \geq 2$. The k-regular self-similar fragmentation process of index $\alpha \in \mathbb{R}$ is the self-similar fragmentation process $(I_t)_{t\geq 0}$ starting with the single interval $I_0 := [0, 1)$ in which an interval of size $u \in (0, 1]$ in I_t waits an exponential time with mean $u^{-\alpha}$, and after this time breaks into k equally sized intervals.

Note that by listing the sizes of the intervals of $(I_t)_{t\geq 0}$ in decreasing order, $(I_t)_{t\geq 0}$ gives rise to an $\mathcal{S}$-valued process $(Y_t)_{t\geq 0}$. The dislocation measure associated with the k-regular case belongs to a form of dislocation measures which Goldschmidt and Haas [15] call 'geometric', in that fragment sizes always take the form of a geometric progression $(r^n : n \geq 0)$ for some $r \in (0, 1)$. Goldschmidt and Haas remark that geometric fragmentation processes possess genuinely different properties from nongeometric fragmentations, and should not be regarded as a degenerate special case. The reader is referred to [15], Section 8, for a discussion, wherein various other relevant references may be found, for example, Athreya [2].

The relative simplicity of the mechanism means that we are endowed with a variety of exact formulas associated with various functionals of the processes, most notably allowing us to study an alternative representation for the process, where the fragments of sizes k^{-n} are viewed as the nth generation of a discrete k-ary tree. These exact formulas lead to sharp statements about the asymptotics of the size of the smallest and largest fragments in the process at large times.

In the remainder of the paper, we restrict our attention to the case $\alpha > 0$. Before stating our results in full in Section 2, we conclude the Introduction by giving the principal applications of our main results, showing that we can characterise the sizes of both the largest and smallest fragment at large times to a surprising degree of precision.

In the sequel for $x \in \mathbb{R}$, we write $\lceil x \rceil$ for the least integer greater than or equal to x, and will denote by $\mathbb{R}_+$ the set of nonnegative real numbers $[0, \infty)$. Finally, let us introduce the parameters

$$\gamma := \log k \quad \text{and} \quad \kappa := \frac{1}{\gamma\alpha}.$$

Our main result on the largest fragment is a considerable sharpening of Bertoin's estimate (2), stating that if k^{-m_t} is the size of the largest fragment at time t, then m_t has very concentrated behaviour.

THEOREM A. *Let k^{-m_t} be the size of the largest fragment in the system at time t. Then for most times t, m_t is likely to be the smallest integer above*

$$g(t) = \kappa\big(\log t - \log\log t - \log(\gamma\kappa)\big).$$

More precisely, let $\mu_1 := \kappa + 2/\gamma$. Then there exists almost surely a $t_0 \in \mathbb{R}_+$ such that for all $t \geq t_0$

$$\left\lceil g(t) - \mu_1 \frac{\log\log t}{\log t} \right\rceil \leq m_t \leq \left\lceil g(t) + \mu_1 \frac{\log\log t}{\log t} \right\rceil.$$

Roughly speaking, we have that for most values of t the quantities $\lceil g(t) - \mu_1 \log\log t/\log t\rceil$ and $\lceil g(t) + \mu_1 \log\log t/\log t\rceil$ coincide. Hence, Theorem A guarantees that for such t we have $m_t = \lceil g(t)\rceil$. Occasionally an integer n separates $g(t) - \mu_1 \log\log t/\log t$ and $g(t) + \mu_1 \log\log t/\log t$; it is in these time windows that m_t has an opportunity to 'jump' from n to $n+1$.

We now turn our attention to the size k^{-M_t} of the smallest fragment. Here we find that for large t, the law of the random variable M_t is also highly concentrated.

THEOREM B. *Let k^{-M_t} be the size of the smallest fragment in the system at time t. Then for most times t, M_t is likely to be the smallest integer above*

$$h(t) := \kappa\left(\log t + \sqrt{2\gamma \log t} - \frac{1}{2}\log\log t + c\right),$$

where $c := -\frac{1}{2\kappa} - \log\kappa + \gamma - \frac{1}{2}\log(2\gamma) + 1$ is a constant. More precisely, let $\mu_2 := 2\kappa^{2/3}$. Then there exists almost surely a $t_0 \in \mathbb{R}_+$ such that for all $t \geq t_0$

$$\left\lceil h(t) - \mu_2 \frac{1}{\log^{1/3} t} \right\rceil \leq M_t \leq \left\lceil h(t) + \mu_2 \frac{1}{\log^{1/3} t} \right\rceil.$$

In words, Theorems A and B state that the logarithm of the sizes of all fragments in the system at a large time t have the same first order approximation, which is of order $(1 + o(1))\kappa \log t$. For the largest fragment, we have a correction of order $\log\log t$, while we see a correction of order $\sqrt{\log t}$ for the smallest fragment. Moreover, the largest and smallest fragments are both with high probability pinpointed to specific integers.

The rest of the paper is organised as follows. In Section 2, we describe the representation of the fragmentation process as a certain time-inhomogeneous branching random walk, which is key to our proofs. We also give our main results on point process convergence (Theorem C and Corollary 2.1) for the branching random walk, which explain why the sizes of the largest and smallest fragment satisfy such a sharp concentration property. In Section 3, we study weighted sums of exponential random variables and their relation to the q-Markov chain: the increasing Markov chain on $\{0, 1, 2, \dots\}$ which jumps from a site j to $j+1$ at rate q^j (in our case, $q = k^{-\alpha}$). In Section 4, we prove our results on the largest fragments in the process. The final two sections, Section 5 and Section 6, are dedicated to our work on the smallest fragments in the process.

2. The associated branching random walk. This section is dedicated to giving a complete statement of our main results in their general form. Through the majority of the proofs in the paper, we consider the fragments as vertices in a k-ary tree, where the offspring of an interval are the k intervals it splits into. We study the time of the appearance of the fragments using the fact that the fragmentation can be represented as a certain inhomogeneous branching random walk which we shall now describe; see Figure 1.

2.1. *Representation as BRW.* We now explain the representation of the process in terms of an expanding branching random walk, see [2]. For the k-regular self-similar fragmentation of index α, we set

$$q := k^{-\alpha}, \tag{3}$$

and note that $q < 1$ under our assumption $\alpha \in (0, \infty)$. Let k^{-X_t} be the size of the fragment containing 0, in other words the fragment of the form $[0, k^{-X_t})$, present in the system at time t. Then the process $(X_t)_{t \geq 0}$ forms a Markov chain on $\{0, 1, 2, \dots\}$ satisfying $X_0 = 0$ and

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}(X_{t+h} = j | X_t = i) = \begin{cases} \lambda_i & \text{if } j = i + 1, \\ 0 & \text{otherwise,} \end{cases}$$

where $\lambda_i = q^i$. For $n \geq 0$, define the time of fragmentation of the fragment $[0, k^{-n})$ into k fragments of sizes $k^{-(n+1)}$ to be $S_n := \sup\{t \geq 0 : X_t = n\}$. It follows that $\{X_t = n\} = \{S_n > t, S_{n-1} \leq t\}$, and that $\{X_t \leq n\} = \{S_n > t\}$. Moreover, we may write

$$S_n = \sum_{i=0}^{n} \lambda_i^{-1} W_i, \tag{4}$$

where, for each $i \in \{0, 1, \dots, n\}$, $\lambda_i^{-1} W_i$ is the amount of time X_t spends at the state i, and hence W_i is a standard exponential random variable.

The same analysis can be carried through on every interval. The dependence structure in the resulting system can be described using branching processes. Each interval breaks into k pieces and thus we may consider each interval v of size k^{-n} living for some time period as a vertex v within the nth generation of a k-regular tree. Indeed, let $\mathbb{T}_n$ denote the set of subintervals of the form $[m/k^n, (m+1)/k^n)$, $m \in \{0, 1, \dots, k^n - 1\}$ so that $\mathbb{T}_n$ has k^n elements. Write $\mathbb{T} = \bigcup_{n \in \mathbb{N}} \mathbb{T}_n$ for the set of all subintervals that can appear in the system. For $v \in \mathbb{T}$, let $k^{-|v|}$ denote the size of v, in other words $|v| = n$ for $v \in \mathbb{T}_n$. Finally for intervals $v, w \in \mathbb{T}$ let $v \wedge w$ denote the smallest (in the sense of inclusion) element of $\mathbb{T}$ containing v and w. Then $v \wedge w$ is the most recent common ancestor of v and w. We will also write $v \geq w$ whenever $v \subseteq w$. Letting $S(v)$ denote the time at which an element v of $\mathbb{T}_n$ of size k^{-n} breaks into k pieces of sizes $k^{-(n+1)}$, we now see that the set $\{S(v) : v \in \mathbb{T}_n\}$ coincides with the positions of the nth generation of a certain branching random walk in which the step size distribution changes from generation to generation. The time of the first splitting is a standard exponential random variable $S([0, 1)) = W^{([0,1))}$ and each particle in generation n has exactly k children in generation $n + 1$. If $w \in \mathbb{T}_{n+1}$ is a child of $v \in \mathbb{T}_n$, then

$$S(w) = S(v) + q^{-|w|} W^{(w)}, \tag{5}$$

where $W^{(w)}$ is a standard exponential random variable which is independent of $S(v)$. In fact, the random variable $q^{-|w|} W^{(w)}$ is equal to the length of time that the interval w exists in the process until it splits. Since $q < 1$, this random walk gets slower and slower as n becomes

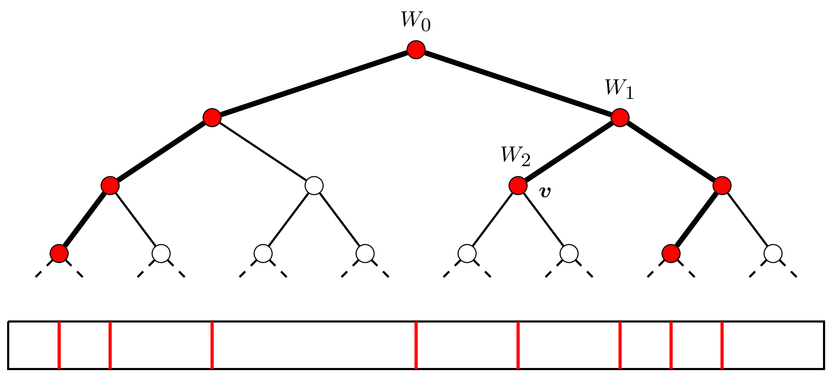

FIG. 1. *Visualization of a 2-regular self-similar fragmentation of index α, and the genealogical tree of its associated branching random walk at some time $t \geq 0$. All sites present in the tree at time t are marked in red. Note that the i.i.d. standard exponential random variables $(W_i)_{i\in\mathbb{N}}$ in the definition of $S(v)$ for the site v must satisfy $\sum_{i=0}^{|v|} 2^{\alpha i} W_i \leq t$.*

large. We will refer to $S = (S(v))_{v\in\mathbb{T}}$ as the *expanding branching random walk* as proposed in [2]. It is natural to consider for all $v \in \mathbb{T}$ the rescaled quantities

$$K(v) := q^{|v|} S(v).$$

We will refer to $(K(v))_{v\in\mathbb{T}}$ as the *rescaled expanding branching random walk*. The jumps in the rescaled branching random walk have the simple description that if w is a child of v, then

$$K(w) = qK(v) + W^{(w)};$$

that is, a particle inherits q times their parent's position, plus a standard exponential. It is easily seen that as n becomes large, for a typical $v \in \mathbb{T}_n$, the sum $K(v)$ has order 1. In fact, the marginal law of each random variable $K(v)$ for $v \in \mathbb{T}_n$ is equal in distribution to the weighted sum

$$K_n := \sum_{i=0}^{n} q^i W_i, \tag{6}$$

where the W_i are i.i.d. standard exponential random variables. The collection $\{K_n\}_{n\in\mathbb{N}}$ forms a perpetuity sequence with almost sure limit

$$K_\infty := \sum_{i=0}^{\infty} q^i W_i, \tag{7}$$

being the solution to

$$K_\infty \stackrel{d}{=} qK_\infty + W, \quad K_\infty \text{ independent from } W, \tag{8}$$

where W is a standard exponential random variable. Random variables of this type were studied in the literature [11, 21] with a heavy emphasis on the right tail behaviour $\mathbb{P}(K_\infty > t)$ as $t \to \infty$. A careful and delicate analysis of the upper and lower tails of K_∞ will play an important role in our study of the asymptotics of the largest and smallest fragments of the process.

We conclude this section on the representation with a branching random walk emphasizing the scaling on which the process may be viewed. Indeed, consider the interval $[0, k^{-n})$—a representative of the typical interval of size k^{-n}—which exists for a random period of time during the process. This random period of time is equal in law to

$$[q^{-(n-1)} K_{n-1}, q^{-(n-1)} K_{n-1} + q^{-n} W),$$

where W is a standard exponential random variable and K_n is given by (6) (so that in particular, K_n has unit order when n is large). In particular, loosely speaking we have:

The times t for which the intervals of size k^{-n} exist in the process are of order $q^{-(1+o(1))n}$.

Inverting this relation gives:

The intervals of sizes k^{-n} existing at a time t have the order $n = \big(1 + o(1)\big)\kappa \log t$,

where, as in the Introduction, $\kappa = 1/\log(1/q)$. In particular, this discussion sketches the first order scale on which the process lives: the typical interval at time t has size $k^{-(1+o(1))\kappa \log t} = t^{-1/\alpha}$, where $\alpha > 0$ is the index of self-similarity. Note that Theorem A and Theorem B state that indeed, up to first order, *every* interval has this size.

