A TIGHTNESS CRITERION FOR FRAGMENTATIONS

GABRIEL BERZUNZA OJEDA, CECILIA HOLMGREN, AND SVANTE JANSON

Abstract. This note presents a simple criterion for the tightness of stochastic fragmentation processes. Our work is motivated by an application to a fragmentation process derived from deleting edges in a conditioned Galton-Watson tree [8]. In that paper, while finite-dimensional convergence was established, the claimed functional convergence relied on [12, Lemma 22], which is unfortunately incorrect. We show how our results can correct the proof in [8].

Furthermore, we show the applicability of our results by establishing tightness for fragmentation processes derived from various random tree models previously studied in the literature, such as Cayley trees, trees with specified degree sequences, and **p**-trees.

1. Introduction

We consider fragmentation processes. Intuitively, a (deterministic or random) fragmentation process describes how a mass is split into smaller and smaller pieces as time increases. We are only interested in the masses (or sizes) of the pieces, which are positive real numbers that we arrange in decreasing order. (We use "positive", "increasing" and "decreasing" in the weak sense.) The number of pieces may be finite or (countably) infinite, but we will always write the masses as an infinite sequence $(x_i)_1^\infty$, extending a finite sequence by zeroes. The space of possible mass configurations (at any given time $t \geqslant 0$) is thus

$$\mathbb{S}^{\downarrow} := \bigl\{(x_i)_1^\infty : x_1 \geqslant x_2 \geqslant \ldots \geqslant 0\bigr\}. \tag{1.1}$$

(We are usually interested only in smaller subspaces of this space, see Section 2, but we try to state results in great generality.) We define also a partial order $\preceq$ on $\mathbb{S}^{\downarrow}$; the intuitive meaning of $(y_i)_1^\infty \preceq (x_i)_1^\infty$ is that the fragmentation $(y_i)_1^\infty$ can be obtained from $(x_i)_1^\infty$ by further fragmenting each of the pieces in some way; see (2.5) for a formal definition.

A (deterministic or stochastic) *fragmentation process* is a function $X(t) = (X_i(t))_1^\infty$ from $\mathbb{R}_+ := [0,\infty)$ to $\mathbb{S}^{\downarrow}$, which is càdlàg (for the product topology on $\mathbb{S}^{\downarrow}$, see Section 2), and furthermore is decreasing for the order $\preceq$; in other words,

$$X(t_2) \preceq X(t_1) \qquad \text{when } 0 \leqslant t_1 \leqslant t_2. \tag{1.2}$$

Note that we do not assume any other properties, such as, for example, Markov. (We state all results for fragmentation processes defined on $[0,\infty)$; the results hold also for processes defined on $[0,1]$, *mutatis mutandis.*)

The purpose of this note is to state and prove a simple criterion for tightness of a sequence of stochastic fragmentation processes, see Section 3. The main purpose

Date: 6 July, 2025 (typeset July 28, 2025 8.34).

2020 *Mathematics Subject Classification.* 60F17.

Supported by the Knut and Alice Wallenberg Foundation; Ragnar Söderberg's Foundation; the Swedish Research Council.

of this result is to use it in applications where one can prove finite-dimensional convergence of the fragmentation processes; if one also can show that our condition holds, then one can conclude that the processes converge in the Skorohod space $D(\mathbb{R}_+, \mathbb{S}^\downarrow)$ (Theorem 3.3).

The motivation for this note is such an application to a certain fragmentation process derived from deleting edges in a random tree [8, Lemma 5]. In that paper first finite-dimensional convergence was proved, and then convergence in $D(\mathbb{R}_+, \mathbb{S}^\downarrow)$ was claimed using a general compactness result in [12, Lemma 22] – however, this lemma in [12] is unfortunately incorrect (see Section 7). We therefore show in Section 6 how the proof in [8] can be corrected using our results below.

We state our general results for fragmentation processes in Section 3, with proofs in Section 4. In the remaining sections we apply the general results to fragmentations of trees. We give general results in Section 5, with weighted versions in Section 9; Sections 6–8 and 10 give applications to some specific random trees.

2. Notation

Recall the definition (1.1) of $\mathbb{S}^\downarrow$. We equip $\mathbb{S}^\downarrow$ with the product topology, i.e., the topology as a subspace of $\mathbb{R}^\infty$. Convergence $(x_i^{(n)})_1^\infty \to (x_i)_1^\infty$ is thus equivalent to $x_i^{(n)} \to x_i$ for each i. Then $\mathbb{S}^\downarrow$ is a Polish space, i.e., a complete separable metric space. The choice of metric is (as usually) not important; one convenient standard choice is

$$d\big((x_i)_1^\infty, (y_i)_1^\infty\big) := \sum_{i=1}^\infty 2^{-i} \min\big(|x_i - y_i|, 1\big). \tag{2.1}$$

(Note that $\mathbb{S}^\downarrow$ is a closed subset of $\mathbb{R}^\infty$, and thus complete for this metric.)

We consider also the subspaces

$$\mathbb{S}^\downarrow_{\leqslant 1} := \Big\{(x_i)_1^\infty \in \mathbb{S}^\downarrow : 0 \leqslant \sum_{i=1}^\infty x_i \leqslant 1\Big\}, \tag{2.2}$$

$$\mathbb{S}^\downarrow_1 := \Big\{(x_i)_1^\infty \in \mathbb{S}^\downarrow : \sum_{i=1}^\infty x_i = 1\Big\}. \tag{2.3}$$

We recall some well-known facts.

Lemma 2.1. (i) *On $\mathbb{S}^\downarrow_{\leqslant 1}$, the ℓ^q metric is equivalent to the product metric* (2.1) *for any $q \in (1,\infty]$. $\mathbb{S}^\downarrow_{\leqslant 1}$ is a compact metric space (for any of these equivalent metrics).*

(ii) *On $\mathbb{S}^\downarrow_1$, the ℓ^1 metric is equivalent to the product metric* (2.1). *(And hence so is the ℓ^q metric for any $q \in [1,\infty]$.) $\mathbb{S}^\downarrow_1$ is a non-compact Polish space; the ℓ^1 metric is a complete separable metric.*

Sketch of proof. (i): Compactness follows since $\mathbb{S}^\downarrow_{\leqslant 1}$ is a closed subset of $[0,1]^\infty$. Convergence in the metric (2.1), and thus pointwise, implies convergence in ℓ^q for any $q \in (1,\infty)$ by dominated convergence, using the fact that $(x_i) \in \mathbb{S}^\downarrow_{\leqslant 1}$ implies $0 \leqslant x_i \leqslant 1/i$ for every i.

(ii): Elements of $\mathbb{S}^\downarrow_1$ can be regarded as probability measures on $\mathbb{N}$. Hence pointwise convergence implies convergence in ℓ^1 by Scheffé's lemma [18, Theorem 5.6.4]. Completeness follows since $\mathbb{S}^\downarrow_1$ is a closed subset of the Banach space ℓ^1.

If we define

$$x_i^{(n)} := \frac{1}{n}\mathbf{1}\{i \leqslant n\} \qquad \text{and} \qquad x^{(n)} := (x_i^{(n)})_1^\infty \in \mathbb{S}_1^{\downarrow}, \tag{2.4}$$

then $x^{(n)} \to 0$ in $\mathbb{S}^{\downarrow}$ (and thus in $\mathbb{S}_{\leqslant 1}^{\downarrow}$) in the product metric (2.1), but not in ℓ^1. This example shows that $\mathbb{S}_1^{\downarrow}$ is not a closed subset of $\mathbb{S}_{\leqslant 1}^{\downarrow}$ or $\mathbb{S}^{\downarrow}$, and hence $\mathbb{S}_1^{\downarrow}$ is not compact. (It shows also that the ℓ^1 metric is not an equivalent metric on $\mathbb{S}_{\leqslant 1}^{\downarrow}$.) $\square$

As said in the introduction, we define a partial order on $\mathbb{S}^{\downarrow}$; the formal definition is

$$(y_i)_1^\infty \preceq (x_i)_1^\infty \iff \exists \alpha : \mathbb{N} \to \mathbb{N} \text{ such that } \sum_{j:\alpha(j)=i} y_j \leqslant x_i \text{ for all } i \in \mathbb{N}. \tag{2.5}$$

The intuitive interpretation is that the piece corresponding to y_j is a fragment of the piece corresponding to $x_{\alpha(j)}$; in other words, the piece with mass x_i is fragmented into pieces with masses $(y_j : j \in \alpha^{-1}(i))$. Note the inequality in the definition (2.5); we allow mass to disappear (or, equivalently, becoming invisible dust); in particular, $\alpha^{-1}(i)$ may be empty. (For processes in $\mathbb{S}_1^{\downarrow}$, strict inequality is not possible so there is no dust.)

Recall that we define a (deterministic or stochastic) *fragmentation process* as a function $X(t) = (X_i(t))_1^\infty$ from $\mathbb{R}_+ := [0, \infty)$ to $\mathbb{S}^{\downarrow}$, which is càdlàg in the topology above, and decreasing for the order $\preceq$, i.e., (1.2) holds.

For any metric space $\mathbb{S}$, the space $D(\mathbb{R}_+, \mathbb{S})$ consist of all càdlàg functions $\mathbb{R}_+ \to \mathbb{S}$, equipped with the Skorohod J_1 topology; if $\mathbb{S}$ is complete and separable, then so is $D(\mathbb{R}_+, \mathbb{S})$, see e.g. [10, Chapter 3] or [16, Chapter 3]. Note that equivalent metrics in $\mathbb{S}$ define equivalent metrics and thus the same topology in $D(\mathbb{R}_+, \mathbb{S})$, e.g. as a consequence of [16, Problem 3.11.13].

Given a stochastic process $X(t)$, $t \geqslant 0$, we say that a random variable τ is an X-*stopping time* if it is a stopping time with respect to the natural filtration generated by X, i.e., the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ with $\mathcal{F}_t := \sigma\{X(s) : s \leqslant t\}$, $t \geqslant 0$.

We let $\mathbb{N} := \{1, 2, 3, \dots\}$.

For stochastic processes $X^{(n)}(t)$ and $X(t)$, *finite-dimensional convergence*, denoted $X_n(t) \xrightarrow{\text{f.d.}} X(t)$, means joint convergence in distribution $(X_n(t_i))_{i=1}^m \xrightarrow{\text{d}} (X(t_i))_{i=1}^m$ for any fixed finite sequence $t_1, \dots, t_m$. (In general, it suffices that this holds for $t_1, \dots t_m$ in the dense set of times t where $X(t)$ a.s. does not jump. In our applications this is the entire $\mathbb{R}_+$.)

If X_n are random variables and a_n positive numbers, then $X_n = O_{\mathrm{p}}(a_n)$ means that the variables X_n/a_n are bounded in probability (tight).

Unspecified limits are as $n \to \infty$.

3. Main results

One of our main results is the following, which is proved in Section 4, based on Aldous's criterion for tightness.

Theorem 3.1. *Let $X^{(n)}(t) = (X_i^{(n)}(t))_{i=1}^\infty$ be a sequence of fragmentation processes, and define*

$$Q^{(n)}(t) := \sum_{i=1}^\infty \big(X_i^{(n)}(t)\big)^2. \tag{3.1}$$

Suppose that for every uniformly bounded sequence of X^n-stopping times τ_n and every sequence h_n of positive numbers with $h_n \to 0$, we have, as $n \to \infty$,

$$Q^{(n)}(\tau_n) - Q^{(n)}(\tau_n + h_n) \xrightarrow{\mathrm{p}} 0. \tag{3.2}$$

Then, the family of processes $X^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$.

It is easy to see, and verified in detail in Lemma 4.2 below, that $Q^{(n)}(t)$ is a decreasing function, and thus $Q^{(n)}(\tau_n) \geqslant Q^{(n)}(\tau_n + h_n)$. Theorem 3.1 thus has the following corollary.

Corollary 3.2. *With the notation in Theorem 3.1, suppose that for every uniformly bounded sequence of X^n-stopping times τ_n and every sequence h_n of positive numbers with $h_n \to 0$, we have, as $n \to \infty$,*

$$\mathbb{E}\, Q^{(n)}(\tau_n) - \mathbb{E}\, Q^{(n)}(\tau_n + h_n) \to 0. \tag{3.3}$$

Then, the family of processes $X^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$.

We assume tacitly that the left-hand side of (3.3) is not of the form $\infty - \infty$.

As a simple consequence, we obtain the following, also proved in Section 4.

Theorem 3.3. *Let $X^{(n)}(t) = (X_i^{(n)}(t))_{i=1}^{\infty}$ be a sequence of fragmentation processes with $X^{(n)}(t) \in \mathbb{S}^{\downarrow}$, and suppose that we have finite-dimensional convergence*

$$X^{(n)}(t) \xrightarrow{\text{f.d.}} X(t) \tag{3.4}$$

for some process $X(t) \in D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$. Define $Q^{(n)}(t)$ by (3.1) *and suppose that the condition* (3.2) *(or the stronger* (3.3)*) holds for every uniformly bounded sequence of X^n-stopping times τ_n and every sequence h_n of positive numbers with $h_n \to 0$. Then, $X^{(n)}(t) \xrightarrow{\mathrm{d}} X(t)$ in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$.*

We have defined fragmentation processes as processes taking values in $\mathbb{S}^{\downarrow}$, and therefore the results above are stated for $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$. In applications we frequently have fragmentation processes taking values in $\mathbb{S}^{\downarrow}_{\leqslant 1}$ or $\mathbb{S}^{\downarrow}_{1}$, and then the results above often may be combined with the following simple lemmas to conclude the corresponding results in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$ or $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{1})$. Note that it is potentially more useful to use these spaces since, as shown in Lemma 2.1, we may use stronger metrics on $\mathbb{S}^{\downarrow}_{\leqslant 1}$ and $\mathbb{S}^{\downarrow}_{1}$ than on $\mathbb{S}^{\downarrow}$, for example the ℓ^2 and ℓ^1 metrics, respectively.

Lemma 3.4. *Let $X^{(n)}(t) = (X_i^{(n)}(t))_{i=1}^{\infty}$ be a sequence of fragmentation processes with $X^{(n)}(t) \in \mathbb{S}^{\downarrow}_{\leqslant 1}$.*

(i) *The family of processes $X^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$ if and only if it is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$.*
(ii) *Suppose further that $X(t)$ is a process in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$. Then $X^{(n)}(t) \xrightarrow{\mathrm{d}} X(t)$ in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$ if and only if $X^{(n)}(t) \xrightarrow{\mathrm{d}} X(t)$ in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$.*

Proof. Immediate, since $\mathbb{S}^{\downarrow}_{\leqslant 1}$ is a closed subspace of $\mathbb{S}^{\downarrow}$, and thus $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$ is a closed subspace of $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$ with the subspace topology. (This is a consequence of e.g. [16, Proposition 3.5.3].) $\square$

Lemma 3.5. *Let $X^{(n)}(t) = (X_i^{(n)}(t))_{i=1}^\infty$ be a sequence of fragmentation processes with $X^{(n)}(t) \in \mathbb{S}_1^\downarrow$. Suppose further that $X(t)$ is a process in $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$. Then $X^{(n)}(t) \xrightarrow{\mathrm{d}} X(t)$ in $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$ if and only if $X^{(n)}(t) \xrightarrow{\mathrm{d}} X(t)$ in $D(\mathbb{R}_+, \mathbb{S}^\downarrow)$, and also if and only if $X^{(n)}(t) \xrightarrow{\mathrm{d}} X(t)$ in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^\downarrow)$.*

Proof. Immediate, since $\mathbb{S}_1^\downarrow$ is a subspace of $\mathbb{S}_{\leqslant 1}^\downarrow$ and $\mathbb{S}^\downarrow$, and thus $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$ is a subspace of both $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^\downarrow)$ and $D(\mathbb{R}_+, \mathbb{S}^\downarrow)$, with the subspace topology. (Again by e.g. [16, Proposition 3.5.3].) □

Remark 3.6. Note that there is no result for $\mathbb{S}_1^\downarrow$ corresponding to Lemma 3.4(i): even if all $X^{(n)}(t) \in \mathbb{S}_1^\downarrow$, we cannot conclude tightness in $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$ without further assumptions. (This is because $\mathbb{S}_1^\downarrow$ is not a closed subspace of $\mathbb{S}_{\leqslant 1}^\downarrow$ and $\mathbb{S}^\downarrow$.) A counterexample is provided by $X^{(n)}(t) := x^{(n)}$ in (2.4) for all $t \geqslant 0$, or (if we want $X^{(n)}(0)$ fixed) by $X^{(n)}(t) := x^{(n)}$ for $t \geqslant U$ and $X^{(n)}(t) := (1, 0, 0, \dots)$ for $0 \leqslant t < U$, with $U \sim \mathrm{Exp}(1)$, say. △

Remark 3.7. Yet another alternative, used in for example [8] is to consider fragmentation processes as taking values in the space

$$\mathbb{S}_{<\infty}^\downarrow := \Big\{ (x_i)_1^\infty \in \mathbb{S}^\downarrow : 0 \leqslant \sum_{i=1}^\infty x_i < \infty \Big\}, \tag{3.5}$$

equipped with the ℓ^1 metric. Note that $\mathbb{S}_1^\downarrow$ is a closed subspace of $\mathbb{S}_{<\infty}^\downarrow$. (See Lemma 2.1.) Consequently, analogously to Lemma 3.5, for processes taking values in $\mathbb{S}_1^\downarrow$, tightness or convergence in $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$ is equivalent to tightness or convergence in $D(\mathbb{R}_+, \mathbb{S}_{<\infty}^\downarrow)$. △

4. Proof of main results

We begin with two simple deterministic lemmas.

Lemma 4.1. *Suppose that $(y_i)_1^\infty \in \mathbb{S}^\downarrow$ and that $x \geqslant \sum_i y_i$. Then*

$$x(x - y_1) \leqslant x^2 - \sum_{i=1}^\infty y_i^2. \tag{4.1}$$

Proof. We have

$$x^2 - x(x - y_1) = xy_1 \geqslant \sum_{i=1}^\infty y_i y_1 \geqslant \sum_{i=1}^\infty y_i^2. \tag{4.2}$$

□

Lemma 4.2. *Suppose that $X(t) = (X_i(t))_{i=1}^\infty$, $t \geqslant 0$, is a fragmentation process, and let, for $k \geqslant 1$,*

$$S_k(t) := \sum_{i=1}^k X_i(t) \tag{4.3}$$

be the sum of the k largest masses. Define

$$Q(t) := \sum_{i=1}^\infty [X_i(t)]^2. \tag{4.4}$$

Then, for any two times $0 \leqslant t_1 \leqslant t_2$ *we have*

$$Q(t_1) \geqslant Q(t_2) \tag{4.5}$$

and, for every $k \geqslant 1$*,*

$$0 \leqslant S_k(t_1) - S_k(t_2) \leqslant 2\sqrt{k(Q(t_1) - Q(t_2))}. \tag{4.6}$$

Proof. By our definition of a fragmentation process, we have $(X_i(t_2))_1^\infty \preceq (X_i(t_1))_1^\infty$. Let $\alpha : \mathbb{N} \to \mathbb{N}$ be a corresponding function as in the definition (2.5) and let $J_i := \alpha^{-1}(i)$, $i \geqslant 1$. Then (J_i) is a partition of $\mathbb{N}$, and

$$\sum_{j \in J_i} X_j(t_2) \leqslant X_i(t_1). \tag{4.7}$$

For the first inequality in (4.6), let $I_k := \{\alpha(j) : j \leqslant k\}$. Then $\{1, \dots, k\} \subseteq \bigcup_{i \in I_k} J_i$, and thus by summing (4.7) for $i \in I_k$,

$$S_k(t_2) \leqslant \sum_{i \in I_k} \sum_{j \in J_i} X_j(t_2) \leqslant \sum_{i \in I_k} X_i(t_1). \tag{4.8}$$

The set I_k has at most k elements, and thus the right-hand side of (4.8) is the sum of at most k masses at time t_1, whence it is at most $S_k(t_1)$. Hence $S_k(t_2) \leqslant S_k(t_1)$ as we wanted to show.