2.2. *Largest fragments in the process.* Recall Theorem A in the Introduction, which stated that if k^{-m_t} is the size of the largest fragment in the process at time t, then with high probability for all large times t, m_t is one of the integers neighbouring the quantity

$$\kappa \log t - \kappa \log\log t - \kappa \log(\gamma\kappa).$$

In fact, this is explained by a far more descriptive result, which we now elucidate from the branching random walk perspective. Given an element v of $\mathbb{T}_n$, for each $0 \le i \le n$, let v_i be the unique ancestor of v in generation i, that is, in $\mathbb{T}_i$. One key property of the process $(K(v))_{v\in\mathbb{T}}$ is that the majority of mass in each quantity $K(v)$ is due to recent ancestors. Indeed, we have the representation

$$K(v) = \sum_{i=0}^{|v|} q^{|v|-i} W^{(v_i)}, \tag{9}$$

so that most of the mass in $K(v)$ is due to recent ancestors of v: those terms $q^{|v|-i}W^{(v_i)}$ where i is close to $|v|$. Intuitively, this implies that to a large extent, the random variables $(K(v) : v \in \mathbb{T}_n)$ are asymptotically independent. We note for further reference that (9) implies that for $m < n$, with v_m denoting the ancestor of v in generation m,

$$K(v) = q^{n-m} K(v_m) + \widetilde{K}_{n-m+1}, \tag{10}$$

where $\widetilde{K}_{n-m+1}$ is independent of $K(v_m)$ and has the same law as K_{n-m+1} defined in (6). Moreover, we show in Lemma 3.1 that the upper tails of the K_n take the form

$$\mathbb{P}(K_n > s) = \big(1 + o(1)\big)e^{-s}/\varphi_n(q) \quad \text{for large } s,$$

where

$$\varphi_n(q) := \prod_{j=1}^{n} (1 - q^j), \quad n \ge 1, \qquad \varphi_0(q) := 1. \tag{11}$$

In particular, the maximal elements of the collection $(K(v))_{v\in\mathbb{T}_n}$ behave a lot like the maximum of k^n independent random variables with exponential tails: namely, like a Gumbel random variable. We mention in passing that $\varphi_n(q)$ is a decreasing function of n, and that as $n \to \infty$, $\varphi_n(q)$ converges to the Euler function $\varphi_\infty(q) := \prod_{i=1}^{\infty}(1 - q^i)$, which takes strictly positive values for $q \in (0, 1)$. This may be seen, for instance, from Eulers pentagonal number theorem, see [1], or from the well-known fact that for $0 < a_i < 1$, we have $\Pi(1 - a_i) > 0$ if and only if $\sum a_i < \infty$.

Let N_n be the point process on the real line given by

$$N_n = \sum_{v\in\mathbb{T}_n} \delta_{J(v)}, \qquad J(v) := K(v) - \gamma|v|, \tag{12}$$

where we recall that $\gamma = \log k$. Our main result states that the elements of $(J(v))_{v\in\mathbb{T}_n}$ behave like a Poisson point process on the real line.

THEOREM C. *As $n \to \infty$, the point process N_n converges in distribution* (*in the sense of vague convergence from* [19]) *to a Poisson point process with intensity measure*

$$e^{-s}\,\mathrm{d}s/\varphi_\infty(q). \tag{13}$$

Moreover, the neighbouring point processes are asymptotically independent, in the sense that for any $\ell \geq 1$, the vector of point processes $(N_n, \dots, N_{n+\ell-1})$ converges in distribution to a vector of ℓ independent Poisson processes with intensity given in (13).

Let us now consider the large fragments. One immediate consequence of Theorem C is the following result on the asymptotic behaviour of

$$K_n^{\max} := \max\{K(v) : v \in \mathbb{T}_n\} \quad \text{and} \quad \tau_n := K_n^{\max} - \gamma n = \max\{J(v) : v \in \mathbb{T}_n\}. \tag{14}$$

COROLLARY 2.1. *Let τ_n be defined as in* (14). *Then, as $n \to \infty$, τ_n converges in distribution to a shifted Gumbel random variable, that is,*

$$\lim_{n\to\infty} \mathbb{P}(\tau_n \leq s) = \exp\bigl(-e^{-s}/\varphi_\infty(q)\bigr).$$

Corollary 2.1 explains the concentration of the size of the largest fragment given in Theorem A. Note that we may write

$$\{m_t \leq n\} = \{q^{-n}K_n^{\max} > t\}. \tag{15}$$

In particular, using the definition of τ_n in (14), we have

$$\mathbb{P}(m_t \leq n) = \mathbb{P}\bigl(\tau_n > q^n t - \gamma n\bigr). \tag{16}$$

Now, since τ_n converges in distribution, if $q^n t - \gamma n \to \infty$ for $n \to \infty$, the probability on the right-hand side of (16) goes to 0, and if $q^n t - \gamma n \to -\infty$ for $n \to \infty$, the probability on the right-hand side of (16) goes to 1. In order to give a proof of the almost-sure statement Theorem A we will need some uniform estimates for τ_n which we will develop in the sequel. The full proof of Theorem A is given at the beginning of Section 4.

2.3. *Smallest fragments in the process.* We saw in Section 2.2 that the behaviour of the largest fragments in the k-regular self-similar fragmentation process is intimately connected with the largest values $K_n^{\max} := \max_{v\in\mathbb{T}_n} K(v)$ in the rescaling of the expanding branching random walk. Analogously, it is the behaviour of the smallest value $K_n^{\min} := \min_{v\in\mathbb{T}_n} K(v)$ that ultimately dictates the asymptotics of the smallest fragments in the fragmentation process. In this direction we have the following result.

THEOREM 2.2. *Define $K_n^{\min} := \min\{K(v) : v \in \mathbb{T}_n\}$ and define $w_n = w_n(\kappa, \gamma)$ by*

$$w_n := \sqrt{\frac{2\gamma}{\kappa}n} - \frac{1}{2}\log n - \frac{1}{2\kappa} - \frac{1}{2}\log\kappa + 1 - \frac{1}{2}\log(2\gamma). \tag{17}$$

Then there exists almost surely an n_0 in $\mathbb{N}$ such that for all $n \geq n_0$ we have

$$\log K_n^{\min} \in \left[-w_n - \frac{1}{n^{1/3}}, -w_n + \frac{1}{n^{1/3}}\right].$$

In Section 6 we prove Theorem 2.2, and thereafter use Theorem 2.2 to prove Theorem B.

One of the key tasks in proving Theorem 2.2 is a careful analysis of the $s \downarrow 0$ asymptotics of the left tails $\mathbb{P}(K_\infty \leq s)$ of the random variable K_∞ given in (7). Indeed, we note that since

K_n (defined in (6)) is stochastically dominated by K_∞, and K_n is a sum of $n+1$ independent exponentials, for any n we have

$$\mathbb{P}(K_\infty \le s) \le \mathbb{P}(K_n \le s) \le C(q,n)s^{n+1}$$

for some $C(q,n)$ independent of s. In particular, as $s \downarrow 0$, the probability $\mathbb{P}(K_\infty < s)$ goes to zero faster than any power of s. The following result, which we believe to be of independent interest, gives a fine characterisation of these fast asymptotics.

THEOREM 2.3. *There exists a constant C_q such that for all $s \in (0, 1/e^2]$ and for all $n \ge \kappa(\log \frac{1}{s} + \log\log \frac{1}{s})$, including possibly $n = \infty$, we have*

$$\frac{1}{C_q}\exp\bigl(-F_q(s)\bigr) \le \mathbb{P}(K_n \le s) \le C_q \exp\bigl(-F_q(s)\bigr), \tag{18}$$

where

$$F_q(s) := \frac{\kappa}{2}\left(\log\frac{1}{s} + \log\log\frac{1}{s} + \frac{1}{2\kappa} + \log\kappa - 1\right)^2 + \left(\frac{1}{2}+\kappa\right)\log\log\frac{1}{s}. \tag{19}$$

Theorem 2.3 is proven in Section 5. We remark that the restriction $s \le 1/e^2$ ensures $\log\log\frac{1}{s} > 0$. Let us also note from Theorem 2.3 that for fixed s, provided n is sufficiently large compared to $1/s$, the left tail $\mathbb{P}(K_n \le s)$ takes the same order as $\mathbb{P}(K_\infty \le s)$.

That completes the section on statements of our main results. In the next section we begin setting the foundations for proofs of these statements by looking at formulas surrounding the random variables K_n and the associated Markov chains. Thereafter we provide a simple lemma suitable for converting statements about the expanding branching random walk to those about the fragmentation process.

3. Preliminaries on the rescaled expanding branching random walk.

Throughout the rest of this paper, $C_\Omega \in (0,\infty)$ is a constant which is not of particular interest, and which may vary from line to line, but depends only on the set of parameters $\Omega \subseteq \{q, \vec{p}, k, t_0\}$ (with parameters $\vec{p}$ and t_0 yet to be defined). We stress that constants C_Ω do not depend on $n, m \in \mathbb{N}$ and $t > 0$.

3.1. *Transition probabilities of birth processes.* Recall that k^{-X_t} denotes the length of the interval containing 0 present in the system at time t. As noted in Section 2.1 the moment of the nth splitting of this interval, $S_n = \sup\{t \ge 0 : X_t = n\}$ has an explicit representation, see (4). Using (4) one can compute directly (see, for instance, Feller [12], I.13 Problem 12) that

$$\mathbb{P}(S_n \in \mathrm{d}t) = \left(\prod_{i=0}^{n}\lambda_i\right)\sum_{j=0}^{n}\frac{e^{-\lambda_j t}}{\prod_{0\le k\le n, k\ne j}(\lambda_k - \lambda_j)}\,\mathrm{d}t, \tag{20}$$

where $\lambda_i = q^i$. Integrating both sides of (20), we obtain

$$\mathbb{P}(S_n > t) = \left(\prod_{i=0}^{n}\lambda_i\right)\sum_{j=0}^{n}\frac{e^{-\lambda_j t}}{\lambda_j\prod_{0\le k\le n, k\ne j}(\lambda_k - \lambda_j)}. \tag{21}$$

Consider now calculating $\mathbb{P}(X_t = n) = \mathbb{P}(S_n > t, S_{n-1} \le t)$. We claim that

$$\mathbb{P}(X_t = n) = \left(\prod_{i=0}^{n-1}\lambda_i\right)\sum_{j=0}^{n}\frac{e^{-\lambda_j t}}{\prod_{0\le k\le n, k\ne j}(\lambda_k - \lambda_j)}. \tag{22}$$

The most natural way to prove (22) is by writing $\mathbb{P}(X_t = n) = \mathbb{P}(S_n > t) - \mathbb{P}(S_{n-1} > t)$, and then applying (21). However, there is a far slicker route, writing

$$\mathbb{P}\big(S_n \in [t, t+h)\big) = \mathbb{P}(X_t \le n, X_{t+h} > n) = \mathbb{P}(X_t = n, X_{t+h} = n+1) + o(h).$$

Using the Markov property results in

$$\mathbb{P}(S_n \in \mathrm{d}t) = \lambda_n \mathbb{P}(X_t = n)\,\mathrm{d}t. \tag{23}$$

In particular, by using (23) and (20), we immediately obtain (22). With a view towards tackling the equations (20), (21) and (22) with $\lambda_i = q^i$, recall the definition (11) of $\varphi_n(q)$, and note that for any $0 \le j \le n$,

$$\prod_{0 \le k \le n, k \ne j} (q^k - q^j) = (-1)^{n-j} q^{j(n-j/2-1/2)} \varphi_j(q)\varphi_{n-j}(q). \tag{24}$$

By replacing j with $n - j$, and using (24), (21) and $\lambda_i = q^i$ we have

$$\mathbb{P}(S_n > t) = \sum_{j=0}^{n} \frac{(-1)^j q^{j(j+1)/2}}{\varphi_j(q)\varphi_{n-j}(q)} \exp(-q^{n-j} t).$$

Recall that K_n is given in (6), and is equal in distribution to $q^n S_n$. Thus, we have

$$\mathbb{P}(K_n > t) = \sum_{j=0}^{n} \frac{(-1)^j q^{j(j+1)/2}}{\varphi_j(q)\varphi_{n-j}(q)} \exp(-q^{-j} t). \tag{25}$$

By differentiating both sides of (25) with respect to t, we see that the density f_n of K_n is given by

$$f_n(t) = \sum_{j=0}^{n} \frac{(-1)^j q^{j(j-1)/2}}{\varphi_j(q)\varphi_{n-j}(q)} \exp(-q^{-j} t). \tag{26}$$

From (6) it is plain that $K_n \le K_{n+1}$, and that almost surely, as $n \to \infty$, the random variables $(K_n)_{n\in\mathbb{N}}$ converge to a finite limit K_∞, which is given by (7). It is straightforward to verify, using the monotone convergence theorem and (25), that

$$\mathbb{P}(K_\infty > t) = \frac{1}{\varphi_\infty(q)} \sum_{j=0}^{\infty} \frac{(-1)^j q^{j(j+1)/2}}{\varphi_j(q)} \exp(-q^{-j} t). \tag{27}$$

That the right-hand side of (27) is equal to 1 when $t = 0$ is a consequence of the identity

$$\sum_{j=0}^{\infty} \frac{\zeta^j q^{j(j+1)/2}}{\varphi_j(q)} = \prod_{i=1}^{\infty} (1 + q^i \zeta), \qquad \zeta \in \mathbb{R},$$

which is a well known fact in q-combinatorics; see for instance Exercise 4 in Section I.2 of Macdonald [20]. Using (27), we can control the second order term in the asymptotic expansion of the right tail of K_∞, which will be useful in the sequel.

LEMMA 3.1. *For every $t \ge 0$, we have the following tail and density bounds for K_n, $n \in \mathbb{N} \cup \{0\} \cup \{\infty\}$:*

$$\left| \mathbb{P}(K_n > t) - \frac{e^{-t}}{\varphi_n(q)} \right| \le C_q e^{-t/q}$$

and

$$\left| f_n(t) - \frac{e^{-t}}{\varphi_n(q)} \right| \le C_q e^{-t/q}.$$

PROOF. Recall (25). We use the triangle inequality, and the facts that $\varphi_j(q)$ is decreasing in j, and that we have $q<1$, to see that

$$\begin{aligned}\left|\mathbb{P}(K_n>t)-\frac{e^{-t}}{\varphi_n(q)}\right| &= \left|\sum_{j=1}^{n}\frac{(-1)^j q^{j(j+1)/2}}{\varphi_j(q)\varphi_{n-j}(q)}\exp(-q^{-j}t)\right| \\ &\leq \sum_{j=1}^{n}\frac{q^j}{\varphi_n(q)^2}\exp(-t/q).\end{aligned}$$

Now note that $\varphi_n(q)\geq\varphi_\infty(q)$, which gives the first claim. A similar argument yields the second claim. □

We will also find occasion to use the crude bounds

$$f_n(t)\leq C'_q e^{-t} \quad \text{and} \quad \mathbb{P}(K_n>t)\leq C'_q e^{-t}, \quad t\geq 0, \tag{28}$$

both of which are direct consequences of Lemma 3.1. Note that the latter bound can be significantly improved, as we will see in Section 5 when proving Theorem 2.3.

3.2. *From the branching random walk back to the fragmentation process.* In this brief section, we give a basic lemma for bounding values of increasing functions $f:\mathbb{R}_+\to\mathbb{N}$ in terms of the times at which they jump. This allows us to convert the results on the rescaled expanding branching random walk to statements about the fragmentation process, see Theorem A and Theorem B. The proof follows from a standard computation and will therefore be omitted.

LEMMA 3.2. *Let $t_0\in\mathbb{R}_+$ and let $n_0\in\mathbb{N}$. Suppose $f:[t_0,\infty)\to\{n_0,n_0+1,n_0+2,\dots\}$ is a surjective and nondecreasing right-continuous function. For each $n\geq n_0$ define*

$$T_n:=\sup\{t\geq t_0: f(t)=n\}$$

to be the point in $[t_0,\infty)$ at which $f(t)$ jumps from n to $n+1$. Suppose $a,b:[n_0,\infty)\to\mathbb{R}_+$ are two strictly increasing continuous functions such that $a(s),b(s)\to\infty$ with $s\to\infty$, and for each $n\geq n_0$

$$a(n)\leq T_n\leq b(n).$$

Then for all $t\in[t_0,\infty)$ large enough, we have

$$\lceil b^{-1}(t)\rceil\leq f(t)\leq\lceil a^{-1}(t)\rceil$$

where a^{-1} and b^{-1} are the respective inverse functions of a and b.

4. The rightmost particles in the rescaled expanding branching random walk. In this section, we study the rightmost particles in the rescaled expanding branching random walk, which are connected to the largest fragments in the fragmentation process. We begin in Section 4.1 with a proof of Theorem A concerning the concentration in law of the size of the largest fragment at large times. In the remainder of Section 4, we study point processes associated with the largest particles in the rescaled branching random walk, ultimately proving Theorem C.

4.1. *Proof of Theorem A.* We recall from Section 2 that we associate with the fragmentation an expanding branching random walk $\{S(v) : v \in \mathbb{T}\}$: the elements v in the nth generation $\mathbb{T}_n$ of $\mathbb{T}$ correspond to the intervals of size k^{-n}, with $S(v)$ denoting the time at which the interval v fragments. In particular, the quantity

$$\max_{v \in \mathbb{T}_n} S(v) = \sup\{t \geq 0 : m_t \leq n\}$$

is the last time at which there is an interval of size k^{-n} in the process. We recall further that $K(v) := q^{|v|} S(v)$ denotes the rescaling of the expanding BRW, and that $\tau_n := \max_{v \in \mathbb{T}_n} K(v) - \gamma n$. In particular, up to scaling and translation, the behaviour of τ_n dictates that of the maximal fragment.

We now obtain upper bounds on both the upper and lower tails of τ_n. Considering first the upper tail, by using the union bound to obtain the first inequality below, and then the tail bound (28) on K_n to obtain the second, we have

$$\mathbb{P}(\tau_n > s) = \mathbb{P}\Big(\max_{|v|=n} J(v) > s\Big) \leq k^n \mathbb{P}(K_n > s + \gamma n) \leq C_q e^{-s},$$

for all $n \geq 0$ and $s \in \mathbb{R}$. In particular, $\mathbb{P}(\tau_n > 2 \log n) \leq C_q / n^2$ is summable in n, so that by the Borel–Cantelli lemma

$$\mathbb{P}(\tau_n \leq 2 \log n \text{ for all but finitely many } n) = 1. \tag{29}$$

On the other hand, by the construction of the rescaled expanding branching random walk (see (9)), $\max_{|v|=n} J(v)$ stochastically dominates $\max_{1 \leq j \leq k^n}(W_j - \gamma n)$, where $W_1, \dots, W_{k^n}$ are i.i.d. standard exponential random variables. Hence,

$$\begin{aligned}\mathbb{P}(\tau_n < s) &= \mathbb{P}\Big(\max_{|v|=n} J(v) < s\Big) \leq \big(1 - \mathbb{P}(W_1 > s + \gamma n)\big)^{k^n} \leq \exp\big(-k^n \mathbb{P}(W_1 > s + \gamma n)\big) \\ &= \exp\big(-e^{-s}\big).\end{aligned}$$

In particular $\mathbb{P}(\tau_n < -\log(2 \log n)) \leq 1/n^2$ is summable in n, so that again by Borel–Cantelli we have

$$\mathbb{P}\big(\tau_n \geq -\log(2 \log n) \text{ for all but finitely many } n\big) = 1. \tag{30}$$

To summarise, from (29) and (30) we have seen that almost surely

$$-\log(2 \log n) \leq \tau_n \leq 2 \log n \quad \text{for all but finitely many } n. \tag{31}$$

Let $T_n := \max_{|v|=n} S(v)$ denote the last time at which there was an interval of size k^{-n}, so that $T_n = q^{-n}(\tau_n + \gamma n)$. Rephrasing (31) we have, almost surely,

$$q^{-n}\big(\gamma n - \log(2 \log n)\big) \leq T_n \leq q^{-n}(\gamma n + 2 \log n) \quad \text{for all but finitely many } n. \tag{32}$$

Note that by definition $m_t := \sup\{t \geq 0 : T_n \leq t\}$. Moreover, for $a(x) := q^{-x}(\gamma x - \log(2 \log x))$ and $b(x) := q^{-x}(\gamma x + 2 \log x)$, we are in the setting of Lemma 3.2, so that almost surely there exists a $t_0 \in \mathbb{R}$ such that for all $t \geq t_0$,

$$\lceil b^{-1}(t) \rceil \leq m_t \leq \lceil a^{-1}(t) \rceil. \tag{33}$$

It remains to obtain explicit functions from $b^{-1}(t)$ and $a^{-1}(t)$. The reader is invited to verify using the fact that $\frac{1}{\kappa} := \log \frac{1}{q}$ that with $g(t)$ as in the statement of Theorem A, we have

$$a^{-1}(t) = g(t) + \kappa \frac{\log \log t}{\log t} + o\Big(\frac{\log \log t}{\log t}\Big)$$

and

$$b^{-1}(t) = g(t) + \left(\kappa - \frac{2}{\gamma}\right)\frac{\log\log t}{\log t} + o\left(\frac{\log\log t}{\log t}\right).$$

Setting $\mu_1 = (\kappa + \frac{2}{\gamma})$, for all sufficiently large t we have

$$g(t) - \mu_1 \frac{\log\log t}{\log t} \le b^{-1}(t) \le a^{-1}(t) \le g(t) + \mu_1 \frac{\log\log t}{\log t}. \tag{34}$$

In particular, combining (33) and (34) we see that there exists t_1 such that for all $t \ge t_1$ we have

$$\left\lceil g(t) - \mu_1 \frac{\log\log t}{\log t}\right\rceil \le m_t \le \left\lceil g(t) + \mu_1 \frac{\log\log t}{\log t}\right\rceil,$$

which is precisely the statement of Theorem A.

4.2. *The rescaled point process.* We define a sequence of point processes $(N_n)_{n\ge 1}$ on the real line as follows. The number of points $N_n(A)$ lying in a Borel set $A \subseteq \mathbb{R}$ is given by

$$N_n(A) := \sum_{v\in\mathbb{T}_n} \delta_{J(v)}(A) = \#\{v \in \mathbb{T}_n : J(v) \in A\},$$

where we recall from (12) that $J(v) = K(v) - \gamma|v|$ for $\gamma = \log k$. It follows from the linearity of expectation and Lemma 3.1 that

$$\mathbb{E}\big[N_n([t,\infty))\big] = k^n \mathbb{P}(K_n > \gamma n + t) = \big(1 + o(1)\big)e^{-t}/\varphi_\infty(q).$$

That is, as n grows, the point process N_n has a unit order number of particles in each compact interval in terms of expectations. We will prove that the point process N_n converges in distribution to a Poisson point process, denoted by N_∞ and with intensity $e^{-s}\,\mathrm{d}s/\varphi_\infty(q)$. In fact, we will in the following establish a stronger statement.

THEOREM 4.1. *Let ℓ be a positive integer. Then as $n \to \infty$, the ℓ-tuple $(N_n, \dots, N_{n+\ell-1})$ of point processes converge in distribution to a ℓ-tuple of i.i.d. Poisson point processes on the real line with intensity measures $e^{-s}\,\mathrm{d}s/\varphi_\infty(q)$.*

Theorem 4.1 is simply a reformulation of Theorem C, which was stated in Section 2. In order to prove Theorem 4.1, we will use a moment argument based on *factorial measures*, which we now introduce using some definitions from the theory of point processes following Section 4.3 of [19]. Given a point process $Y = \sum_k \delta_{x_k}$ on a set E, for every integer $p \ge 1$ we may define a new point process $Y^{[p]}$ on E^p by letting $Y^{[p]}(A)$ be the ordered p-tuples of distinct points of Y in $A \subseteq E^p$. Given a measure λ on E, we define the pth factorial measure $\lambda^{[p]}$ on E^p by setting

$$\lambda^{[p]}(A) = \mathbb{E}\big[Y^{[p]}(A)\big], \quad A \subseteq E^p.$$

If Y is a Poisson point process on E with intensity measure λ, then the pth factorial measure is simply given by the product measure $\lambda^{\otimes p}$ on E^p [19]. The following technical lemma, the proof of which we only outline based on existing literature, guarantees that convergence of moments implies convergence in distribution to a Poisson process.

LEMMA 4.2. *Suppose X is a Poisson process on E with nonatomic intensity measure λ. Let $(X_n)_{n\ge1}$ be a sequence of point processes on a set E satisfying*

$$\lim_{n\to\infty} \mathbb{E}\big[X_n^{[p]}(A)\big] = \lambda^{\otimes p}(A)$$

for every measurable subset A of E^p for which $\lambda^{\otimes p}(A)$ is finite. Then X_n converges in distribution to X.

PROOF. Using the fact that the limiting process is simple, that is, it assigns at most a unit mass to each point, we use [16], Theorem 4.18, which asserts that it is sufficient to show that the avoidance functions and intensity measures converge, that is, we have for any Borel set $A \subseteq \mathbb{R}$,

$$\mathbb{P}(X_n(A)=0) \to \mathbb{P}(X_\infty(A)=0) \quad \text{and} \quad \mathbb{E}[X_n(A)] \to \mathbb{E}[X_\infty(A)]. \tag{35}$$

The factorial measures can be used to represent the avoidance function via [10], formula 5.4.10,

$$\mathbb{P}(Y(A)=0) = \sum_{p=0}^{\infty} (-1)^p \frac{\lambda^{[p]}(A^{(p)})}{p!}, \quad A^{(p)} = \prod_{j=1}^{p} A, A \in E.$$

It follows that if $\lim_{n\to\infty} \mathbb{E}[X_n^{[p]}(A)] = \lambda^{\otimes p}(A)$ for every p, then (35) holds, establishing convergence in distribution. □

Next, suppose $\mathbf{Y} := (Y_0, \dots, Y_{\ell-1})$ is an ℓ-tuple of independent Poisson processes on $\mathbb{R}$, each with intensity $e^{-s}\,\mathrm{d}s/\varphi_\infty(q)$. Then $\mathbf{Y}$ may as well be regarded as a single Poisson process on $\{0, \dots, \ell-1\} \times \mathbb{R}$ with intensity measure $\frac{1}{\varphi_\infty(q)} \sum_{i=0}^{\ell-1} \delta_i \otimes e^{-s}\,\mathrm{d}s$. Let $\vec{p} = (p_0, \dots, p_{\ell-1})$ denote a collection of nonnegative integers and $(t_{i,j} : 0 \le i \le \ell-1, 1 \le j \le p_i)$ be real numbers. Note that sets of the form

$$E(t_{i,j}) := \prod_{i=0}^{\ell-1} \prod_{j=1}^{p_i} \{i\} \times [t_{i,j}, \infty)$$

yield a π-system generating $(\{0, \dots, \ell-1\} \times \mathbb{R})^{|\vec{p}|}$, where $|\vec{p}| := p_0 + \dots + p_{\ell-1}$. Using the fact that $\int_{t_{i,j}}^{\infty} e^{-s}\,\mathrm{d}s/\varphi_\infty(q) = e^{-t_{i,j}}/\varphi_\infty(q)$, we have

$$\begin{aligned} \mathbb{E}[\mathbf{Y}^{[|\vec{p}|]}(E(t_{i,j}))] &= \mathbb{E}\left[\prod_{i=0}^{\ell-1} Y_i^{[p_i]}([t_{i,1}, \infty) \times \dots \times [t_{i,p_i}, \infty))\right] \\ &= \prod_{i=0}^{\ell-1} \prod_{j=1}^{p_i} e^{-t_{i,j}}/\varphi_\infty(q). \end{aligned} \tag{36}$$

Here $Y_i^{[p_i]}([t_{i,1}, \infty) \times \dots \times [t_{i,p_i}, \infty))$ denotes the number of p_i-tuples $(x_1, \dots, x_{p_i})$ of distinct points of Y_i for which $x_j \ge t_{i,j}$ for every $1 \le j \le p_i$.