We turn to the second inequality in (4.6). Let, for each $i \geqslant 1$,

$$Y_i := \max_{j \in J_i} X_j(t_2), \tag{4.9}$$

interpreted as 0 if $J_i = \emptyset$. Note that the sets J_i are disjoint; thus $Y_1, \dots, Y_k$ are the masses of k different fragments at time t_2, and thus $\sum_{i=1}^k Y_i \leqslant S_k(t_2)$. Consequently,

$$S_k(t_1) - S_k(t_2) \leqslant S_k(t_1) - \sum_{i=1}^k Y_i = \sum_{i=1}^k (X_i(t_1) - Y_i). \tag{4.10}$$

Furthermore, by Lemma 4.1 applied to $(X_j(t_2) : j \in J_i)$, using (4.7) and (4.9),

$$X_i(t_1)\bigl(X_i(t_1) - Y_i\bigr) \leqslant X_i(t_1)^2 - \sum_{j \in J_i} X_j(t_2)^2. \tag{4.11}$$

Summing over all i yields

$$0 \leqslant \sum_{i=1}^\infty X_i(t_1)\bigl(X_i(t_1) - Y_i\bigr) \leqslant \sum_{i=1}^\infty X_i(t_1)^2 - \sum_{i=1}^\infty \sum_{j \in J_i} X_j(t_2)^2 = Q(t_1) - Q(t_2). \tag{4.12}$$

This implies (4.5). Moreover, let $a > 0$. Then (4.12) implies

$$\sum_{i: X_i(t_1) \geqslant a} \bigl(X_i(t_1) - Y_i\bigr) \leqslant a^{-1} \sum_{i: X_i(t_1) \geqslant a} X_i(t_1)\bigl(X_i(t_1) - Y_i\bigr) \leqslant a^{-1}\bigl(Q(t_1) - Q(t_2)\bigr). \tag{4.13}$$

Furthermore, obviously

$$\sum_{i \leqslant k: X_i(t_1) < a} \bigl(X_i(t_1) - Y_i\bigr) \leqslant \sum_{i \leqslant k: X_i(t_1) < a} X_i(t_1) \leqslant ka. \tag{4.14}$$

Consequently, by combining (4.13) and (4.14), for every $a > 0$,

$$\sum_{i=1}^{k}\bigl(X_i(t_1) - Y_i\bigr) \leqslant a^{-1}\bigl(Q(t_1) - Q(t_2)\bigr) + ka. \tag{4.15}$$

Optimizing by choosing $a := \sqrt{(Q(t_1) - Q(t_2))/k}$ yields the second inequality in (4.6). ☐

Proof of Theorem 3.1. Let, similarly to (4.3), $S_k^{(n)} := \sum_{i=1}^{k} X_i^{(n)}(t)$. Let stopping times τ_n and positive h_n be as in the statement. Then (3.2) holds by assumption and thus Lemma 4.2 implies that for every fixed $k \geqslant 1$,

$$S_k^{(n)}(\tau_n) - S_k^{(n)}(\tau_n + h_n) \xrightarrow{\text{p}} 0. \tag{4.16}$$

Since $X_i^{(n)}(t) = S_i^{(n)}(t) - S_{i-1}^{(n)}(t)$ (with $S_0^{(n)}(t) := 0$), it follows that

$$\bigl|X_i^{(n)}(\tau_n) - X_i^{(n)}(\tau_n + h_n)\bigr| \xrightarrow{\text{p}} 0 \tag{4.17}$$

for every i. The definition (2.1) implies that for every $k \geqslant 1$

$$d\bigl(X^{(n)}(\tau_n), X^{(n)}(\tau_n + h_n)\bigr) \leqslant \sum_{i=1}^{k} 2^{-i}\bigl|X_i^{(n)}(\tau_n) - X_i^{(n)}(\tau_n + h_n)\bigr| + 2^{-k}, \tag{4.18}$$

and consequently it follows easily from (4.17) that

$$d\bigl(X^{(n)}(\tau_n), X^{(n)}(\tau_n + h_n)\bigr) \xrightarrow{\text{p}} 0. \tag{4.19}$$

This shows that $X^{(n)}(t)$ satisfies Aldous's criterion [3] (see also e.g. [21, Theorem 16.11], or [10, Theorem 16.10] for a slightly different formulation), and thus the family $X^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$. ☐

Proof of Corollary 3.2. We have $Q^{(n)}(\tau_n) - Q^{(n)}(\tau_n + h_n) \geqslant 0$ by (4.5) in Lemma 4.2. Thus (3.3) is equivalent to $\mathbb{E}\,|Q^{(n)}(\tau_n) - Q^{(n)}(\tau_n + h_n)| \to 0$, which implies (3.2). ☐

Proof of Theorem 3.3. The family $X^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$ by Theorem 3.1 (or Corollary 3.2), which together with the assumed finite-dimensional convergence implies convergence in distribution in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$, see e.g. [10, Theorem 13.1]. ☐

5. An application: fragmentation of a tree

Consider now the special case of a fragmentation process of a (random) tree, where edges are deleted sequentially one by one in random order. More precisely, we assume that we are given a random tree $\mathcal{T}^{(n)}$ of size n for each $n \geqslant 1$ and we assume that conditioned on the tree, each edge e is deleted at some random time T_e, independent of all other edges. (The results extend to trees with arbitrary sizes $|\mathcal{T}^{(n)}| \to \infty$ with only notational changes.) We further assume, for convenience, that the random times T_e are i.i.d. with a common $\mathrm{Exp}(r_n)$ distribution, for some given rates r_n. (The argument works with a minor modification also for $T_e \sim U(0, t_n)$ with $t_n \to \infty$; alternatively this case follows easily from the exponential case studied below, see Corollary 5.5.)

Let $\mathcal{T}^{(n)}(t)$ be the fragmented forest at time t, so $\mathcal{T}^{(n)}(0) = \mathcal{T}^{(n)}$, and denote its components by $\mathcal{T}_i^{(n)}(t)$, $i = 1, \dots, \nu^{(n)}(t)$, arranged in order of decreasing size. Here, $\nu^{(n)}(t)$ denotes the number of components at time t. We define a fragmentation process $X^{(n)}(t) = (X_i^{(n)}(t))_{i=1}^{\infty}$ as the normalized component sizes, i.e., formally by

$X_i^{(n)}(t) := |\mathcal{T}_i^{(n)}(t)|/n$, with $X_i^{(n)}(t) := 0$ for $i > \nu^{(n)}(t)$. Note that $X^{(n)}(t) \in \mathbb{S}_1^{\downarrow}$ for all $t \geqslant 0$, with $X^{(n)}(0) = (1, 0, 0, \dots)$.

Theorem 5.1. *Let $\mathcal{T}^{(n)}(t)$ and $X^{(n)}(t)$ be as above and suppose that, with $Q^{(n)}(t)$ given by* (3.1)*,*

$$\limsup_{n\to\infty}\bigl(1 - \mathbb{E}\, Q^{(n)}(t)\bigr) \to 0 \qquad \textit{as } t \searrow 0. \tag{5.1}$$

Then, the family of processes $X^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$.

Proof. When an edge e is deleted at the time T_e, some component $\mathbf{t}$ of $\mathcal{T}^{(n)}(T_e-)$ is split into two $\mathbf{t}'$ and $\mathbf{t}''$. Thus, assuming as we may that all T_e are distinct,

$$\begin{aligned} Q^{(n)}(T_e-) - Q^{(n)}(T_e) &= \frac{1}{n^2}\bigl(|\mathbf{t}|^2 - |\mathbf{t}'|^2 - |\mathbf{t}''|^2\bigr) = \frac{2}{n^2}|\mathbf{t}'||\mathbf{t}''| \\ &= \frac{1}{n^2}\sum_{v,w\in\mathcal{T}^{(n)}} \mathbf{1}\{v \text{ and } w \text{ are connected in } \mathcal{T}^{(n)}(T_e-) \text{ but not in } \mathcal{T}^{(n)}(T_e)\} \\ &= \frac{1}{n^2}\sum_{v,w\in\mathcal{T}^{(n)}} \mathbf{1}\bigl\{e \text{ lies in the path between } v \text{ and } w \text{ in } \mathcal{T}^{(n)}(0), \\ &\qquad\qquad \text{and this path remains intact in } \mathcal{T}^{(n)}(T_e-)\bigr\}. \end{aligned} \tag{5.2}$$

Conditioning on $\mathcal{T}^{(n)}$ and all $T_{e'}$, $e' \neq e$, each term in the final sum in (5.2) is a decreasing function of T_e, and since T_e is independent of the past (unless it already has occured), it follows that for any stopping time τ_n and constant $h_n > 0$,

$$\begin{aligned} &\mathbb{E}\Bigl(\bigl(Q^{(n)}(T_e-) - Q^{(n)}(T_e)\bigr)\mathbf{1}\{\tau_n < T_e \leqslant \tau_n + h_n\}\Bigr) \\ &\qquad\qquad \leqslant \mathbb{E}\Bigl(\bigl(Q^{(n)}(T_e-) - Q^{(n)}(T_e)\bigr)\mathbf{1}\{0 < T_e \leqslant h_n\}\Bigr). \end{aligned} \tag{5.3}$$

Summing over all edges e yields

$$\mathbb{E}\bigl[Q^{(n)}(\tau_n) - Q^{(n)}(\tau_n + h_n)\bigr] \leqslant \mathbb{E}\bigl[Q^{(n)}(0) - Q^{(n)}(h_n)\bigr] = 1 - \mathbb{E}\, Q^{(n)}(h_n). \tag{5.4}$$

Hence, the condition (3.3) reduces to

$$\mathbb{E}\, Q^{(n)}(h_n) \to 1 \qquad \text{for any sequence } h_n \to 0. \tag{5.5}$$

This follows easily from (5.1). In fact, let $\varepsilon > 0$. Then (5.1) shows that there exists $t_0 > 0$ such that $\limsup_{n\to\infty}(1 - \mathbb{E}\, Q^{(n)}(t_0)) < \varepsilon$. Hence, for some n_0, if $n > n_0$ then $1 - \mathbb{E}\, Q^{(n)}(t_0) < \varepsilon$. Consequently, if n is so large that $n \geqslant n_0$ and $h_n < t_0$, then $Q^{(n)}(h_n) \geqslant Q^{(n)}(t_0) > 1 - \varepsilon$, and thus (5.5) follows. (In fact, it is easy to see that (5.5) and (5.1) are equivalent, and also $\sup_{n\geqslant 1}\bigl(1 - \mathbb{E}\, Q^{(n)}(t)\bigr) \to 0$ as $t \searrow 0$, using that trivially $\mathcal{T}^{(n)}(t) \to \mathcal{T}^{(n)}(0)$ and thus $Q^{(n)}(t) \to 1$ as $t \searrow 0$, for every fixed n.)