In light of Lemma 4.2 and (36), in order to prove Theorem 4.1 it is sufficient to show that for all nonnegative integers $p_0, \dots, p_{\ell-1}$, and all real numbers $(t_{i,j} : 0 \le i \le \ell-1, 1 \le j \le p_i)$, we have

$$\lim_{n\to\infty} \mathbb{E}\left[\prod_{i=0}^{\ell-1} N_{n+i}^{[p_i]}([t_{i,1}, \infty) \times \dots \times [t_{i,p_i}, \infty))\right] = \prod_{i=0}^{\ell-1} \prod_{j=1}^{p_i} e^{-t_{i,j}}/\varphi_\infty(q), \tag{37}$$

where, by the definition of $N_n(A)$, for each $0 \le i \le \ell-1$, $N_{n+i}^{[p_i]}([t_{i,1}, \infty) \times \dots \times [t_{i,p_i}, \infty))$ is the number of p_i-tuples of elements $(u_{i,1}, \dots, u_{i,p_i})$ in generation $n+i$ of the expanding branching random walk for which $J(u_{i,j}) > t_{i,j}$.

To this end, for $\vec{p} := (p_0, \dots, p_{\ell-1})$, we define

$$\mathbb{T}_n^{\vec{p}} := \{\mathbf{u} := (u_{i,j} : 0 \le i \le \ell-1, 1 \le j \le p_i) : (u_{i,1}, \dots, u_{i,p_i}) \in \mathbb{T}_{n+i} \text{ are distinct}\}. \tag{38}$$

From the linearity of expectation, we have

$$\mathbb{E}\left[\prod_{i=0}^{\ell-1} \prod_{j=1}^{p_i} N_{n+i}^{[p_i]}([t_{i,1}, \infty) \times \dots \times [t_{i,p_i}, \infty))\right] = \sum_{\mathbf{u} \in \mathbb{T}_n^{\vec{p}}} \mathbb{P}\left(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\}\right). \tag{39}$$

We split the task of proving (37) over the next two sections, first dealing with the easier lower bound, and then with the more difficult upper bound.

4.3. *The lower bound in* (37). This section is dedicated to proving the lower bound

$$\liminf_{n\to\infty}\mathbb{E}\left[\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} N_{n+i}^{[p_i]}\big([t_{i,1},\infty)\times\cdots\times[t_{i,p_i},\infty)\big)\right]\ge\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} e^{-t_{i,j}}/\varphi_\infty(q). \tag{40}$$

We begin with a lemma estimating the cardinality of $\mathbb{T}_n^{\vec{p}}$.

LEMMA 4.3. *There is a constant $C_{\vec{p}}\in(0,\infty)$ depending on $\vec{p}=(p_0,\dots,p_{\ell-1})$, but independent of n, such that*

$$\prod_{i=0}^{\ell-1}k^{p_i(n+i)}\ge\#\mathbb{T}_n^{\vec{p}}\ge\big(1-C_{\vec{p}}k^{-n}\big)\prod_{i=0}^{\ell-1}k^{p_i(n+i)}.$$

PROOF. Consider that

$$\begin{aligned}\#\mathbb{T}_n^{\vec{p}}&=\prod_{i=0}^{\ell-1}k^{n+i}\big(k^{n+i}-1\big)\cdots\big(k^{n+i}-p_i+1\big)\\&=\prod_{i=0}^{\ell-1}k^{p_i(n+i)}\big(\big(1-1/k^{n+i}\big)\cdots\big(1-(p_i-1)/k^{n+i}\big)\big).\end{aligned}$$

The upper bound is now trivial. The lower bound follows from noting that

$$\prod_{i=0}^{\ell-1}\big(\big(1-1/k^{n+i}\big)\cdots\big(1-(p_i-1)/k^{n+i}\big)\big)\ge1-\sum_{i=0}^{\ell-1}\sum_{j=1}^{p_i-1}j/k^{n+i}\ge1-k^{-n}\sum_{i=0}^{\ell-1}\binom{p_i}{2},$$

so that we may take $C_{\vec{p}}=\sum_{i=0}^{\ell-1}\binom{p_i}{2}$. □

The following lemma is an FKG-type inequality for correlated events on the tree. Since the proof is not related to the rest of our arguments, it will be given in the Appendix. (We remark that our proof of Lemma 4.4 uses the upcoming equation (51); the derivation of this equation is independent of the rest of the paper.)

LEMMA 4.4. *For any $n\in\mathbb{N}$, $\vec{p}=(p_0,\dots,p_{\ell-1})$ and any tuple $\mathbf{u}:=(u_{i,j}:0\le i\le\ell-1,1\le j\le p_i)$, we have*

$$\mathbb{P}\left(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{J(u_{i,j})>t_{i,j}\}\right)\ge\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\mathbb{P}\big(J(u_{i,j})>t_{i,j}\big) \tag{41}$$

for any choice of real numbers $(t_{i,j})$.

With Lemma 4.4 at hand, we are now ready to prove the lower bound (40).

PROOF OF (40). By Lemma 3.1, recalling $J(v)=K(v)-\gamma|v|$ from (12), we have for all i, j that

$$\mathbb{P}\big(J(u_{i,j})>t_{i,j}\big)\ge\frac{e^{-t_{i,j}-\gamma(n+i)}}{\varphi_{n+i}(q)}\big(1-C_qe^{-(q^{-1}-1)(\gamma n+t_0)}\big), \tag{42}$$

where $t_0 := \min_{i,j}\{t_{i,j}\}$ and $C_q \in (0,\infty)$ is some constant depending on q, but independent of n and $(t_{i,j})$. Now, for a sufficiently large constant $C_{\vec{p},q,t_0} \in (0,\infty)$ depending on q, t_0 and $\vec{p} = (p_0,\dots,p_{\ell-1})$, but again independent of n and $(t_{i,j})$, we have that by the Bernoulli inequality for sufficiently large n,

$$\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\big(1 - C_q e^{-(q^{-1}-1)(\gamma n+t_0)}\big) = \big(1 - C_q e^{-(q^{-1}-1)(\gamma n+t_0)}\big)^{|\vec{p}|} \geq 1 - C_{q,\vec{p},t_0} e^{-(q^{-1}-1)\gamma n}, \tag{43}$$

where $|\vec{p}| = p_0 + p_1 + \cdots + p_{\ell-1}$. Combining (41), (42) and (43), for any tuple $(u_{i,j})$ in $\mathbb{T}_n^{\vec{p}}$, we have

$$\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{J(u_{i,j}) > t_{i,j}\}\Bigg) \geq \big(1 - C_{q,\vec{p},t_0} e^{-(q^{-1}-1)\gamma n}\big)\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} e^{-t_{i,j}-\gamma(n+i)}/\varphi_{n+i}(q).$$

Plugging this in the probability summation formula (39), we obtain

$$\mathbb{E}\Bigg[\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} N_{n+i}^{[p_i]}\big([t_{i,1},\infty)\times\cdots\times[t_{i,p_i},\infty)\big)\Bigg] \geq \#\mathbb{T}_n^{\vec{p}}\big(1 - C_{q,\vec{p},t_0} e^{-(q^{-1}-1)\gamma n}\big)\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} e^{-t_{i,j}-\gamma(n+i)}/\varphi_{n+i}(q).$$

Now using the fact that $\gamma = \log k$, and the lower bound in Lemma 4.3, we find that

$$\#\mathbb{T}_n^{\vec{p}}\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} e^{-\gamma(n+i)} \geq \big(1 - C_{\vec{p}} k^{-n}\big).$$

Combining the last two estimates we obtain

$$\mathbb{E}\Bigg[\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} N_{n+i}^{[p_i]}\big([t_{i,1},\infty)\times\cdots\times[t_{i,p_i},\infty)\big)\Bigg] \geq \big(1 - C_{\vec{p}} k^{-n}\big)\big(1 - C_{q,\vec{p},t_0} e^{-(q^{-1}-1)\gamma n}\big)\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} e^{-t_{i,j}}/\varphi_{n+i}(q).$$

Taking $n \to \infty$ concludes the proof of (40). $\square$

4.4. *The hard direction in Theorem C: An overview.* In this section, we work towards proving the upper bound in (37). Namely, the goal is to show that

$$\limsup_{n\to\infty}\mathbb{E}\Bigg[\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} N_{n+i}^{[p_i]}\big([t_{i,1},\infty)\times\cdots\times[t_{i,p_i},\infty)\big)\Bigg] \leq \prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i} e^{-t_{i,j}}/\varphi_\infty(q). \tag{44}$$

In light of the probability summation formula (39), to tackle the hard direction, we need to obtain effective upper bounds on the *exceedance probabilities*

$$\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{J(u_{i,j}) > t_{i,j}\}\Bigg)$$

for tuples $(u_{i,j} : 0 \leq i \leq \ell-1, 1 \leq j \leq p_i)$ with $u_{i,j} \in \mathbb{T}_{n+i}^{\vec{p}}$. We now overview the main idea in proving an inequality of the form (44). Given integers $0 \leq m \leq n$ and a tuple $\mathbf{u} = (u_{i,j} :$

$0 \le i \le \ell - 1, 1 \le j \le p_i$), we define the number $1 \le P_{n-m}(\mathbf{u}) \le |\vec{p}| = p_0 + \cdots + p_{\ell-1}$ by setting

(45) $\quad P_{n-m}(\mathbf{u}) :=$ number of different ancestors in generation $n-m$ of the vertices $u_{i,j}$.

For a special choice of m which we give below, we distinguish between two different types of tuples:

- We say a tuple $\mathbf{u}$ in $\mathbb{T}_n^{\vec{p}}$ is *distantly related* (in generation $n-m$) if
$$P_{n-m}(\mathbf{u}) = p_0 + \cdots + p_{\ell-1} = |\vec{p}|.$$
We will see that provided $m \le \theta n$ for some constant $\theta < 1$, the overwhelming number of tuples in $\mathbb{T}_n^{\vec{p}}$ are distantly related as m and n become large. Since the bulk of particles are of this form, we will require a fairly delicate m-dependent control on the exceedance probabilities; see Lemma 4.5 below.
- For an integer $\nu < |\vec{p}|$, we say $\mathbf{u}$ in $\mathbb{T}_n^{\vec{p}}$ is ν-*closely related* (in generation $n-m$) if
$$P_{n-m}(\mathbf{u}) = \nu.$$
We find that for such tuples, the exceedance probability $\mathbb{P}(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\})$ has a larger order than for distantly related tuples. However, it turns out that this order is negligible when compared with the relative size of the number of closely related tuples. Indeed, we show in Lemma 4.7 that the number of tuples $\mathbf{u}$ with $P_{n-m}(\mathbf{u}) = \nu$ has the order $k^{\nu n}$, while Lemma 4.6 tells us that the associated exceedance probabilities are of order $o(k^{-\nu n})$.

The next two lemmas are the main results of this section, controlling respectively, the exceedance probabilities associated with distantly and closely related tuples.

LEMMA 4.5. *Let $m \in \mathbb{N}$ such that $|\vec{p}|q^{m-1} \le 1/2$, and suppose that $P_{n-m}(\mathbf{u}) = |\vec{p}|$. Then*

$$\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\}\Bigg) \le (1 + \varepsilon_{n,m}) \prod_{i=0}^{\ell-1} \prod_{j=1}^{p_i} \big(e^{-\gamma(n+i)-t_{i,j}}/\varphi_\infty(q)\big), \tag{46}$$

for all real numbers $t_{i,j}$, where

$$\varepsilon_{n,m} = C_{\vec{p},q,t_0}\big(e^{-\frac{1}{q}\gamma n} + q^m\big)$$

for $t_0 := \min_{i,j}\{t_{i,j}\}$.

LEMMA 4.6. *Let $m \in \mathbb{N}$ be sufficiently large so that $2|\vec{p}|q^m < 1/2$. Let $t_0 := \min_{i,j}\{t_{i,j}\}$ for real numbers $t_{i,j}$, and $\mathbf{u}$ in $\mathbb{T}_n^{\vec{p}}$ be such that $P_{n-m}(\mathbf{u}) = \nu < |\vec{p}|$. Then*

$$\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\}\Bigg) = o\big(k^{-\nu n}\big).$$

More specifically, there is a constant $C_{q,t_0} \in (0, \infty)$ such that

$$\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\}\Bigg) \le C_{q,t_0} \exp\big(-(\theta_q + \nu)\gamma n\big),$$

where $\theta_q := \min\{\frac{1}{q}, 2-q\} - 1 > 0$.

The proofs of Lemma 4.5 and Lemma 4.6 are both lengthy, and we defer them to Sections 4.5 and 4.6, respectively. We now conclude this overview with the following short lemma on the number of closely related tuples, which will be used in conjunction with Lemma 4.5 and Lemma 4.6 to prove the upper bound (44).

LEMMA 4.7. *Recall* (38) *and* (45). *We have the following bound on the number of* ν-*closely related tuples in generation* $n-m$:

$$\#\{\mathbf{u}\in\mathbb{T}_n^{\vec{p}}:P_{n-m}(\mathbf{u})=\nu\}\le C_{k,\ell,\vec{p}}k^{n\nu}k^{(|\vec{p}|-\nu)m}.$$

PROOF. There are at most $k^{(n-m)\nu}$ ways of choosing ν different ancestors in generation $n-m$. Each individual in generation $n-m$ has k^{m+i} descendents in generation $n+i$. Using the crude bound that $k^{m+i}\le k^{m+\ell-1}$ whenever $i\le\ell-1$, the number of ν-closely related tuples in generation $n-m$ is bounded from above by

$$k^{(n-m)\nu}\cdot(k^{m+\ell-1})^{|\vec{p}|}=C_{k,\ell,\vec{p}}k^{\nu n}k^{(|\vec{p}|-\nu)m},$$

where $C_{k,\ell,\vec{p}}:=k^{|\vec{p}|(\ell-1)}$. $\square$

We now show how Lemma 4.5, Lemma 4.6 and Lemma 4.7 are combined to obtain (44).

PROOF OF (44) ASSUMING LEMMA 4.5 AND LEMMA 4.6. By the probability summation formula (39), for any m we have

$$\begin{aligned}
&\mathbb{E}\Bigg[\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}N_{n+i}^{[p_i]}([t_{i,1},\infty)\times\cdots\times[t_{i,p_i},\infty))\Bigg]\\
&\quad=\sum_{\mathbf{u}:P_{n-m}(\mathbf{u})=|\vec{p}|}\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{J(u_{i,j})>t_{i,j}\}\Bigg)\\
&\qquad+\sum_{\nu=1}^{|\vec{p}|-1}\sum_{\mathbf{u}:P_{n-m}(\mathbf{u})=\nu}\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{J(u_{i,j})>t_{i,j}\}\Bigg).
\end{aligned}\tag{47}$$

We begin by controlling the contribution from distantly related tuples. Since there are at most $\prod_{i=0}^{\ell-1}k^{p_i(n+i)}$ elements in $\mathbb{T}_n^{\vec{p}}$, using Lemma 4.5, we obtain

$$\begin{aligned}
&\sum_{\mathbf{u}:P_{n-m}(\mathbf{u})=|\vec{p}|}\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{J(u_{i,j})>t_{i,j}\}\Bigg)\\
&\quad\le(1+\varepsilon_{n,m})\prod_{i=0}^{\ell-1}k^{p_i(n+i)}\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\frac{e^{-\gamma(n+i)-t_{i,j}}}{\varphi_\infty(q)}\\
&\quad=(1+\varepsilon_{n,m})\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\frac{e^{-t_{i,j}}}{\varphi_\infty(q)}\\
&\quad\le\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}(e^{-t_{i,j}}/\varphi_\infty(q))+\varepsilon_{n,m}\bigg(\frac{e^{-t_0}}{\varphi_\infty(q)}\bigg)^{-|\vec{p}|},
\end{aligned}\tag{48}$$

where $\varepsilon_{n,m}$ is as in the statement of Lemma 4.5, that is, $\varepsilon_{n,m}=C_{\vec{p},q,t_0}(e^{-\frac{1}{q}(\gamma n-t_0)}+q^m)$, and $t_0=\min_{i,j}\{t_{i,j}\}$. We now control the contribution from closely related tuples. Indeed, combining Lemma 4.6 with Lemma 4.7, provided that $2|\vec{p}|q^m < 1/2$, for each $1 \le \nu \le |\vec{p}| - 1$, we have

$$\sum_{\mathbf{u}: P_{n-m}(\mathbf{u})=\nu} \mathbb{P}\left(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\}\right) \le C_{k,\ell,\vec{p},q} k^{n\nu} k^{(|\vec{p}|-\nu)m} \exp(-\theta_q \gamma n - \nu\gamma n) \tag{49}$$
$$\le C_{k,\ell,\vec{p},q} \exp\big(\gamma\big((|\vec{p}|-1)m - \theta_q n\big)\big).$$

Now for each $n \in \mathbb{N}$, we set $m := \lceil \tilde{\theta}_q n \rceil$, where $|\vec{p}|\tilde{\theta}_q < \theta_q$, and send $n \to \infty$. Now by using the bounds (48) and (49) in (47), we obtain (44). □

This finishes the proof of the upper bound (44), and thereby completes the proof of Theorem 4.1, respectively, Theorem C. It remains to prove Lemma 4.5 and Lemma 4.6, which we do in the next two sections.

4.5. *Bounding exceedance probabilities of distantly related tuples.*

PROOF OF LEMMA 4.5. Let $P_{n-m}(\mathbf{u}) = |\vec{p}|$. For each i, j, let $v_{i,j}$ denote the ancestor of $u_{i,j}$ in generation $(n-m)$. Since $P_{n-m}(\mathbf{u}) = |\vec{p}|$, the sites $v_{i,j}$ are $|\vec{p}|$ distinct elements of generation $n-m$. Now by construction, we have

$$\mathbb{P}\left(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\}\right) = \mathbb{P}\left(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{K(u_{i,j}) > t_{i,j} + \gamma(n+i)\}\right) \tag{50}$$
$$= \mathbb{P}\left(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{q^{m+i} K(v_{i,j}) + K_{m+i}^{(i,j)} > t_{i,j} + \gamma(n+i)\}\right),$$

where the variables $\{K_{m+i}^{(i,j)} : 0 \le i \le \ell-1, 1 \le j \le p_i\}$ are independent, and each $K_{m+i}^{(i,j)}$ is distributed as K_{m+i}.

Consider the following general fact. If for a finite indexing set $\mathcal{E}$, $(A_e)_{e\in\mathcal{E}}$ are identically distributed (and possibly dependent) random variables with the same law as A, and $(B_e)_{e\in\mathcal{E}}$ are (possibly dependent but) independent of $(A_e)_{e\in\mathcal{E}}$ and A with any distributions, then we have that

$$\mathbb{P}\left(\bigcap_{e\in\mathcal{E}} \{A_e + B_e > c_e\}\right) \le \mathbb{P}\left(\bigcap_{e\in\mathcal{E}} \{A + B_e > c_e\}\right). \tag{51}$$

To see that (51) holds just note that

$$\min_{e\in\mathcal{E}} A_e + B_e - c_e \le A_{e^*} + B_{e^*} - c_{e^*} \stackrel{d}{=} A + \min_{e\in\mathcal{E}} B_e - c_e,$$

where e^* is a (random) element of $\mathcal{E}$ for which $B_{e^*} - c_{e^*} = \min_{e\in\mathcal{E}} B_e - c_e$. Using (51) in (50) with $A_{i,j} = K(v_{i,j})$ and $B_{i,j} = K_{m+i+1}^{(i,j)}$, we may replace $K(v_{i,j})$ in (50) with a single copy of K_{n-m}, so that we have the upper bound

$$\mathbb{P}\left(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\}\right) \tag{52}$$
$$\le \mathbb{P}\left(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{q^{m+i} K_{n-m} + K_{m+i}^{(i,j)} > t_{i,j} + \gamma(n+i)\}\right).$$

Finally, using the fact that K_∞ stochastically dominates K_n for each $n \in \mathbb{N}$, as well as the fact that $q^{m+i} \le q^m$, we may simplify several matters of indexing by extracting from (52) the upper bound

$$\begin{aligned}\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{J(u_{i,j}) > t_{i,j}\}\Bigg) &\le \mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{q^{m+i}K_\infty + K_\infty^{(i,j)} > t_{i,j} + \gamma(n+i)\}\Bigg)\\ &= \mathbb{E}\Bigg[\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\mathbb{P}\big(K_\infty^{(i,j)} > t_{i,j} + \gamma(n+i) - q^m K_\infty | K_\infty\big)\Bigg],\end{aligned} \tag{53}$$

where K_∞ and $K_\infty^{(i,j)}$ are independent copies of K_∞, and the last equality above follows from using the definition of conditional expectation. We now control the terms inside the product in the expectation on the right-hand side of (53). Indeed, by Lemma 3.1 for each i, j we have

$$\begin{aligned}\mathbb{P}\big(K_\infty^{(i,j)} > t_{i,j} + \gamma(n+i) - q^m K_\infty | K_\infty\big) \le{}& \frac{e^{-(t_{i,j}+\gamma(n+i)-q^m K_\infty)}}{\varphi_\infty(q)}\\ &\cdot\big(1 + C_q e^{-(1/q-1)(t_0+\gamma n - q^m K_\infty)_+}\big),\end{aligned}$$

where we recall $t_0 := \min_{i,j}\{t_{i,j}\}$, and for a real number x, we let x_+ denote the maximum of x and 0.

Now for every $n \in \mathbb{N}$ and every $C_q \in (0,\infty)$, there is a second constant $C_{q,n} \in (0,\infty)$ such that $(1 + C_q w)^n \le 1 + C_{q,n} w$ for all $w \in [0,1]$. In particular, setting $w = e^{-(1/q-1)(t_0+\gamma n-q^m K_\infty)_+}$ we have

$$\begin{aligned}&\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\mathbb{P}\big(K_\infty^{(i,j)} > t_{i,j} + \gamma(n+i) - q^m K_\infty | K_\infty\big)\\ &\quad\le \big(1 + C_{\vec{p},q,t_0} e^{-(1/q-1)(\gamma n - q^m K_\infty)}\big) e^{|\vec{p}| q^m K_\infty} \prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\big(e^{-t_{i,j}-\gamma(n+i)}/\varphi_\infty(q)\big).\end{aligned} \tag{54}$$

Plugging (54) into (53), we obtain

$$\begin{aligned}&\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{J(u_{i,j}) > t_{i,j}\}\Bigg)\\ &\quad\le \big(\mathbb{E}\big[e^{|\vec{p}|q^m K_\infty}\big] + C_{\vec{p},q,t_0} e^{-(1/q-1)\gamma n}\mathbb{E}\big[e^{(|\vec{p}|q^m + (1/q-1)q^m)K_\infty}\big]\big)\\ &\qquad\cdot\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\big(e^{-t_{i,j}-\gamma(n+i)}/\varphi_\infty(q)\big)\end{aligned} \tag{55}$$

$$\le \big(1 + C_{\vec{p},q,t_0} e^{-(1/q-1)\gamma n}\big)\mathbb{E}\big[e^{|\vec{p}|q^{m-1}K_\infty}\big]\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\big(e^{-t_{i,j}-\gamma(n+i)}/\varphi_\infty(q)\big), \tag{56}$$

where the final inequality above follows from the fact that both $|\vec{p}|q^m$ and $|\vec{p}|q^m + (1/q - 1)q^m$ are bounded from above by $|\vec{p}|q^{m-1}$. Now by (28) there is a constant $C_q \in (0,\infty)$ such that whenever $\theta \le 1/2$, we have

$$\mathbb{E}\big[e^{\theta K_\infty}\big] \le 1 + C_q\theta.$$

In particular, provided that $m \in \mathbb{N}$ is sufficiently large so that $|\vec{p}|q^{m-1} \le 1/2$, using this inequality in (55), we obtain

$$\begin{aligned}
&\mathbb{P}\Bigg(\bigcap_{i=0}^{\ell-1}\bigcap_{j=1}^{p_i}\{J(u_{i,j}) > t_{i,j}\}\Bigg) \\
&\quad \le \big(1 + C_{\vec{p},q,t_0} e^{-(1/q-1)\gamma n}\big)\big(1 + C_q|\vec{p}|q^{m-1}\big)\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\big(e^{-t_{i,j}-\gamma(n+i)}/\varphi_\infty(q)\big) \\
&\quad \le (1 + \varepsilon_{n,m})\prod_{i=0}^{\ell-1}\prod_{j=1}^{p_i}\big(e^{-t_{i,j}-\gamma(n+i)}/\varphi_\infty(q)\big),
\end{aligned}$$

where $\varepsilon_{n,m} = C_{\vec{p},q,t_0}(e^{-(1/q-1)\gamma n} + q^{m-1})$ for a sufficiently large $C_{\vec{p},q,t_0} \in (0, \infty)$. □

4.6. *Bounding exceedance probabilities of closely related tuples.*

LEMMA 4.8. *Let w and w' be distinct elements in $\mathbb{T} = \bigcup_{n\in\mathbb{N}} \mathbb{T}_n$. Then there exists a constant $C_q \in (0, \infty)$ such that for all $L > 0$,*

$$\mathbb{P}\big(K(w) > L, K\big(w'\big) > L\big) \le C_q \exp(-\lambda_q L),$$

where $\lambda_q := \min\{\frac{1}{q}, 2 - q\}$.

PROOF. Let $v = w \wedge w'$ be the most recent common ancestor of w and w', so that v is in generation n and w and w' are in generations $n + c$ and $n + c'$ respectively. Since $w \neq w'$, we have $\max\{c, c'\} \ge 1$. Without loss of generality, we can assume that $c' \ge 1$. Then, taking into account (10),

$$\mathbb{P}\big(K(w) > L, K\big(w'\big) > L\big) = \mathbb{P}\big(q^c K_n + \tilde{K}_c > L, q^{c'} K_n + \bar{K}_{c'} > L\big),$$

where $\tilde{K}_c$, $\bar{K}_{c'}$ and K_n are independent, and $\tilde{K}_c$ is distributed as K_c, $\bar{K}_{c'}$ is distributed as $K_{c'}$. Taking a rather generous bound using the facts that $q \le 1$, and that K_n is stochastically dominated by K_∞, we have

$$\mathbb{P}\big(K(w) > L, K\big(w'\big) > L\big) \le \mathbb{P}(K_\infty + \tilde{K}_\infty > L, qK_\infty + \bar{K}_\infty > L),$$

where K_∞, $\tilde{K}_\infty$ and $\bar{K}_\infty$ are i.i.d., recalling (7). By conditioning on the value of K_∞, we have

$$\mathbb{P}(K_\infty + \tilde{K}_\infty > L, qK_\infty + \bar{K}_\infty > L) = \int_0^\infty f_\infty(s)\mathbb{P}(K_\infty > L - qs)\mathbb{P}(K_\infty > L - s)\,\mathrm{d}s.$$

Due to (28) and Lemma 3.1, there is a constant $C_q \in (0, \infty)$ such that $f_\infty(s) \le C_q e^{-s}$ and $\mathbb{P}(K_\infty > M) \le C_q e^{-M_+}$, and we obtain

$$\begin{aligned}
&\mathbb{P}\big(K_\infty + K'_\infty > L, qK_\infty + K''_\infty > L\big) \\
&\quad \le C_q \int_0^\infty \exp\big(-s - (L - qs)_+ - (L - s)_+\big)\,\mathrm{d}s \\
&\quad \le C_q\bigg[\int_0^L e^{-2L+qs}\,\mathrm{d}s + \int_L^{L/q} e^{-s-(L-qs)}\,\mathrm{d}s + \int_{L/q}^\infty e^{-s}\,\mathrm{d}s\bigg] \\
&\quad \le C_q \exp\bigg(-\min\bigg\{\frac{1}{q}, 2 - q\bigg\}L\bigg)
\end{aligned}$$

for a sufficiently large constant $C_q \in (0, \infty)$. This proves the claim. □