Thus, by Corollary 3.2, the sequence of processes $X_n(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$. This, in turn, implies tightness in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$ by Lemma 3.4, since $X_n(t) \in \mathbb{S}_{\leqslant 1}^{\downarrow}$. $\square$

Theorem 5.2. *Let $\mathcal{T}^{(n)}(t)$ and $X^{(n)}(t)$ be as above and suppose that we have finite-dimensional convergence $X^{(n)}(t) \xrightarrow{\text{f.d.}} X(t)$ to some stochastic process $X(t)$ in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$. Then $X^{(n)}(t) \xrightarrow{\text{d}} X(t)$ in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$.*

Proof. For each fixed $t \geqslant 0$ we have by assumption $X^{(n)}(t) \xrightarrow{\mathrm{d}} X(t)$, where $X^{(n)}(t) \in \mathbb{S}_1^{\downarrow} \subseteq \mathbb{S}_{\leqslant 1}^{\downarrow}$. Since $\mathbb{S}_{\leqslant 1}^{\downarrow}$ is a closed subset of $\mathbb{S}^{\downarrow}$, this implies $X(t) \in \mathbb{S}_{\leqslant 1}^{\downarrow}$ a.s. and $X^{(n)}(t) \xrightarrow{\mathrm{d}} X(t)$ in $\mathbb{S}_{\leqslant 1}^{\downarrow}$. Moreover, by Lemma 2.1, the ℓ^2 metric is an equivalent metric on $\mathbb{S}_{\leqslant 1}^{\downarrow}$; in particular, the ℓ^2 norm is a continuous function on $\mathbb{S}_{\leqslant 1}^{\downarrow}$, and thus we obtain, for every fixed $t \geqslant 0$,

$$Q^{(n)}(t) = \sum_{i=1}^{\infty} \bigl(X_i^{(n)}(t)\bigr)^2 \xrightarrow{\mathrm{d}} \sum_{i=1}^{\infty} \bigl(X_i(t)\bigr)^2 = Q(t). \tag{5.6}$$

Since $Q^{(n)}(t) \leqslant 1$, we obtain by (5.6) and dominated convergence,

$$\mathbb{E}\, Q^{(n)}(t) \to \mathbb{E}\, Q(t). \tag{5.7}$$

By definition, $X^{(n)}(0) = (1, 0, 0, \dots)$ and thus $Q^{(n)}(0) = 1$ for every n. Hence (5.6) implies $Q(0) = 1$ a.s. We assume that $X(t) \in D(\mathbb{R}_+, \mathbb{S}^{\downarrow})$, and in particular $X(t) \to X(0)$ a.s. as $t \searrow 0$. By the same argument as above, this implies $X(t) \to X(0)$ in ℓ^2 and $Q(t) \to Q(0)$ a.s., and thus $\mathbb{E}\, Q(t) \to \mathbb{E}\, Q(0) = 1$ as $t \searrow 0$. Consequently, using (5.7),

$$\lim_{t\to 0} \limsup_{n\to\infty} \bigl(1 - \mathbb{E}\, Q^{(n)}(t)\bigr) = \lim_{t\to 0} \bigl(1 - \mathbb{E}\, Q(t)\bigr) = 0, \tag{5.8}$$

which shows that (5.1) holds. Thus Theorem 5.1 applies and yields tightness of the family $X^{(n)}(t)$ in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$. This and the finite-dimensional convergence yield convergence in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$ (as in the proof of Theorem 3.3). ☐

We may also sometimes estimate $Q^{(n)}(t)$ directly.

Theorem 5.3. *With notation as in Theorem 5.1, let $d^{(n)}(v, w)$ denote the graph distance between vertices v and w in $\mathcal{T}^{(n)}$. Then, for every $t \geqslant 0$,*

$$1 - \mathbb{E}\, Q^{(n)}(t) \leqslant t\, r_n\, \mathbb{E}\, d^{(n)}(V_1, V_2), \tag{5.9}$$

where V_1 and V_2 are two independent uniformly random vertices of $\mathcal{T}^{(n)}$.

Consequently, if

$$r_n\, \mathbb{E}\, d^{(n)}(V_1, V_2) = O(1), \tag{5.10}$$

then the family of processes $X^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$.

Proof. Fix a time $t \geqslant 0$. Then, similary to (5.2),

$$\begin{aligned} Q^{(n)}(t) &= \frac{1}{n^2} \sum_{v,w \in \mathcal{T}^{(n)}} \mathbf{1}\{v \text{ and } w \text{ are connected in } \mathcal{T}^{(n)}(t)\} \\ &= \frac{1}{n^2} \sum_{v,w \in \mathcal{T}^{(n)}} \mathbf{1}\{T_e > t \text{ for every } e \text{ in the path between } v \text{ and } w \text{ in } \mathcal{T}^{(n)}\}, \end{aligned} \tag{5.11}$$

and thus

$$\begin{aligned} 1 - \mathbb{E}\left(Q^{(n)}(t) \mid \mathcal{T}^{(n)}\right) &= \mathbb{E}\left(Q^{(n)}(0) - Q^{(n)}(t) \mid \mathcal{T}^{(n)}\right) \\ &= \frac{1}{n^2} \sum_{v,w \in \mathcal{T}^{(n)}} \mathbb{P}\bigl(T_e \leqslant t \text{ for some } e \text{ in the path between } v \text{ and } w \text{ in } \mathcal{T}^{(n)} \mid \mathcal{T}^{(n)}\bigr) \end{aligned}$$

$$\leqslant \frac{1}{n^2} \sum_{v,w \in \mathcal{T}^{(n)}} r_n t d^{(n)}(v,w) = r_n t\,\mathbb{E}\bigl(d^{(n)}(V_1,V_2) \mid \mathcal{T}^{(n)}\bigr), \tag{5.12}$$

which shows (5.9). The final statement follows from (5.9) and Theorem 5.1. □

The following corollary, shows that we can replace the condition (5.10) on the average distance by corresponding conditions on the diameter or (in a rooted tree) the total path-length; it shows also that we then can weaken the condition from bounded in expectation to bounded in probability. If $\mathcal{T}^{(n)}$ is a rooted tree, with root denoted by ρ, then its *total path length* is defined as

$$\Upsilon^{(n)} := \Upsilon(\mathcal{T}^{(n)}) := \sum_{v \in \mathcal{T}^{(n)}} d^{(n)}(\rho, v) = n\,\mathbb{E}\bigl(d^{(n)}(\rho, V) \mid \mathcal{T}^{(n)}\bigr), \tag{5.13}$$

where V is a uniformly random vertex in $\mathcal{T}^{(n)}$.

For convenience in applications, we state the corollary with several versions of the conditions although some are redundant.

Corollary 5.4. *Let $\mathcal{T}^{(n)}$, r_n, and $X^{(n)}(t)$ be as above. Let $D^{(n)}$ be the diameter of $\mathcal{T}^{(n)}$ and, provided the tree $\mathcal{T}^{(n)}$ is rooted, let $\Upsilon^{(n)}$ be its total pathlength. Suppose that one of the following conditions holds:*

(i) $r_n\,\mathbb{E}\,D^{(n)} = O(1)$.

(ii) $r_n D^{(n)} = O_{\mathrm{p}}(1)$. *(I.e., the sequence $r_n D^{(n)}$ is tight.)*

(iii) $r_n\,\mathbb{E}\,\Upsilon^{(n)} = O(n)$.

(iv) $r_n \Upsilon^{(n)} = O_{\mathrm{p}}(n)$. *(I.e., the sequence $\frac{r_n}{n}\Upsilon^{(n)}$ is tight.)*

Then, the family of processes $X^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$.

Proof. (i): Let V_1 and V_2 be independent random vertices as in Theorem 5.3. Then $d^{(n)}(V_1,V_2) \leqslant D^{(n)}$, and thus (i) implies (5.10) and the tightness follows by Theorem 5.3.

(iii): Similarly, if $\mathcal{T}^{(n)}$ has the root ρ, then (5.13) implies

$$\mathbb{E}\bigl(d^{(n)}(V_1,V_2) \mid \mathcal{T}^{(n)}\bigr) \leqslant \mathbb{E}\bigl(d^{(n)}(V_1,\rho) + d^{(n)}(\rho, V_2) \mid \mathcal{T}^{(n)}\bigr) = \frac{2}{n}\Upsilon^{(n)}. \tag{5.14}$$

Consequently, $\mathbb{E}\,d^{(n)}(V_1,V_2) \leqslant \frac{2}{n}\,\mathbb{E}\,\Upsilon^{(n)}$, and thus (5.10) follows from (iii).

(ii): Let $\varepsilon > 0$. By (ii) there exists a constant C such that $\mathbb{P}(r_n D^{(n)} > C) < \varepsilon$ for every n. We may apply (i) to the conditioned trees $(\mathcal{T}^{(n)} \mid r_n D^{(n)} \leqslant C)$, and conclude that the conditioned processes $(X^{(n)}(t) \mid r_n D^{(n)} \leqslant C)$ form a tight family in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$. Hence, there exists a compact set K in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$ such that, for every n, $\mathbb{P}\bigl(X^{(n)}(t) \notin K \mid r_n D^{(n)} \leqslant C\bigr) \leqslant \varepsilon$, and consequently

$$\mathbb{P}\bigl(X^{(n)}(t) \notin K\bigr) \leqslant \mathbb{P}\bigl(X^{(n)}(t) \notin K \mid r_n D^{(n)} \leqslant C\bigr) + \mathbb{P}\bigl(r_n D^{(n)} > C\bigr) \leqslant 2\varepsilon. \tag{5.15}$$

This implies that the family $X^{(n)}(t)$ is tight.

(iv): Similar, now using (iii). □

Now, let us consider the case where each edge e in $\mathcal{T}^{(n)}$ is deleted at a random time $\widehat{T}_e \sim U(0, t_n)$, independently for all edges. Let $\widehat{\mathcal{T}}^{(n)}(t)$ be the fragmented forest at time t, and let $\widehat{X}^{(n)}(t) = (\widehat{X}^{(n)}_i(t))_{i \geqslant 1}$ be the corresponding fragmentation process of normalized component sizes, defined analogously as before in this section.

Corollary 5.5. *With notations as above, suppose that $t_n \to \infty$ and that one of the following conditions holds:*

(i) $\mathbb{E}\, d^{(n)}(V_1, V_2) = O(t_n)$.
(ii) $\mathbb{E}\, D^{(n)} = O(t_n)$.
(iii) $\mathbb{E}\, \Upsilon^{(n)} = O(nt_n)$.
(iv) $D^{(n)} = O_{\mathrm{p}}(t_n)$.
(v) $\Upsilon^{(n)} = O_{\mathrm{p}}(nt_n)$.