PROOF OF LEMMA 4.6. Let $\mathbf{u} = (u_{i,j} : 0 \leq i \leq \ell - 1, 1 \leq j \leq p_i)$ be a ν-closely related tuple in generation $n - m$. Since $\nu < p_0 + \cdots + p_{\ell-1}$, by the pigeonhole principle, there exists an element v_0 of generation $n - m$ that has more than one descendent among the set $\{u_{i,j} : 0 \leq i \leq \ell - 1, 1 \leq j \leq p_i\}$. Let (w, w') be any two distinct members of the tuple $\mathbf{u}$ that are descendants of v_0. Let $\{v_1, \ldots, v_{\nu-1}\}$ be the other $\nu - 1$ ancestors of $\mathbf{u}$ in generation $n - m$, and for each $1 \leq i \leq \nu - 1$, let w_i be an element of $\mathbf{u}$ which has ancestor v_i. Define integers $s, s', s_1, \ldots, s_{\nu-1} \in \{0, 1, \ldots, \ell - 1\}$ to be the generations such that $w \in \mathbb{T}_{n+s}$, $w' \in \mathbb{T}_{n+s'}$, $w_i \in \mathbb{T}_{n+s_i}$. For $t_0 = \min\{t_{i,j}\}$, we have the simple relation

$$\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\} \subseteq \{J(w) > t_0\} \cap \{J(w') > t_0\} \cap \bigcap_{i=1}^{\nu-1} \{J(w_i) > t_0\} =: A_0 \cap \bigcap_{i=1}^{\nu-1} A_i,$$

where $A_0 := \{J(w) > t_0\} \cap \{J(w') > t_0\}$ and $A_i := \{J(w_i) > t_0\}$. In particular, we have the rather generous upper bound on the exceedance probability

$$\mathbb{P}\left(\bigcap_{i=0}^{\ell-1} \bigcap_{j=1}^{p_i} \{J(u_{i,j}) > t_{i,j}\}\right) \leq \mathbb{P}\left(A_0 \cap \bigcap_{i=1}^{\nu-1} A_i\right). \tag{57}$$

For integers N, let $\mathcal{F}_N := \sigma(J(v) : v \in \mathbb{T}_i, i \leq N)$. We note that the events $A_0, \ldots, A_{\nu-1}$ are conditionally independent given $\mathcal{F}_{n-m}$, each A_i conditionally depending only on $J(v_i) = K(v_i) - \gamma|v_i|$. In particular,

$$\mathbb{P}\left(A_0 \cap \bigcap_{i=1}^{\nu-1} A_i\right) = \mathbb{E}\left[\mathbb{P}\left(A_0 \cap \bigcap_{i=1}^{\nu-1} A_i | \mathcal{F}_{n-m}\right)\right] = \mathbb{E}\left[\prod_{i=0}^{\nu-1} \psi_i(K(v_i))\right], \tag{58}$$

where $\psi_i(x) := \mathbb{P}(A_i | K(v_i) = x)$. We now obtain effective upper bounds on the functions $\psi_i(x)$, first looking at the case $i \geq 1$, and then treating the $i = 0$ case separately.

For $i \geq 1$, using the definition of $(J(v))_{v \in \mathbb{T}}$ for the second equality below,

$$\begin{aligned}\psi_i(x) &:= \mathbb{P}(J(w_i) > t_0 | K(v_i) = x) \\ &= \mathbb{P}(K(w_i) > t_0 + \gamma(n + s_i) | K(v_i) = x) \leq \mathbb{P}(K(w_i) > L_0 | K(v_i) = x),\end{aligned}$$

where $L_0 := t_0 + \gamma n \leq t_0 + \gamma(n + s_i)$. Now, continuing this calculation, we use the definition of the rescaled expanding branching random walk $(K(v))_{v \in \mathbb{T}}$ to obtain the equality below. We take further generous bounds to obtain the following inequality in the second line below, and then the tail bound Lemma 3.1 to obtain the inequality in the third line below, yielding

$$\begin{aligned}\mathbb{P}(K(w_i) > L_0 | K(v_i) = x) &= \mathbb{P}(q^{s_i+m} x + K_{s_i+m} > L_0) \\ &\leq \mathbb{P}(K_\infty > L_0 - q^m x) \leq C_q \exp(-(L_0 - q^m x)) \frac{1}{\varphi_\infty(q)}.\end{aligned}$$

In summary, for each $1 \leq i \leq \nu - 1$ we have

$$\psi_i(x) \leq C_q \exp(-(L - q^m x)), \tag{59}$$

where $L = t_0 + \gamma n - \log \varphi_\infty(q) > L_0$. We now turn to estimating $\psi_0(x)$ using Lemma 4.8. Indeed, using Lemma 4.8 to obtain the inequality below we have, provided $L_0 > 0$,

$$\psi_0(x) = \mathbb{P}(K(w) > L_0, K(w') > L_0 | K(v_0) = x) \leq C_q \exp(-\lambda_q(L_0 - q^m x)). \tag{60}$$

Combining (57) with (58), and then using the bounds (59) and (60), we have

$$\mathbb{P}\left(\bigcap_{i,j} \{J(u_{i,j}) > t_{i,j}\}\right) \leq C_q e^{-(\nu-1+\lambda_q)L_0} \mathbb{E}\left[\exp\left(\lambda_q q^m K(v_0) + q^m \sum_{i=1}^{\nu-1} K(v_i)\right)\right].$$

Let $m \in \mathbb{N}$ be sufficiently large so that $\lambda_q \nu q^m < 2\nu q^m < 1/2$ holds. We have

$$\mathbb{E}\left[\exp\left(\lambda_q q^m K(v_0) + q^m \sum_{i=1}^{\nu-1} K(v_i)\right)\right] \le \mathbb{E}\left[\exp\left(\frac{1}{2\nu}\sum_{i=0}^{\nu-1} K(v_i)\right)\right] \le \mathbb{E}[e^{\frac{1}{2}K_\infty}] = C_q,$$

where we used the fact that $\exp(\frac{1}{n}\sum_{i=0}^{n-1} a_i) \le \frac{1}{n}\sum_{i=0}^{n-1}\exp(a_i)$ for the second inequality. In particular, we have

$$\mathbb{P}\left(\bigcap_{i,j}\{J(u_{i,j}) > t_{i,j}\}\right) \le C_q e^{-(\nu-1+\lambda_q)L_0}.$$

Since $L_0 = t_0 + \gamma n$, the claim follows. $\square$

5. Left tails for the geometric sum of exponentials. In this section, we work towards proving Theorem 2.3. Before presenting the proof of Theorem 2.3, we will give two preliminary lemmas.

LEMMA 5.1. *For all $m \in \mathbb{N}$ and $s \ge 0$, we have*

$$\frac{s^m}{m!} q^{-m(m-1)/2} \exp\left(-\frac{sq^{-m}}{(q^{-1}-1)m}\right) \le \mathbb{P}(K_{m-1} \le s) \le \frac{s^m}{m!} q^{-m(m-1)/2}. \tag{61}$$

PROOF. From the definition (6) of K_{m-1}, we see that

$$\mathbb{P}(K_{m-1} \le s) = \int_{\mathbb{R}_+^m} 1_{\{u_0+u_1+\cdots u_{m-1} \le s\}} \prod_{i=0}^{m-1} q^{-i}\exp(-q^{-i}u_i)\,\mathrm{d}u_0\,\mathrm{d}u_1\cdots\mathrm{d}u_{m-1}$$

holds for all $s \ge 0$. The integral is taken over the s-scaled unit m-simplex of volume $s^m/m!$,

$$s\Delta^m = \left\{(u_0,\dots,u_{m-1}) : u_i \ge 0, \sum_{i=0}^{m-1} u_i \le s\right\}.$$

In particular, we may write

$$\mathbb{P}(K_{m-1} \le s) = \frac{s^m}{m!} q^{-m(m-1)/2}\mathbb{E}\left[\exp\left(-\sum_{i=0}^{m-1} q^{-i}\zeta_i\right)\right],$$

where $(\zeta_0,\dots,\zeta_{m-1})$ is a random vector uniformly distributed on $s\Delta^m$. The upper bound in (61) follows from the simple estimate $\mathbb{E}[\exp(-s\sum_{i=0}^{m-1} q^{-i}\zeta_i)] \le 1$. To prove the lower bound note that by Jensen's inequality, we have

$$\mathbb{E}\left[\exp\left(-\sum_{i=0}^{m-1} q^{-i}\zeta_i\right)\right] \ge \exp\left(-\mathbb{E}\left[\sum_{i=0}^{m-1} q^{-i}\zeta_i\right]\right).$$

The lower bound in (61) now follows from noting that, since $\mathbb{E}[\zeta_i] = s/m$ for each $i \in \{0,\dots,m-1\}$, we have

$$\mathbb{E}\left[\sum_{i=0}^{m-1} q^{-i}\zeta_i\right] = \frac{s}{m}\frac{q^{-m}-1}{q^{-1}-1} \le \frac{s}{m}\frac{q^{-m}}{q^{-1}-1}.$$

$\square$

We emphasize that thanks to Lemma 5.1, it can be seen that whenever the quantity $\frac{sq^{-m}}{(q^{-1}-1)m}$ is small, the quantity $s^m q^{-m(m-1)/2}/m!$ is a good estimate for $\mathbb{P}(K_{m-1} \le s)$. Our proofs of both the upper and lower bounds in Theorem 2.3 will involve combining monotonicity arguments—namely that for $n \ge m$, K_n stochastically dominates K_m—with taking an optimal choice of m. For the latter, we have the following lemma, which identifies a critical choice of $m(s)$ so that $s^m q^{-m(m-1)/2}/m!$ has the order $e^{-F_q(s)}$.

LEMMA 5.2. *For each $s \in (0, 1/e^2]$, letting $m(s)$ be the smallest integer greater than $\kappa(\log \frac{1}{s} + \log\log \frac{1}{s})$, we have*

$$\frac{1}{C_q} \exp\big(-F_q(s)\big) \le s^{m(s)} q^{-m(s)(m(s)-1)/2}/m(s)! \le C_q \exp\big(-F_q(s)\big),$$

where $F_q(s)$ is as in Theorem 2.3.

PROOF. Note that by using the Stirling bounds $\sqrt{2\pi} m^{m+1/2} e^{-m} \le m! \le e m^{m+1/2} e^{-m}$, as well as the definition $q = e^{-1/\kappa}$, we have

$$\frac{1}{C} \exp\big(f(s, m)\big) \le s^m q^{-m(m-1)/2}/m! \le C \exp\big(f(s, m)\big),$$

where for all $x, y > 0$

$$f(x, y) := \frac{y^2}{2\kappa} - \left(\log \frac{1}{x} - 1 + \frac{1}{2\kappa}\right) y - (y + 1/2) \log y. \tag{62}$$

Using the shorthand $S := \log \frac{1}{s}$, by setting $m(s) := \kappa(S + \log S + \delta(s))$, where $\delta(s) \in (0, 1/\kappa]$ is such that $m(s)$ is an integer, a calculation tells us that

$$\begin{aligned} f\big(s, m(s)\big) = &- \frac{\kappa}{2}\big(S + \log S + \delta(s)\big)\left(S + \log S - \delta(s) - 2 + \frac{1}{\kappa} + 2\log\kappa + 2\varepsilon(s)\right) \\ &- \frac{1}{2} \log S - \frac{1}{2} \log\kappa - \frac{1}{2}\varepsilon(s), \end{aligned}$$

where $\varepsilon(s) := \log m(s) - \log(\kappa S) = \log(1 + \frac{\log S + \delta(s)}{S})$. Using the fact that $\log(1 + x) - x = O(x^2)$ as $x \to 0$, we obtain

$$\big(S + \log S + \delta(s)\big)\varepsilon(s) = \log S + r(q, s),$$

where $r(q, s)$ is bounded in $s \le 1/e^2$ for each q. In particular, by the last two displays, we have

$$\begin{aligned} f\big(s, m(s)\big) = &- (1/2 + \kappa) \log S \\ &- \frac{\kappa}{2}\big(S + \log S + \delta(s)\big)\left(S + \log S - \delta(s) - 2 + \frac{1}{\kappa} + 2\log\kappa\right) S + r'(q, s), \end{aligned}$$

where $r'(q, s)$ is uniformly bounded in $s \le 1/e^2$. Using the identity $(x + a)(x + b) = (x + \frac{a+b}{2})^2 - (\frac{a-b}{2})^2$, we obtain

$$f\big(s, m(s)\big) = -\left(\frac{1}{2} + \kappa\right) \log S - \frac{\kappa}{2}\left(S + \log S + \frac{1}{2\kappa} + \log\kappa - 1\right)^2 + r''(q, s)$$

for a $r''(q, s)$ uniformly bounded in $s \le 1/e^2$. This completes the proof. □

PROOF OF THEOREM 2.3. Let $n \in \mathbb{N}$ be such that $n \ge \kappa(\log \frac{1}{s} + \log\log \frac{1}{s})$. Then by construction, n is at least $m(s)$ for all n sufficiently large. Hence, by stochastic domination, we have

$$\mathbb{P}(K_{n-1} \le s) \le \mathbb{P}(K_{m(s)-1} \le s).$$

It then follows from the upper bound in Lemma 5.1 and the upper bound in Lemma 5.2 that for every $s \le 1/e^2$,

$$\mathbb{P}(K_{n-1} \le s) \le \mathbb{P}(K_{m(s)-1} \le s) \le C_q e^{-F_q(s)},$$

completing the proof of the upper bound in (18).

We now turn to proving the more difficult lower bound in (18). Since K_∞ stochastically dominates K_{n-1} for every $n \in \mathbb{N}$, it is sufficient to prove the lower bound for $n = \infty$. To this end, with $m(s)$ as in Lemma 5.2, iterating (8) $m(s)$ times yields

$$K_\infty \overset{(d)}{=} K_{m(s)-1} + q^{m(s)} K_\infty, \quad K_{m(s)-1} \text{ independent from } K_\infty.$$

Our strategy is as follows. For a carefully chosen $\varepsilon(s) > 0$, we use the bound

$$\mathbb{P}(K_\infty \le s) \ge \mathbb{P}\big(K_{m(s)-1} \le \big(1-\varepsilon(s)\big)s\big)\mathbb{P}\big(q^{m(s)} K_\infty \le \varepsilon(s)s\big). \tag{63}$$

It transpires that the best choice of $\varepsilon(s)$ to be taken is so that $\varepsilon(s)sq^{-m(s)}$ has unit order. Indeed, with $S = \log \frac{1}{s}$ as above, set

$$\varepsilon(s) := 1/S.$$

Using again $m(s) = \kappa(\log \frac{1}{s} + \log\log \frac{1}{s} + \delta(s))$, a calculation tells us that

$$\varepsilon(s)sq^{-m(s)} = e^{\delta(s)} \ge 1$$

for every s, so that in particular

$$\mathbb{P}\big(q^{m(s)} K_\infty \le \varepsilon(s)s\big) = \mathbb{P}\big(K_\infty \le \varepsilon(s)sq^{-m(s)}\big) \ge \mathbb{P}(K_\infty \le 1) \ge C_q.$$

Moreover, by (63) with $\varepsilon(s) = 1/S$, we have

$$\mathbb{P}(K_\infty \le s) \ge C_q \mathbb{P}\big(K_{m(s)-1} \le \big(1-\varepsilon(s)\big)s\big). \tag{64}$$

By Lemma 5.1, we can write

$$\mathbb{P}(K_{m-1} \le s) \ge C_q \exp\big(f(s,m) - g(s,m)\big), \tag{65}$$

where $f(s,m)$ is given as in (62), and

$$g(s,m) := \frac{e^{\kappa^{-1}m}s}{(q^{-1}-1)m}.$$

Set $w(s) := (1-\varepsilon(s))s$ and let $m(s) = \kappa(S + \log S + \delta(s))$ be defined as above. A calculation yields

$$g\big(w(s), m(s)\big) \le \frac{e^{\delta(s)}}{q^{-1}-1} \le C_q.$$

We now turn to computing $f(w(s), m(s))$. Again, a calculation similar to the one in the proof of Lemma 5.2 tells us that

$$f\big(w(s), m(s)\big) = F_q(s) + r(q,s)$$

where for each q, $r(q,s)$ is bounded uniformly in $s \le 1/e^2$. In particular, we see that the difference between $f(w(s), m(s))$ and $f(s, m(s))$ is bounded. Using (65), we have that

$$\mathbb{P}\big(K_{m(s)-1} \le w(s)\big) \ge C_q \exp\big(-F_q(s)\big),$$

and by (64)

$$\mathbb{P}(K_\infty \le s) \ge C_q \exp\big(-F_q(s)\big).$$

This completes the proof of the lower bound in (18). □

6. The leftmost particles in the rescaled expanding branching random walk.

6.1. *Three preliminary estimates.* Recall the rescaled expanding branching random walk $(K(v))_{v\in\mathbb{T}}$ defined in Section 2. We now give three simple estimates on the random variables $(K(v))_{v\in\mathbb{T}}$, which will be used in the proof of Theorem 2.2. Combined with the estimates on the lower tails of $K(v)$ from Theorem 2.3, this allows us to determine a sharp concentration of the size of the smallest fragment in Theorem B. The following lemma gives an upper bound on the joint tails of $K(v)$ and $K(w)$, which will be useful when w is close to v.

LEMMA 6.1. *Let $v, w \in \mathbb{T}_n$ and suppose that $|v \wedge w| = n - m - 1$ for some $m \geq 0$. Then for all $x \geq 0$, we have that*

$$\mathbb{P}\big(K(v) \leq x, K(w) \leq x\big) \leq \mathbb{P}(K_n \leq x)\mathbb{P}(K_m \leq x).$$

PROOF. Recall the expanding branching random walk $S = (S(x))_{x\in\mathbb{T}}$ defined in Section 2. We have

$$\begin{aligned}\mathbb{P}\big(S(v) \leq t, S(w) \leq t\big) &= \mathbb{P}\big(S(v) \leq t, S(v \wedge w) + S(w) - S(v \wedge w) \leq t\big)\\ &\leq \mathbb{P}\big(S(v) \leq t, S(w) - S(v \wedge w) \leq t\big)\\ &= \mathbb{P}\big(S(v) \leq t\big)\mathbb{P}\big(S(w) - S(v \wedge w) \leq t\big).\end{aligned}$$

Now use the fact that $\{q^n S(v), q^n S(w)\} = \{K(v), K(w)\}$, where we have $q^n S(v) \stackrel{d}{=} K_n$ as well as that $S(w) - S(v \wedge w) \stackrel{d}{=} q^{-(n-m)} S_m \stackrel{d}{=} q^{-n} K_m$, and conclude by substituting $t = q^{-n}x$. □

LEMMA 6.2. *Let $v, w \in \mathbb{T}_n$ and suppose that $|v \wedge w| = n - m - 1$ for some $m \geq 0$. Then for all $s, x \geq 0$, we have*

$$\mathbb{P}\big(K(v) \leq s, K(w) \leq s\big) \leq \mathbb{P}(K_n \leq s)\big(\mathbb{P}\big(K_n \leq s + q^{m+1}x\big) + \mathbb{P}(K_{n-m-1} > x)\big).$$

PROOF. Lemma 6.1 gives $\mathbb{P}(K(v) \leq s, K(w) \leq s) \leq \mathbb{P}(K(v) \leq s)\mathbb{P}(K_m \leq s)$. For a pair of independent random variables (K_m, K_{n-m-1}) as defined in (6), set $\widetilde{K}_n := K_m + q^{m+1}K_{n-m-1}$. Note that we have

$$\begin{aligned}\mathbb{P}(K_m \leq s) &\leq \mathbb{P}(K_m \leq s, K_{n-m-1} \leq x) + \mathbb{P}(K_{n-m-1} > x)\\ &\leq \mathbb{P}\big(\widetilde{K}_n \leq s + q^{m+1}x\big) + \mathbb{P}(K_{n-m-1} > x)\end{aligned}$$

for all $x, s \geq 0$. Since $\widetilde{K}_n$ has the same law as $K(v)$, we conclude the proof. □

The following lemma provides an estimate for the probability that $K(v)$ is contained in a small interval.

LEMMA 6.3. *For all $s, z \geq 0$ and $n \in \mathbb{N}$, we have that*

$$\mathbb{P}\big(K_n \in [s, s+z]\big) \leq z\mathbb{P}\big(K_{n-1} \leq q^{-1}(s+z)\big).$$

PROOF. Consider the following general fact. If A is any nonnegative random variable and W is an independent standard exponential random variable, then since the density function of W is bounded above by one we have

$$\mathbb{P}\big(A + W \in [s, s+z]\big) \leq z\mathbb{P}(A \leq s+z).$$

The result in question follows from this general fact by noting that $K_n \stackrel{d}{=} qK_{n-1} + W$ where W is an independent standard exponential. □

6.2. *Second moment method.* In order to prove Theorem 2.2, we will apply the second moment method with respect to the following sum of indicator random variables:

$$M_n(s) = \sum_{v \in \mathbb{T}_n} I_{\{K(v) \le s\}}.$$

We start with the following bound on the expectation of $M_n(s)$. Define $z_n := z_n(\kappa, \gamma)$ as the unique solution to the equation

$$z_n + \log z_n + \frac{1}{2\kappa} + \log \kappa - 1 = \sqrt{\frac{2\gamma}{\kappa} n}. \tag{66}$$

It is easily verified that

$$z_n = \sqrt{\frac{2\gamma}{\kappa} n} - \frac{1}{2} \log n - \frac{1}{2\kappa} - \frac{1}{2} \log \kappa + 1 - \frac{1}{2} \log(2\gamma) + O\left(\frac{\log n}{\sqrt{n}}\right). \tag{67}$$

LEMMA 6.4. *With z_n as in (66), for $n \in \mathbb{N}$ define the quantities*

$$s_n^- := \exp(-z_n - z_n^{-1} \log^2 z_n) \quad \textit{and} \quad s_n^+ := \exp(-z_n + z_n^{-1} \log^2 z_n) \tag{68}$$

Then

$$\mathbb{E}[M_n(s_n^-)] = k^n \mathbb{P}(K_n \le s_n^-) \le \frac{1}{n^2} \quad \textit{and} \quad \mathbb{E}[M_n(s_n^+)] = k^n \mathbb{P}(K_n \le s_n^+) \ge n^2$$

for all $n \in \mathbb{N}$ sufficiently large.

PROOF. By the definition of z_n, we have that $F_q(s)$ given in (19) satisfies

$$F_q(\exp(-z_n + y_n)) = \left(\frac{1}{2} + \kappa\right) \log z_n + \frac{\kappa}{2} \left(\sqrt{\frac{2\gamma}{\kappa} n} + y_n\right)^2 + O(y_n)$$

for all $(y_n)_{n \in \mathbb{N}}$ with $\lim_{n \to \infty} y_n = 0$. Hence, using Theorem 2.3 and the fact that $\log^2(z_n) = O(\log^2 n)$, we see that

$$\log(k^n \mathbb{P}(\log K_n \le -z_n - z_n^{-1} \log^2 z_n)) \le -\frac{1}{2} \log n - \frac{\kappa}{2} \log^2 z_n \le -2 \log n$$

for all n large enough. Similarly, we apply Theorem 2.3 to obtain

$$\log(k^n \mathbb{P}(\log K_n \ge -z_n + z_n^{-1} \log^2 z_n)) \ge -\frac{1}{2} \log n + \frac{\kappa}{2} \log^2 z_n \ge 2 \log n$$

for all n large enough, which concludes the proof. $\square$

We now have all tools to prove Theorem 2.2.

PROOF OF THEOREM 2.2. We begin by proving the slightly stronger statement that there exists almost surely an $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$ we have

$$\log K_n^{\min} \in [-z_n - z_n^{-1} \log^2 z_n, -z_n + z_n^{-1} \log^2 z_n]. \tag{69}$$

To this end, we see from Lemma 6.4 that for s_n^- defined in (68)

$$\mathbb{P}(\exists v \in \mathbb{T}_n : \log K(v) \le \log s_n^-) = \mathbb{P}(M_n(s_n^-) \ge 1) \le \mathbb{E}[M_n(s_n^-)] \le \frac{1}{n^2}.$$

This gives the $\mathbb{P}$-almost sure lower bound on $\log K_n^{\min}$ in (69), that is, $\log K_n^{\min} \ge \log s_n^-$ almost surely for $n \ge 0$ large enough. For the corresponding upper bound, we will now estimate

$\mathrm{Var}(M_n(s))$, where we set $s = s_n^+$. Partitioning according to the generation of the most recent common ancestor, we have that

$$\begin{aligned}\mathrm{Var}(M_n(s)) &= \sum_{v,w\in\mathbb{T}_n} \big(\mathbb{P}\big(K(v)\le s, K(w)\le s\big) - \mathbb{P}\big(K(v)\le s\big)\mathbb{P}\big(K(w)\le s\big)\big) \\ &\le \mathbb{E}\big[M_n(s)\big] \\ &\quad + \sum_{m=1}^{n-1} k^{n+m}\big(\mathbb{P}\big(K(v_m)\le s, K(w_m)\le s\big) - \mathbb{P}\big(K(v_m)\le s\big)\mathbb{P}\big(K(w_m)\le s\big)\big),\end{aligned}$$

where $v_m, w_m \in \mathbb{T}_n$ are chosen for all $m \ge 1$ such that the equality $|v_m \wedge w_m| = n-m$ holds. Splitting the last sum at $n/2$, we have, using Lemma 6.1, that

$$\begin{aligned}\sum_{m=1}^{n/2-1} k^{n+m}\mathbb{P}\big(K(v_m)\le s, K(w_m)\le s\big) &\le \sum_{m=1}^{n/2-1} k^n\mathbb{P}(K_n\le s)k^m\mathbb{P}(K_m\le s) \\ &= \mathbb{E}\big[M_n(s)\big]\sum_{m=1}^{n/2-1} k^m\mathbb{P}(K_m\le s).\end{aligned}$$

Recall that z_n is of order $\sqrt{n}$. Hence, since $s_m^- \ge s_n^+$ for all $m \le n/2$ when n is large enough, we can use Lemma 6.4 to see that $\sum_{m=0}^{n/2} k^m\mathbb{P}(K_m \le s)$ is bounded above uniformly in n. For the remaining terms, we apply Lemma 6.2 to get that

$$\begin{aligned}&k^{n+m}\big(\mathbb{P}\big(K(v_m)\le s, K(w_m)\le s\big) - \mathbb{P}\big(K(v_m)\le s\big)^2\big) \\ &\quad\le \mathbb{E}\big[M_n(s)\big]k^m\big(\mathbb{P}\big(K_n\in\big[s, s+q^m x\big]\big) + \mathbb{P}(K_\infty > x)\big)\end{aligned}$$

holds for all $x \ge 0$. Let $x = n^2$ and use Lemma 6.3 to obtain

$$\sum_{m=n/2}^{n-1} k^m\big(\mathbb{P}\big(K_n\in\big[s, s+q^m n^2\big]\big)\big) \le n^2 q^{n/2}\sum_{m=n/2}^{n-1} k^m\mathbb{P}\big(K_{n-1}\le q^{-1}2s\big) \tag{70}$$

as $s \ge q^{n/2}n^2$ for all n large enough by (67) and (68). Note that $s = s_n^+$ and n satisfy the conditions in Theorem 2.3. Since

$$F_q\big(q^{-1}2s\big) = \frac{\kappa}{2}\big(z_n^2 + z_n\log(z_n)\big) + O(z_n)$$

we get that

$$k^m\big(\mathbb{P}\big(K_n\in\big[s, s+q^m n^2\big]\big)\big) \le q^{-n^{2/3}}$$

for all n large enough and $m \le n$. Plugging this estimate into the right-hand side in (70), we obtain

$$\sum_{m=n/2}^{n-1} k^m\big(\mathbb{P}\big(K_n\in\big[s, s+q^m n^2\big]\big)\big) \le 1$$

for all n large enough.