Then, the family of processes $\widehat{X}^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$.

Proof. Define $a_n(t) := t_n(1 - e^{-t/t_n})$. Observe that $\widehat{\mathcal{T}}^{(n)}(a_n(t))$ represents the fragmented forest at time t resulting from independently deleting each edge e in $\mathcal{T}^{(n)}$ at a random time $T_e := -t_n \log(1 - \widehat{T}_e/t_n) \sim \mathrm{Exp}(r_n)$, where $r_n := t_n^{-1}$ and the T_e's are i.i.d. In particular, $X^{(n)}(t) = \widehat{X}^{(n)}(a_n(t))$. If (i) holds, then we have $r_n \,\mathbb{E}\, d^{(n)}(V_1, V_2) = O(1)$, and thus Theorem 5.3 implies that the family of processes $X^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$. Similarly, if one of the other conditions holds, then the same conclusion holds by Corollary 5.4.

Define $b_n(t) := -t_n \ln(\max(1 - t/t_n, 0))$. Observe that $b_n(t)$ converges to the identity function uniformly on each compact interval $[0, s]$, $s \in \mathbb{R}_+$. Thus, since $\widehat{X}^{(n)}(t) = X^{(n)}(b_n(t))$ for $t < t_n$, our claim follows. ☐

6. Fixing the proof of tightness in [8]

Let $\mu = (\mu(k), k \geqslant 0)$ be a probability distribution on the non-negative integers satisfying $\sum_{k=0}^{\infty} k\mu(k) = 1$. In addition, we always implicitly assume that $\mu(0) + \mu(1) < 1$ (or equivalently, $\mu(0) > 0$), and that μ is aperiodic. We say that μ belong to the domain of attraction of a stable law of index $\alpha \in (1, 2]$ if for a sequence $(Y_i)_{i\geqslant 1}$ of i.i.d. random variables with distribution μ, there exists a sequence of positive real numbers $(B_n)_{n\geqslant 1}$ such that $B_n \to \infty$ and

$$\frac{Y_1 + Y_2 + \cdots + Y_n - n}{B_n} \xrightarrow{\mathrm{d}} Z_\alpha, \quad \text{as} \quad n \to \infty, \tag{6.1}$$

where Z_α is a random variable with Laplace transform $\mathbb{E}[\exp(-\lambda Z_\alpha)] = \exp(\lambda^\alpha)$ whenever $\alpha \in (1, 2)$, and $\mathbb{E}[\exp(-\lambda Z_2)] = \exp(\lambda^2/2)$ if $\alpha = 2$, for every $\lambda > 0$. (See [17, Section XVII.5], and note that since $Y_i \geqslant 0$, we may normalize B_n such that (6.1) holds.) In particular, for $\alpha = 2$, we have that Z_2 is distributed as a standard Gaussian random variable. Furthermore, if (6.1) holds, then B_n is of order about $n^{1/\alpha}$; more precisely, $B_n/n^{1/\alpha}$ is a slowly varying sequence.

Let $\mathcal{T}^{(n)}_{\mathrm{GW}}$ be a critical Galton–Watson tree conditioned on having n vertices and whose offspring distribution μ belongs to the domain of attraction of a stable law. Conditioned on the tree, each edge e is deleted at some random time $\widehat{T}_e = \frac{n}{B_n}(1 - U_e)$, independent of all other edges, where as in [8, Section 1], we assume that the U_e are i.i.d. with a common $U(0,1)$ distribution; hence $\widehat{T}_e \sim U(0, n/B_n)$. Let $\widehat{X}^{(n)}(t) = (\widehat{X}^{(n)}_i(t))_{i\geqslant 1}$ be the associated fragmentation process of normalized component sizes, as defined before Corollary 5.5. In the notation of [8, (3)], $\widehat{X}^{(n)}(t) = \mathbf{F}^{(\alpha)}_n(t)$.

Corollary 6.1. *The family of processes $\widehat{X}^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$.*

Proof. $H^{(n)}$ be the height of $\mathcal{T}^{(n)}_{\mathrm{GW}}$. Then, $D^{(n)} \leqslant 2H^{(n)}$, where $D^{(n)}$ is the diameter of $\mathcal{T}^{(n)}_{\mathrm{GW}}$. It follows from [23, Theorem 2] that $\mathbb{E}[D^{(n)}] = O(n/B_n)$. Therefore, our claim follows from Corollary 5.5(ii) (with $t_n = n/B_n$). Alternatively, it follows in the same way from Corollary 5.5(iv) that is suffices to show that $(B_n/n)H^{(n)} = O_{\mathrm{p}}(1)$, which follows from the convergence in distribution of the height process (or of the contour process) [15, Theorem 3.1], see also [22, Theorem 3]. Or from Corollary 5.5(v), since the total path length $\Upsilon^{(n)}$ of $\mathcal{T}^{(n)}$ satisfies $\Upsilon^{(n)} = O_{\mathrm{p}}(n^2/B_n)$, by [14, Corollary 4.11]. $\square$

Remark 6.2. Corollary 6.1 addresses a gap in the tightness argument of [8, Theorem 1]. Specifically, this issue originated from [8, Lemma 5], which is used in the proof of Theorem [8, Theorem 1]. While the convergence of the finite-dimensional distributions in [8, Lemma 5] (and consequently in Theorem [8, Theorem 1]) is correct, the tightness proof within [8, Lemma 5] contains a gap due to the application of the unfortunately incorrect [12, Lemma 22]. We may now instead use Corollary 6.1.

Note that [8] considers the space $\mathbb{S}^{\downarrow}_{<\infty}$ defined in (3.5). However, in this particular case, since $\widehat{X}^{(n)}(t) \in \mathbb{S}^{\downarrow}_{1}$ for every $t \geqslant 0$ (a.s.) and [8] shows the existence of a finite-dimensional limit $\widehat{X}(t) \in \mathbb{S}^{\downarrow}_{1}$, tightness in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$ implies convergence $\widehat{X}_n(t) \to \widehat{X}(t)$ in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{\leqslant 1})$, which by Lemma 3.5 and Remark 3.7 implies convergence in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{1})$ and in $D(\mathbb{R}_+, \mathbb{S}^{\downarrow}_{<\infty})$. $\triangle$

7. Fixing the proof of [12, (2) of Theorem 3]

As we said in the introduction, one motivation for the present note is that a compactness lemma in [12] is incorrect. The lemma concerns subsets of $D(I, \ell^p_{\geqslant 0})$, where I is an interval, but the problem remains for functions with values in a finite-dimensional space, as seen by the following counterexample.

Example 7.1. For $n \geqslant 2$, let g_n be the distribution function of a uniform distribution on $[\frac{1}{2} - \frac{1}{n}, \frac{1}{2}]$, i.e., $g_n(\frac{1}{2} - \frac{1}{n}) = 0$, $g_n(\frac{1}{2}) = 1$ with g_n linear in between and constant outside this interval. Let $f_n(x)$ be the vector $(g_n(x), 1 - g_n(x), 0, 0, \dots)$. It is then easy to see that the set $\{f_n, n \geqslant 2\}$ satisfies the assumptions of [12, Lemma 22] (with $I = [0, 1]$, $p = 1$, and $M = 2$ for every ε). Nevertheless, the set is not relatively compact since no subsequence of (f_n) converges. To see this it suffices to consider the first components $g_n(x)$; it is easy to see that no subsequence can converge since the set of continuous functions is closed in $D(I, \mathbb{R})$, and on this subset, convergence in the Skorohod topology is equivalent to uniform convergence. (The error in the proof in [12] is the assertion that any sequence of [uniformly] bounded non-decreasing functions on I has an accumulation point in $D(I, \mathbb{R})$. This is correct for the weaker M_1 topology, see e.g. [30, Corollary 12.5.1], but not for the standard J_1 topology; a counterexample is provided by g_n above.) $\triangle$

We now provide a corrected proof for the claim made in [12, (2) of Theorem 3]. Let $\mathcal{T}^{(n)}_{\mathrm{Cayley}}$ be a uniform Cayley tree with n vertices (uniformly labelled tree on $\{1, \dots, n\}$). Conditioned on the tree, each edge e is deleted at some random time $\widehat{T}_e = \sqrt{n}(1 - U_e)$, independent of all other edges. As in [12, Section 3.2.3], we assume that the U_e are i.i.d. with a common $U(0, 1)$ distribution. Let $\widehat{X}^{(n)}(t) = (\widehat{X}^{(n)}_i(t))_{i \geqslant 1}$ be the associated fragmentation process of normalized component sizes, as defined before Corollary 5.5.

In [12], the process $\widehat{X}^{(n)}(t) = (\widehat{X}_i^{(n)}(t))_{i\geqslant 1}$ is encoded using a process $\mathbf{y}_+^{n,(t)} = (\mathbf{y}_+^{n,(t)}(x))_{x\in[0,1]}$, as defined in [12, (7)]. More precisely, $\widehat{X}^{(n)}(t)$ is precisely the sequence of excursion lengths of $\mathbf{y}_+^{n,(t)}$ above its minimum sorted in decreasing order as time t evolves. Let $\mathbf{y}_+^{(t)}(x) := \mathbf{e}(x) - tx$, for $x \in [0,1]$ and $t \geqslant 0$, where $\mathbf{e}$ is the normalised Brownian excursion (with unit length). Let $\widehat{X}(t)$ be the sequence of excursion lengths of $\mathbf{y}_+^{(t)}$ above its minimum sorted in decreasing order. Following the notation of [8], $\widehat{X}^{(n)}(t) = \boldsymbol{\gamma}^{+,n}(t)$ and $\widehat{X}(t) = \boldsymbol{\gamma}^{+}(t)$.

Theorem 7.2. $\widehat{X}^{(n)}(t) \xrightarrow{\mathrm{d}} \widehat{X}(t)$ *in* $D(\mathbb{R}_+, \mathbb{S}_1^{\downarrow})$.