Note that $nk^m\mathbb{P}(K_\infty > n^2) \le 1$ holds for all $m \le n$ with n sufficiently large by Lemma 3.1. Hence, combining the previous observations, we obtain that

$$\sum_{m=n/2}^{n-1} k^{n+m}\big(\mathbb{P}\big(K(v_m)\le s, K(w_m)\le s\big) - \mathbb{P}(K_n\le s)^2\big) \le 2\mathbb{E}\big[M_n(s)\big]$$

holds for all $x \geq 0$ and all n large enough. Thus, we conclude that the variance $\mathrm{Var}(M_n(s_n^+))$ is of order at most $\mathbb{E}[M_n(s_n^+)]$. Using the Paley–Zygmund inequality, we have, writing again $s = s_n^+$,

$$\mathbb{P}(\exists v \in \mathbb{T}_n : \log K(v) \leq \log s) = \mathbb{P}(M_n(s) > 0) \geq \frac{\mathbb{E}[M_n(s)^2]}{\mathbb{E}[M_n(s)]^2} = 1 - \frac{\mathrm{Var}(M_n(s))}{\mathbb{E}[M_n(s)]^2} \geq 1 - \frac{c}{n^2}$$

for all n large enough, where we used Lemma 6.4 for the last inequality. Applying the Borel–Cantelli lemma, this gives us the upper bound on $\log K_n^{\min}$ in (69), that is, $\log K_n^{\min} \leq \log s_n^+$ almost surely for n large enough.

We now obtain the statement in Theorem 2.2 from the stronger statement given in (69). First we note that with w_n as in the statement of Theorem 2.2, by (67) we have $w_n = z_n + O(\frac{\log n}{\sqrt{n}})$. In particular, since $\frac{\log^2 z_n}{z_n} = O(\frac{\log^2 n}{\sqrt{n}}) = o(n^{-1/3})$, we obtain that there exists almost surely an n_0 in $\mathbb{N}$ such that for all $n \geq n_0$ we have

$$\log K_n^{\min} \in \left[-w_n - \frac{1}{n^{1/3}}, -w_n + \frac{1}{n^{1/3}}\right]$$

which is precisely the statement of Theorem 2.2. □

6.3. *Proof of Theorem* B. We are now ready to use Theorem 2.2 to give a proof of Theorem B.

PROOF OF THEOREM B USING THEOREM 2.2. Let k^{-M_t} denote the size of the smallest fragment in the process at time t. Let

$$T_n := \sup\{t \geq 0 : M_t = n\}$$

denote the last time at which the smallest fragment in the process has size k^{-n}, that is, T_n is the last time at which all fragments in the process have size k^{-n} or larger. In particular, based on our discussion in Section 2.1, $T_n := \min_{v \in \mathbb{T}_n} S(v)$. Since $S_n^{\min} = q^{-n} K_n^{\min}$, by Theorem 2.2 there exists almost surely some n_0 in $\mathbb{N}$ such that for all $n \geq n_0$

$$\exp\left\{\frac{n}{\kappa} - w_n - \frac{1}{n^{1/3}}\right\} \leq T_n \leq \exp\left\{\frac{n}{\kappa} - w_n + \frac{1}{n^{1/3}}\right\}, \tag{71}$$

where we made use of the fact that $q^{-n} = e^{n/\kappa}$. Set $\tilde{c} := \frac{1}{2\kappa} + \frac{1}{2}\log\kappa - 1 + \frac{1}{2}\log(2\gamma)$ and for $\sigma \in \{-1, +1\}$ define

$$p_\sigma(x) := \exp\left\{\frac{x}{\kappa} - \sqrt{\frac{2\gamma}{\kappa}x} + \log(x)/2 + \tilde{c} + \sigma x^{-1/3}\right\}. \tag{72}$$

Using the definition of w_n given in the statement of Theorem 2.2, (71) reads as saying

$$p_{-1}(n) \leq T_n \leq p_{+1}(n)$$

for all $n \geq n_0$. In particular, we are in the setting of Lemma 3.2, so that there exists almost surely a t_0 in $\mathbb{R}_+$ such that for all $t \geq t_0$

$$\lceil p_{+1}^{-1}(t) \rceil \leq M_t \leq \lceil p_{-1}^{-1}(t) \rceil. \tag{73}$$

Setting $c := -\tilde{c} - \frac{1}{2}\log\kappa + \gamma$, so that it agrees with the constant c given in Theorem B, a brief calculation inverting (72) verifies that for $\sigma \in \{-1, +1\}$, we have

$$p_\sigma^{-1}(t) = \kappa\left(\log t + \sqrt{2\gamma \log t} - \frac{1}{2}\log\log t + c - \frac{\sigma}{(\kappa \log t)^{1/3}} + o\left(\frac{1}{\log^{1/3} t}\right)\right).$$

In particular, with $h(t)$ and $\mu_2 = 2\kappa^{2/3}$ as in the statement of Theorem B, for all t sufficiently large

$$h(t) - \mu_2 \frac{1}{\log^{1/3} t} \le p_{+1}^{-1}(t) \le p_{-1}^{-1}(t) \le h(t) + \mu_2 \frac{1}{\log^{1/3} t}. \tag{74}$$

Combining (73) with (74), we have

$$\left\lceil h(t) - \mu_2 \frac{1}{\log^{1/3} t} \right\rceil \le M_t \le \left\lceil h(t) + \mu_2 \frac{1}{\log^{1/3} t} \right\rceil,$$

which is precisely the statement of Theorem B. $\square$

APPENDIX

We provide here a proof of Lemma 4.4. We will (51) to control the behaviour of the expanding branching random walk $(S(v))_{v\in\mathbb{T}}$ (defined in (5)) by k independent copies of itself. An iteration of this procedure will yield our claim.

PROOF OF LEMMA 4.4. For $m > 0$, denote $\mathcal{E}_m = \bigcup_{j=0}^m \mathbb{T}_j$ and $\mathcal{E}_m(w) = \{v \in \mathcal{E}_m : w \le v\}$ for $w \in \mathbb{T} = \bigcup_n \mathbb{T}_n$. Using (51), we will first prove the following inequality:

$$\mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m} \{S(v) > t_v\}\Big) \ge \prod_{w\in\mathcal{E}_{k-1}} \mathbb{P}\big(S(w) > t_w\big) \prod_{w\in\mathbb{T}_k} \mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m(w)} \{S(v) > t_v\}\Big), \quad k \ge 0, \tag{75}$$

for any choice of real numbers t_v, where $v \in \mathcal{E}_m$. We let $\mathbb{P}(\bigcap_{v\in\mathcal{E}_m(w)}\{S(v) > t_v\}) = 1$ when $\mathcal{E}_m(w) = \varnothing$. Arguing inductively on k, first note that for $k = 0$ both sides of (75) are equal. Take $k > 0$ and note that for $w \in \mathbb{T}_k$ and $v \in \mathcal{E}_m(w)$ we have $S(v) = S(w) + (S(v) - S(w))$, where $S(w)$ is independent from $(S(v) - S(w))_{v\in\mathcal{E}_m(w)}$. Applying (51), we have that

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m(w)} \{S(v) > t_v\}\Big) &= \mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m(w)} \{S(v) - S(w) + S(w) > t_v\}\Big)\\ &\ge \mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m(w)} \{S(v) - S(w) + S_v(w) > t_v\}\Big),\end{aligned} \tag{76}$$

where $(S_v(w))_{v\in\mathcal{E}_m(w)}$ are i.i.d. copies of $S(w)$. The process $(S(v) - S(w))_{v\ge w}$ is distributed as $(q^{-|w|}(S(v) - S([0,1))))_{v\in\mathbb{T}}$ and therefore, conditionally on $(S_v(w))_{v\in\mathcal{E}_m(w)}$, we can invoke the branching property and get

$$\begin{aligned}&\mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m(w)} \{S(v) - S(w) + S_v(w) > t_v\}\Big)\\ &\quad= \mathbb{P}\big(S(w) > t_w\big) \prod_{z\in\mathbb{T}_{k+1}, z\ge w} \mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m(z)} \{S(v) - S(w) + S_v(w) > t_v\}\Big).\end{aligned} \tag{77}$$

Note that this is the exact place where we use the independence of $S_v(w)$'s. For each $w \in \mathbb{T}_k$, the process $(S(v) - S(w) + S_v(w))_{v>w}$ is distributed as $(S(v))_{v>w}$ and thus for any $z \in \mathbb{T}_{k+1}$, $z \ge w$

$$\mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m(z)} \{S(v) - S(w) + S_v(w) > t_v\}\Big) = \mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m(z)} \{S(v) > t_v\}\Big). \tag{78}$$

If we now combine the induction hypothesis with (76), (77) and (78) we arrive at

$$\mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m} \{S(v) > t_v\}\Big) = \prod_{w\in\mathcal{E}_k} \mathbb{P}\big(S(w) > t_w\big) \prod_{z\in\mathbb{T}_{k+1}} \mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m(z)} \{S(v) > t_v\}\Big).$$

This concludes the proof of (75). If we take $k > m$, (75) reads

$$\mathbb{P}\Big(\bigcap_{v\in\mathcal{E}_m}\{S(v) > t_v\}\Big) \ge \prod_{v\in\mathcal{E}_m}\mathbb{P}\big(S(v) > t_v\big). \tag{79}$$

The claim now follows by taking in (79), $m = n - \ell - 1$ and $t_v \to -\infty$ if $u_{i,j} \neq v$ for all $u_{i,j} \in \mathbf{u}$, and $t_v = t_{i,j}q^{|v|} - \gamma|v|$ if $v = u_{i,j}$. □

Acknowledgments. We thank Günter Last for answering questions about point processes and two anonymous referees for a careful reading and helpful suggestions.

Funding. SJ and JP are supported by the Austrian Science Fund (FWF) Project P32405 *Asymptotic geometric analysis and applications* of which JP is principal investigator.

DS thanks the Studienstiftung des deutschen Volkes and the TopMath program for financial support.

The research of PD was supported by the Alexander von Humboldt Foundation.

REFERENCES

[1] APOSTOL, T. M. (1976). *Introduction to Analytic Number Theory*. *Undergraduate Texts in Mathematics*. Springer, New York. MR0434929

[2] ATHREYA, K. B. (1985). Discounted branching random walks. *Adv. in Appl. Probab.* **17** 53–66. MR0778593 https://doi.org/10.2307/1427052

[3] BERESTYCKI, J. (2002). Ranked fragmentations. *ESAIM Probab. Stat.* **6** 157–175. MR1943145 https://doi.org/10.1051/ps:2002009

[4] BERTOIN, J. (2001). Homogeneous fragmentation processes. *Probab. Theory Related Fields* **121** 301–318. MR1867425 https://doi.org/10.1007/s004400100152

[5] BERTOIN, J. (2002). Self-similar fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat.* **38** 319–340. MR1899456 https://doi.org/10.1016/S0246-0203(00)01073-6

[6] BERTOIN, J. (2003). The asymptotic behavior of fragmentation processes. *J. Eur. Math. Soc.* (*JEMS*) **5** 395–416. MR2017852 https://doi.org/10.1007/s10097-003-0055-3

[7] BRENNAN, M. D. and DURRETT, R. (1986). Splitting intervals. *Ann. Probab.* **14** 1024–1036. MR0841602

[8] BRENNAN, M. D. and DURRETT, R. (1987). Splitting intervals. II. Limit laws for lengths. *Probab. Theory Related Fields* **75** 109–127. MR0879556 https://doi.org/10.1007/BF00320085

[9] DADOUN, B. (2017). Asymptotics of self-similar growth-fragmentation processes. *Electron. J. Probab.* **22** Paper No. 27, 30. MR3629871 https://doi.org/10.1214/17-EJP45

[10] DALEY, D. J. and VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes*. *Vol. I*: *Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1950431

[11] DENISOV, D. and ZWART, B. (2007). On a theorem of Breiman and a class of random difference equations. *J. Appl. Probab.* **44** 1031–1046. MR2382943 https://doi.org/10.1239/jap/1197908822

[12] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*. *Vol. II*, 2nd ed. Wiley, New York. MR0270403

[13] FILIPPOV, A. F. (1961). Über das Verteilungsgesetz der Grössen der Teilchen bei Zerstückelung. *Teor. Veroyatn. Primen.* **6** 299–318. MR0140159

[14] GOLDSCHMIDT, C. and HAAS, B. (2010). Behavior near the extinction time in self-similar fragmentations. I. The stable case. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 338–368. MR2667702 https://doi.org/10.1214/09-AIHP317

[15] GOLDSCHMIDT, C. and HAAS, B. (2016). Behavior near the extinction time in self-similar fragmentations II: Finite dislocation measures. *Ann. Probab.* **44** 739–805. MR3456350 https://doi.org/10.1214/14-AOP988

[16] KALLENBERG, O. (1975). *Random Measures*. *Schriftenreihe des Zentralinstituts für Mathematik und Mechanik Bei der Akademie der Wissenschaften der DDR* **23**. Akademie-Verlag, Berlin. MR0431372

[17] KOLMOGOROFF, A. N. (1941). Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. *C. R.* (*Dokl.*) *Acad. Sci. URSS* **31** 99–101. MR0004415

[18] KYPRIANOU, A., LANE, F. and MÖRTERS, P. (2017). The largest fragment of a homogeneous fragmentation process. *J. Stat. Phys.* **166** 1226–1246. MR3610212 https://doi.org/10.1007/s10955-017-1714-1

[19] LAST, G. and PENROSE, M. (2018). *Lectures on the Poisson Process*. *Institute of Mathematical Statistics Textbooks* **7**. Cambridge Univ. Press, Cambridge. MR3791470

[20] MACDONALD, I. G. (2015). *Symmetric Functions and Hall Polynomials*, 2nd ed. *Oxford Classic Texts in the Physical Sciences*. The Clarendon Press, Oxford Univ. Press, New York. With contribution by A. V. Zelevinsky and a foreword by Richard Stanley, Reprint of the 2008 paperback edition [MR1354144]. MR3443860

[21] ROOTZÉN, H. (1986). Extreme value theory for moving average processes. *Ann*. *Probab*. **14** 612–652. MR0832027