Proof. Define $\overline{\mathbf{y}}_+^{n,(t)}(x) := \sup_{s\in[0,x]}(-\mathbf{y}_+^{n,(t)}(s))$ and $\overline{\mathbf{y}}_+^{(t)}(x) := \sup_{s\in[0,x]}(-\mathbf{y}_+^{(t)}(s))$, for $x \in [0,1]$. Following [7, Section 3.1], given an increasing path g in $D([0,1],\mathbb{R})$, we write $\mathbf{F}(g)$ for the sequence of the lengths of the intervals components of the complement of the support of the Stieltjes measure $\mathrm{d}g$, arranged in the decreasing order. In particular, $\widehat{X}^{(n)}(t) = \mathbf{F}(\overline{\mathbf{y}}_+^{n,(t)})$ and $\widehat{X}(t) = \mathbf{F}(\overline{\mathbf{y}}_+^{(t)})$. Then, using [12, Theorem 10] together with [7, Lemma 4 and Lemma 7], one can show the finite-dimensional convergence $\widehat{X}^{(n)}(t) \xrightarrow{\mathrm{f.d.}} \widehat{X}(t)$ in $D(\mathbb{R}_+, \mathbb{S}_1^{\downarrow})$.

On the other hand, Corollary 5.5 (with $t_n = \sqrt{n}$), shows that the family of processes $\widehat{X}^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$, because it was shown by [27] that the diameter $D^{(n)}$ of $\mathcal{T}^{(n)}$ satisfies $\mathbb{E}[D^{(n)}] = O(\sqrt{n})$, see also [2, Corollary 1.3]. Alternatively, we can use that $D^{(n)} = O_{\mathrm{p}}(\sqrt{n})$ by [4; 5], or that the total path length $\Upsilon^{(n)}$ of $\mathcal{T}^{(n)}$ satisfies $\mathbb{E}[\Upsilon^{(n)}] = O(n^{3/2})$ by [28; 29], see also [19, Theorem 3.4]; furthermore, we may also use $\mathbb{E}\, d^{(n)}(V_1, V_2) = O(n^{1/2})$ which was proved in [25]; again see also [19, Theorem 3.4]. Combining this and finite-dimensional convergence, we obtain $\widehat{X}^{(n)}(t) \xrightarrow{\mathrm{d}} \widehat{X}(t)$ in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$. Finally, since $\widehat{X}(t) \in \mathbb{S}_1^{\downarrow}$ for every t a.s. (see [7, Lemma 7]), our claim follows by Lemma 3.5. $\square$

8. Application to the fragmentation of trees with a specified degree sequence

For $n \in \mathbb{N}$, a sequence $\mathbf{s}_n = (N_n(i))_{i\geqslant 0}$ of non-negative integers is the degree sequence of some finite rooted plane tree if and only if $\sum_{i\geqslant 0} N_n(i) = 1 + \sum_{i\geqslant 0} iN_n(i) < \infty$. Then, a random tree $\mathcal{T}_{\mathbf{s}}^{(n)}$ with given degree sequence $\mathbf{s}_n$ is a random variable whose law is uniform on the set of rooted plane trees with $|\mathbf{s}_n| := \sum_{i\geqslant 0} N_n(i)$ vertices amongst which $N_n(i)$ have i offspring for every $i \geqslant 0$, and therefore with $\sum_{i\geqslant 0} iN_n(i)$ edges.

Define $\sigma_n^2 := \sum_{i\geqslant 1} i(i-1)N_n(i)$ and let b_n be a sequence of non-negative real number such that $b_n \to \infty$, $|\mathbf{s}_n|/b_n \to \infty$ and $b_n/\sigma_n \to 1$. Conditioned on the tree, each edge e is deleted at some random time $\widehat{T}_e = \frac{|\mathbf{s}_n|}{b_n}(1 - U_e)$, independent of all other edges. We assume that the U_e are i.i.d. with a uniform $U(0,1)$ distribution. Let $\widehat{X}^{(n)}(t) = (\widehat{X}_i^{(n)}(t))_{i\geqslant 1}$ denote the associated fragmentation process of normalized component sizes, as defined prior to Corollary 5.5.

Corollary 8.1. *The family of processes* $\widehat{X}^{(n)}(t)$ *is tight in* $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$.

Proof. For a vertex v in $\mathcal{T}_{\mathbf{s}}^{(n)}$, we denote its height by $h^{(n)}(v)$. Without loss of generality, by considering n large enough if necessary, we assume that $|\mathbf{s}_n| \geqslant b_n$ and

$\sigma_n > 0$. Let $\Delta_n := \max\{i \geqslant 0 : N_n(i) > 0\}$. Then, since $\sigma_n > 0$, we have $\Delta_n \geqslant 2$ and $\sigma_n \geqslant \Delta_n$. Let V_1 and V_2 be two independent uniformly random vertices of $\mathcal{T}_{\mathbf{s}}^{(n)}$. Observe that $d^{(n)}(V_1, V_2) \leqslant h^{(n)}(V_1) + h^{(n)}(V_2)$. It follows from [24, Proposition 4.5] that there exists two universal constants $c_1, c_2 > 0$ such that

$$\mathbb{P}\big(h^{(n)}(V_1) \geqslant x|\mathbf{s}_n|/b_n\big) \leqslant c_1 e^{-c_2 x \frac{\sigma_n}{b_n}} \tag{8.1}$$

uniformly for $x > 0$ and $n \in \mathbb{N}$. (See also [1, Theorem 1.1] for a similar bound.) Thus, $\mathbb{E}d^{(n)}(V_1, V_2) \leqslant 2\,\mathbb{E}\,h^{(n)}(V_1) = O(|\mathbf{s}_n|/b_n)$ and our claim follows from Corollary 5.5(i). ☐

Remark 8.2. Recently, the fragmentation process of trees with a given degree sequence has been studied in [9], where the convergence $\widehat{X}^{(n)}(t) = (\widehat{X}_i^{(n)}(t))_{i\geqslant 1}$ was established. However, at the time of writing this note, a problem existed in the tightness argument, specifically originating from the unfortunately incorrect [12, Lemma 22]. Now, Corollary 8.1 completes the proof. As in Remark 6.2, the result holds also in $D(\mathbb{R}_+, \mathbb{S}_1^{\downarrow})$ and $D(\mathbb{R}_+, \mathbb{S}_{<\infty}^{\downarrow})$ by Lemma 3.5 and Remark 3.7. △

9. Weighted fragmentations of a tree

In Sections 5–8 we have studied fragmentations of trees, where we have measured the sizes of components by their number of vertices. More generally, we may assume that we are given, for each n, a random tree $\mathcal{T}^{(n)}$ with a nonrandom vertex set $\mathcal{V}^{(n)} = \mathcal{V}(\mathcal{T}^{(n)})$ (finite or infinite), and a family of nonnegative weights $\mathbf{w}^{(n)} = (w_j^{(n)} : j \in \mathcal{V}^{(n)})$ with total mass $\sum_j w_j^{(n)} < \infty$. We consider random fragmentation of the tree $\mathcal{T}^{(n)}$ as above, with edges e deleted at i.i.d. times T_e that are either $\mathrm{Exp}(r_n)$ or $U(0, t_n)$. We then define the fragmentation process $Y^{(n)}(t) = (Y_i^{(n)}(t))_{i=1}^{\infty}$ by letting

$$Y_i^{(n)}(t) := \sum_{j \in \mathcal{T}_i^{(n)}(t)} w_j^{(n)}, \tag{9.1}$$

the total weight of the component $\mathcal{T}_i^{(n)}(t)$, for $i = 1, \dots, \nu^{(n)}(t)$, where we now arrange the components in order of decreasing weight.

For simplicity we will only consider the case when the weight sequence $\mathbf{w}^{(n)}$ is a probability distribution on $\mathcal{V}^{(n)}$, i.e., $\sum_j w_j^{(n)} = 1$. To stress this we use the notation $\mathbf{p}^{(n)} = (p_j^{(n)} : j \in \mathcal{V}^{(n)})$ instead of $\mathbf{w}^{(n)}$. (We leave the modifications for the general case to the reader.)

Theorem 9.1. *With assumptions as above, Theorems 5.1 and 5.2 still hold (with $Y^{(n)}$ and Y instead of $X^{(n)}$ and X), and Theorem 5.3 holds if V_1 and V_2 now are independent vertices of $\mathcal{T}^{(n)}$ with the distribution $\mathbf{p}^{(n)}$.*

Proof. The proofs are essentially the same.

First, for Theorem 5.1, instead of (5.2) we obtain, in the same way,

$$Q^{(n)}(T_e-) - Q^{(n)}(T_e) = \sum_{v,w \in \mathcal{V}^{(n)}} p_v^{(n)} p_w^{(n)} \mathbf{1}\big\{e \text{ lies in the path between } v \text{ and } w \text{ in } \mathcal{T}^{(n)}(0), \text{ and this path remains intact in } \mathcal{T}^{(n)}(T_e-)\big\}. \tag{9.2}$$

The rest of the proof is the same as before.

The proof of Theorem 5.2 is unchanged.

In the proof of Theorem 5.3, we obtain, arguing as in (5.11)–(5.12),

$$1 - \mathbb{E}\left(Q^{(n)}(t) \mid \mathcal{T}^{(n)}\right) \leqslant \sum_{v,w\in\mathcal{V}^{(n)}} p_v^{(n)} p_w^{(n)} r_n t d^{(n)}(v,w) = r_n t\, \mathbb{E}\left(d^{(n)}(V_1,V_2) \mid \mathcal{T}^{(n)}\right), \tag{9.3}$$

which shows (5.9). □

Remark 9.2. The parts of Corollaries 5.4 and 5.5 that use $D^{(n)}$ still hold. However, the parts involving the total path length $\Upsilon^{(n)}$ require us to instead consider the weighted path length $\sum_{v\in V^{(n)}} p_v^{(n)} d^{(n)}(\rho,v)$. Equivalently, these parts can be stated in terms of the average depth $\mathbb{E}\, d^{(n)}(\rho,V)$, where V is a random vertex with distribution $\mathbf{p}^{(n)}$. The part involving $\mathbb{E}\, d^{(n)}(V_1,V_2)$ also holds, but we must consider two independent random vertices V_1 and V_2 of $\mathcal{T}^{(n)}$ chosen according to the distribution $\mathbf{p}^{(n)}$. The details are left to the reader. △

10. Application to the fragmentation of p-trees

For $n \geqslant 1$, let $\mathbf{p}^{(n)} = (p_{ni} : i \geqslant 1)$ be a ranked probability distribution on the positive integers. That is, $p_{n1} \geqslant p_{n2} \geqslant 0$ and $\sum_{i\geqslant 1} p_{ni} = 1$. The support of $\mathbf{p}^{(n)}$ is $\mathcal{V}^{(n)} := \{i \geqslant 1 : p_{ni} > 0\}$. Let $Y_{n,0}, Y_{n,1}, \dots$ be i.i.d. random variables with common distribution $\mathbf{p}^{(n)}$. Define a random discrete tree $\mathcal{T}_{\mathbf{p}}^{(n)}$ to have vertex set $\mathcal{V}^{(n)}$ and undirected edges $\{\{Y_{n,j-1}, Y_{n,j}\} : Y_{n,j} \notin \{Y_{n,0},\dots,Y_{n,j-1}\},\ j \geqslant 1\}$. The tree $\mathcal{T}_{\mathbf{p}}^{(n)}$ is known as the birthday tree with parameter $\mathbf{p}^{(n)}$, or simply as the $\mathbf{p}^{(n)}$-tree (see, e.g., [13], [6, Section 3.1], [26, Section 10.2], and references therein). Observe that when $p_{ni} = 1/n$, for $1 \leqslant i \leqslant n$, the $\mathbf{p}^{(n)}$-tree is a uniform random rooted tree with n vertices (the Cayley tree of size n), discussed above in Section 7.

Define $\sigma_{\mathbf{p}^{(n)}} := \sqrt{\sum_{i\geqslant 1} p_{ni}^2}$. Following [6, Section 3.1], conditioned on $\mathcal{T}_{\mathbf{p}}^{(n)}$, we delete each edge e at some random time $\widehat{T}_e = \sigma_{\mathbf{p}^{(n)}}^{-1}(1-U_e)$, independent of all other edges, where we assume that the U_e are i.i.d. with a common $U(0,1)$ distribution. We define as in Section 9 the fragmentation process $\widehat{Y}^{(n)}(t) = (\widehat{Y}_i^{(n)}(t))_{i\geqslant 1}$ as the sequence of $\mathbf{p}^{(n)}$-measures of the tree-components of the fragmented forest at time t, arranged in decreasing order. So, $\widehat{Y}^{(n)}(t)$ records the $\mathbf{p}^{(n)}$-measures of components obtained when each edge is deleted with probability $t\sigma_{\mathbf{p}^{(n)}}$.

Corollary 10.1. *Suppose that $\sigma_{\mathbf{p}^{(n)}} \to 0$. Then, the family of processes $\widehat{Y}^{(n)}(t)$ is tight in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^{\downarrow})$.*

Proof. This follows from the weighted version of Corollary 5.5 (see Remark 9.2) provided that $\mathbb{E}d^{(n)}(V_1,V_2) = O(\sigma_{\mathbf{p}^{(n)}}^{-1})$, where V_1 and V_2 are two independent random vertices of $\mathcal{T}_{\mathbf{p}}^{(n)}$ chosen according to the distribution $\mathbf{p}^{(n)}$. But this last claim follows from Proposition A.1 in Appendix A, thus concluding our proof. □

Remark 10.2. Define

$$\mathbb{S}_2^{\downarrow} := \left\{(x_i)_1^\infty \in \mathbb{S}^{\downarrow} : 0 \leqslant \sum_{i=1}^{\infty} x_i^2 \leqslant 1\right\} \tag{10.1}$$

and

$$\Theta := \Big\{(x_i)_1^\infty \in \mathbb{S}_2^\downarrow : \sum_{i=1}^\infty x_i^2 < 1 \text{ or } \sum_{i=1}^\infty x_i = \infty\Big\}. \tag{10.2}$$

Aldous and Pitman [6, Proposition 13] proved that if there exists a sequence $\boldsymbol{\theta} = (\theta_i)_1^\infty \in \Theta$ such that

$$\sigma_{\mathbf{p}^{(n)}} \to 0 \quad \text{and} \quad \frac{p_{ni}}{\sigma_{\mathbf{p}^{(n)}}} \to \theta_i, \quad i \geqslant 1, \tag{10.3}$$

then, $\widehat{Y}^{(n)}(t) \xrightarrow{\text{f.d.}} \widehat{Y}(t)$ to some stochastic process $\widehat{Y}(t)$ in $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$. On the other hand, we have that $\widehat{Y}^{(n)}(t) \in \mathbb{S}_1^\downarrow$ and $\widehat{Y}(t) \in \mathbb{S}_1^\downarrow$ for every $t \geqslant 0$ (a.s.) ([6, Lemma 12]). Then, tightness in $D(\mathbb{R}_+, \mathbb{S}_{\leqslant 1}^\downarrow)$ (Corollary 10.1) implies $\widehat{Y}^{(n)}(t) \xrightarrow{\text{d}} \widehat{Y}(t)$ in $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$ by Lemma 3.5. The limit $\widehat{Y}(t)$ is the fragmentation process of the so-called *Inhomogeneous continuum random tree* with parameter $\boldsymbol{\theta}$; see [6, Section 4] for its definition. △

Remark 10.3. Define $M^{(n)}(t) := \widehat{Y}^{(n)}(\sigma_{\mathbf{p}_n}^{-1} e^{-t})$, for $0 \leqslant t < \infty$. Then, $M^{(n)}(t)$ is an additive coalescent with initial state $\mathbf{p}_n$; see [6, Proposition 1]. Indeed, if we define $M(t) := \widehat{Y}(e^{-t})$, for $-\infty < t < \infty$, then $M(t)$ is an (*eternal*) additive coalescent; see [6, Theorem 10].

In [7], a specific case is analysed where $p_{nn} > 0$ and $p_{ni} = 0$ for $i > n$. [7, Theorem 1] proves that the process $(M^{(n)}(-\log \sigma_{\mathbf{p}_n} - \log t), 0 < t < \sigma_{\mathbf{p}_n}^{-1})$ converges in the sense of finite dimensional distributions, as $n \to \infty$, in $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$, to a fragmentation process associated with certain bridges with exchangeable increments. Observe that $M^{(n)}(-\log \sigma_{\mathbf{p}_n} - \log t) = \widehat{Y}^{(n)}(t)$. Then, Corollary 10.1 and Lemma 3.5 imply the tightness of this sequence of processes in $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$. As a consequence, the convergence in [7, Theorem 1] can be strengthened to convergence in distribution in $D(\mathbb{R}_+, \mathbb{S}_1^\downarrow)$. △

Appendix A. An exponential tail bound

In this appendix, we prove an exponential tail bound, used above, on the distance between two independent random vertices of a $\mathbf{p}_n$-tree, chosen independently according to the distribution $\mathbf{p}_n$. For notational simplicity, we will drop the n dependence from the notation introduced in Section 10 and write $\mathbf{p} = (p_i : i \geqslant 1)$ for a ranked probability distribution on the positive integers, and $\mathcal{T}_{\mathbf{p}}$ for the associated $\mathbf{p}$-tree.

Proposition A.1. *Let V_1 and V_2 be two independent random vertices of $\mathcal{T}_{\mathbf{p}}$ chosen according to the distribution $\mathbf{p}$. Then, for all $x \geqslant 8$,*

$$\mathbb{P}(d(V_1, V_2) \geqslant x\sigma_{\mathbf{p}}^{-1}) \leqslant e^{-\frac{x^{1/3}}{3\sigma_{\mathbf{p}}}} + 6e^{-\frac{x^{2/3}}{6}}. \tag{A.1}$$

The proof of Proposition A.1 is based on the Poisson embedding described in [13, Section 4.1]. Let N be a homogeneous Poisson process on $[0,\infty) \times [0,1]$ of rate 1 per unit area, with atoms $((S_j, U_j))_{j \geqslant 0}$ such that $0 < S_0 < S_1 < \cdots$. Then, $(S_j)_{j\geqslant 0}$ are the atoms of a homogeneous Poisson process on $[0,\infty)$ of rate 1 per unit length, and the $(U_j)_{j\geqslant 0}$ are i.i.d. with uniform distribution on $[0,1]$, independent of the $(S_j)_{j\geqslant 0}$. Define

$$N(t) := N([0,t] \times [0,1]) \quad \text{and} \quad N(t-) := N([0,t) \times [0,1]). \tag{A.2}$$

Then, partition the interval $[0,1]$ into sub-intervals $I_1, I_2, \dots$ such that the length of I_i is p_i, for $i \geqslant 0$. For $j \geqslant 0$ define

$$Y_j := \sum_{i \geqslant 1} i \mathbf{1}_{\{U_j \in I_i\}}. \tag{A.3}$$

So, $Y_0, Y_1, \dots$ are i.i.d. random variables with common distribution $\mathbf{p}$.

From now on, assume that the $\mathbf{p}$-tree $\mathcal{T}_{\mathbf{p}}$ is constructed using the random variables defined above; see Section 10. Let R_1 be the index of the first repeated value in the sequence $(Y_j)_{j \geqslant 0}$ and let $T_1 := \inf\{t \geqslant 0 : N(t) > R_1\}$. Then, as a consequence of [13, Corollary 3],

$$N(T_1-) = R_1 \overset{\mathrm{d}}{=} d(V_1, V_2) + 1, \tag{A.4}$$

where V_1 and V_2 are two independent random vertices of $\mathcal{T}_{\mathbf{p}}$ chosen according to the distribution $\mathbf{p}$.

The proof of Proposition A.1 makes use of the following result, which is a straightforward consequence of [13, (26) and (29) in Lemma 9].

Lemma A.2. *For* $0 \leqslant t < (2p_1)^{-1}$, *we have* $\mathbb{P}(T_1 > t) \leqslant e^{-\frac{t^2}{6}\sigma_{\mathbf{p}}^2}$.

Proof. Note that the right-hand side of [13, (29) in Lemma 9] (with $\boldsymbol{\theta} = \mathbf{p}$) can be written as $(t\sigma_{\mathbf{p}})^2 \frac{tp_1}{3(1-tp_1)} \leqslant \frac{t^2}{3}\sigma_{\mathbf{p}}^2$ because $0 \leqslant t < (2p_1)^{-1}$. Thus, our claim follows from [13, (26) and (29)]. □

Proof of Proposition A.1. It follows from (A.4) that, for $h \geqslant 0$ and $0 \leqslant t \leqslant h$,

$$\mathbb{P}(d(V_1, V_2) \geqslant h) \leqslant \mathbb{P}(N(t) \geqslant h) + \mathbb{P}(T_1 > t). \tag{A.5}$$

On the other hand, recalling the following Chernoff-type inequality (see, e.g., [20, Corollary 2.4 and Remark 2.6] or [11, Section 2.2, particularly page 23]), we have that, for $t \leqslant h$,

$$\mathbb{P}(N(t) \geqslant h) \leqslant e^{-t((h/t)\log(h/t) - h/t + 1)}. \tag{A.6}$$

In particular, if $h = 2t$, we have that $t((h/t)\log(h/t) - h/t + 1) = t(2\log 2 - 1) \geqslant t/3$ and thus

$$\mathbb{P}(N(t) \geqslant 2t) \leqslant e^{-\frac{t}{3}}. \tag{A.7}$$

Fix $c \geqslant 2$. If $0 \leqslant c\sigma_{\mathbf{p}}^{-1} < (2p_1)^{-1}$, then taking $t = c\sigma_{\mathbf{p}}^{-1}$ and $h = 2t$, by Lemma A.2, (A.5) and (A.7) we have that

$$\mathbb{P}(d(V_1, V_2) > 2c\sigma_{\mathbf{p}}^{-1}) \leqslant e^{-\frac{c}{3\sigma_{\mathbf{p}}}} + e^{-\frac{c^2}{6}}. \tag{A.8}$$

Now assume that $c\sigma_{\mathbf{p}}^{-1} \geqslant (2p_1)^{-1}$. For any positive real $k \geqslant 2$, if at least two of the variables $U_0, \dots, U_{\lfloor k \rfloor}$ lie in the interval I_1 then $N(T_1-) \leqslant k+1$. Thus, by (A.4),

$$\begin{aligned} \mathbb{P}(d(V_1, V_2) \geqslant k) &\leqslant \mathbb{P}(\#\{j \in \{0, \dots, \lfloor k \rfloor\} : U_j \in I_1\} \leqslant 1) \\ &= (1-p_1)^{\lfloor k \rfloor}(1 - p_1 + (\lfloor k \rfloor + 1)p_1) \\ &\leqslant e^{-p_1 \lfloor k \rfloor}(1 + \lfloor k \rfloor p_1), \end{aligned} \tag{A.9}$$

where we have used that $\#\{j \in \{0, \dots, \lfloor k \rfloor\} : U_j \in I_1\} \sim \text{Binomial}(\lfloor k \rfloor + 1, p_1)$ and that $1 - p_1 \leqslant e^{-p_1}$. Since $\sigma_{\mathbf{p}} \leqslant 1$, it follows that for $k \geqslant 2c\sigma_{\mathbf{p}}^{-1} \geqslant 4$,

$$\mathbb{P}(d(V_1, V_2) \geqslant k) \leqslant 2kp_1 e^{-\frac{kp_1}{2}}, \tag{A.10}$$

where we have used the lower bound on k to deduce that $\lfloor k \rfloor > k/2$ and that $kp_1 \geqslant 2p_1 c\sigma_{\mathbf{p}}^{-1} \geqslant 1$ and so $1 + \lfloor k \rfloor p_1 \leqslant 2kp_1$.

Next, take $k = yc\sigma_{\mathbf{p}}^{-1}$, for $y \geqslant 4$. Recall that we have assumed that $c\sigma_{\mathbf{p}}^{-1} \geqslant (2p_1)^{-1}$ and thus, $kp_1 \geqslant y/2$. Since $z \mapsto 2ze^{-z/2}$ is decreasing for $z \geqslant 2$, the bound (A.10) then implies that

$$\mathbb{P}(d(V_1, V_2) \geqslant yc\sigma_{\mathbf{p}}^{-1}) \leqslant ye^{-\frac{y}{4}}. \tag{A.11}$$

To conclude the proof, we combine (A.8) and (A.11) to obtain a bound that does not depend on the value of p_1. Take $x \geqslant 8$, let $c = x^{1/3} \geqslant 2$ and $y = x^{2/3} \geqslant 4$. Then, $2c \leqslant x$ and $yc = x$. Thus, regardless of the value of p_1, either (A.8) or (A.11) applies, so we obtain that

$$\begin{aligned}\mathbb{P}(d(V_1, V_2) \geqslant x\sigma_{\mathbf{p}}^{-1}) &= \mathbb{P}(d(V_1, V_2) \geqslant yc\sigma_{\mathbf{p}}^{-1}) \\ &\leqslant e^{-\frac{c}{3\sigma_{\mathbf{p}}}} + e^{-\frac{c^2}{6}} + ye^{-\frac{y}{4}} \\ &= e^{-\frac{x^{1/3}}{3\sigma_{\mathbf{p}}}} + e^{-\frac{x^{2/3}}{6}} + x^{2/3}e^{-\frac{x^{2/3}}{4}}. \end{aligned} \tag{A.12}$$

Finally, our claim follows from (A.12) and the observation that $ze^{-z/4} \leqslant 5e^{-z/6}$ for all $z \geqslant 0$. □

References

[1] Louigi Addario-Berry, Anna Brandenberger, Jad Hamdan, Célin Kerriou. Universal height and width bounds for random trees. *Electron. J. Probab.* **27** (2022), Paper No. 118, 24 pp. MR 4479914

[2] Louigi Addario-Berry, Luc Devroye and Svante Janson. Sub-Gaussian tail bounds for the width and height of conditioned Galton–Watson trees. *Ann. Probability* **41** (2013), 1072–1087. MR 3077536

[3] David Aldous. Stopping times and tightness. *Ann. Probability* **6**:2 (1978), 335–340. MR 0474446

[4] David Aldous. The continuum random tree II. An overview. *Stochastic Analysis (Durham, 1990)* **167** (1991), 23–70. MR 1166406

[5] David Aldous. The continuum random tree III. *Ann. Probability* **21**:1 (1993), 248–289. MR 1207226

[6] David Aldous and Jim Pitman. Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent. *Probab. Theory Related Fields* **118**:4 (2000), 455–482. MR 1808372

[7] Jean Bertoin. Eternal additive coalescents and certain bridges with exchangeable increments. *Ann. Probability* **29** (2001), 344–360. MR 1825153

[8] Gabriel Berzunza Ojeda and Cecilia Holmgren. Invariance principle for fragmentation processes derived from conditioned stable Galton–Watson trees. *Bernoulli* **29**:4 (2023), 2745–2770. MR 4632119

[9] Gabriel Berzunza Ojeda, Cecilia Holmgren and Paul Thévenin. Convergence of trees with a given degree sequence and of their associated laminations. *arXiv:2111.07748* (2025+).

[10] Patrick Billingsley. *Convergence of Probability Measures.* 2nd ed., John Wiley & Sons, New York, 1999. MR 0233396

[11] Stéphane Boucheron, Gábor Lugosi and Pascal Massart. *Concentration Inequalities.* Oxford University Press, Oxford, 2013. MR 3185193

[12] Nicolas Broutin and Jean-François Marckert. A new encoding of coalescent processes: applications to the additive and multiplicative cases. *Probab. Theory Related Fields* **166**:1-2 (2016), 515–552. MR 3547745
[13] Michael Camarri and Jim Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electron. J. Probab.* **5**:2 (2000), 18 pp. MR 1741774
[14] Jean-François Delmas, Jean-Stéphane Dhersin and Marion Sciauveau. Cost functionals for large (uniform and simply generated) random trees. *Electron. J. Probab.* **23** (2018), Paper No. 87, 36 pp. MR 3858915
[15] Thomas Duquesne. A limit theorem for the contour process of conditioned Galton–Watson trees. *Ann. Probab.* **31**:2 (2003), 996–1027. MR 1964956
[16] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes. Characterization and Convergence.* John Wiley & Sons, Inc., New York, 1986. MR 0838085
[17] William Feller. *An Introduction to Probability Theory and its Applications. Vol. II.* John Wiley & Sons, Inc., New York-London-Sydney, 1971. MR 270403
[18] Allan Gut. *Probability: A Graduate Course.* 2nd ed., Springer, New York, 2013. MR 2977961
[19] Svante Janson. The Wiener index of simply generated random trees. *Random Structures Algorithms*, **22** (2003), 337–358. MR 1980963
[20] Svante Janson, Tomasz Łuczak & Andrzej Ruciński. *Random Graphs.* John Wiley & Sons, New York, 2000.
[21] Olav Kallenberg. *Foundations of Modern Probability.* 2nd ed., Springer, New York, 2002. MR 1876169
[22] Igor Kortchemski. A simple proof of Duquesne's theorem on contour processes of conditioned Galton–Watson trees. *Séminaire de Probabilités XLV*, Lecture Notes in Math. **2078**, pp. 537–558. Springer, Cham, 2013. MR 3185928
[23] Igor Kortchemski. Sub-exponential tail bounds for conditioned stable Bienaymé-Galton-Watson trees. *Probab. Theory Related Fields* **168**:1-2 (2017), 1–40. MR 3651047
[24] Cyril Marzouk. On scaling limits of random trees and maps with a prescribed degree sequence. *Ann. H. Lebesgue* **5** (2022), 317–386. MR 4443293
[25] Amram Meir and John W. Moon. The distance between points in random trees. *J. Combinatorial Theory* **8** (1970), 99–103. MR 263685
[26] Jim Pitman. *Combinatorial Stochastic Processes.* École d'Été de Probabilités de Saint-Flour XXXII – 2002. Lecture Notes in Math. **1875**, Springer-Verlag, Berlin, 2006. MR 2245368
[27] Alfréd Rényi and George Szekeres. On the height of trees. *J. Austral. Math. Soc.* **7** (1967), 497–507. MR 219440
[28] Lajos Takács. On the total heights of random rooted trees. *J. Appl. Probab.* **29**:3 (1992), 543–556.
[29] Lajos Takács. The asymptotic distribution of the total heights of random rooted trees. *Acta Sci. Math. (Szeged)*, **57**:1-4 (1993), 613–625. MR 1243312
[30] Ward Whitt. *Stochastic-Process Limits.* Springer-Verlag, New York, 2002. MR 1876437

Department of Mathematical Sciences, University of Liverpool, United Kingdom
Email address: gabriel.berzunza-ojeda@liverpool.ac.uk
URL: https://www.liverpool.ac.uk/mathematical-sciences/staff/gabriel-berzunza-ojeda/

Department of Mathematics, Uppsala University, Sweden
Email address: cecilia.holmgren@math.uu.se
URL: https://katalog.uu.se/empinfo/?id=N5-824

Department of Mathematics, Uppsala University, Sweden
Email address: svante.janson@math.uu.se
URL: https://www2.math.uu.se/∼svantejs