Random graphs with additional structure evolving in time

Rivka Mitchell

The Queen's College

University of Oxford

A thesis submitted for the degree of

DPhil in Mathematics

Trinity 2025

Declaration

This thesis is submitted to the University of Oxford for the degree of DPhil in Mathematics. I hereby declare that except for where specific reference is made to the work of others, the contents of this thesis are original.

Chapter 2 consists of joint work with Louigi Addario-Berry, Serte Donderwinkel, and Christina Goldschmidt. Chapter 3 consists of joint work with Anna Brandenberger, Serte Donderwinkel, Céline Kerriou, and Gábor Lugosi. Chapter 4 consists of joint work with Matthew Buckland, Brett Kolesnik, and Tomasz Przybyłowski.

Rivka Mitchell
Trinity Term 2025

Acknowledgments

I have been very lucky to have had so many fantastic mentors, colleagues, and friends over the last few years.

First and foremost I would like to thank my supervisor, Christina Goldschmidt. Through all the twists and turns snakes brought us, I don't think we ever had a meeting free of laughter. This thesis certainly would not have been possible without your generous support, mentorship, thoughtful advice, and kindness. Thank you as well to Louigi Addario-Berry for your mentorship since first taking me on as an undergraduate research student, and for introducing me to a wonderful group of probabilists. You both have an infectious love of mathematics, and have shown me the joys of solving problems with your friends.

Speaking of which — thank you to Céline Kerriou, Serte Donderwinkel, and Anna Brandenberger for the maths, but more importantly the hikes, great food, and long bike trips that we shared along the way. Thank you as well to my collaborators Matthew Buckland, Brett Kolesnik, Gábor Lugosi, and Tomasz Przybyłowski, and to the fantastic members of the probability group, among which I would like to especially thank David, Gabriel, Félix, Joost, and Kim. Further, many thanks to James Martin and Igor Kortchemski for being wonderful DPhil examiners. I greatly enjoyed our discussion, and am grateful for your thoughtful comments and questions.

Thank you to Jacob who encouraged me to keep at it from the beginning, and has held my hand all along the way.

And, last but certainly not least, thank you to my family — to my parents Naomi and Alistair, and to my goofs of siblings, Sacha and Tessa, who have played the key role of keeping things generally ridiculous at all times.

Funding

This thesis was supported by the EPSRC Centre for Doctoral Training in Mathematics of Random Systems: Analysis, Modelling and Simulation (EP/S023925/1), and the Marie Skłodowska-Curie RISE grant RandNET (Randomness and Learning in Networks), grant agreement ID 101007705.

Abstract

This thesis comprises three projects concerning random graphs and random processes on graphs.

In Chapter 2 we prove that if the offspring distribution of a Bienaymé tree is critical and admits a finite third moment, then under a suitable tail condition on the displacements, any globally centered size-conditioned discrete snake with finite global variance converges (upon suitable rescaling) to the Brownian snake driven by a normalized Brownian excursion. We also consider displacement distributions with heavier tails, for which we instead prove convergence in distribution to a variant of the hairy snake introduced by Janson and Marckert. We further give two applications of our main result. Namely, we obtain a scaling limit for the difference between the height function and the Łukasiewicz path of a size-conditioned critical Bienaymé tree; and we obtain a scaling limit for the difference between the height function of a size-conditioned critical Bienaymé tree and the height function of its associated looptree.

In Chapter 3 we prove a threshold for the existence of monotone increasing paths between all pairs of vertices in temporal random geometric graphs. Our result reveals that temporal connectivity appears when the edge density is significantly larger than that required for simple connectivity of the underlying graph. This contrasts with temporal connectivity thresholds for Erdős–Rényi random graphs which occur when the edge density is a constant multiple of that required for connectivity of the underlying graph. Our results hold for a wide family of soft random geometric graphs as well as the standard random geometric graph, and in general dimensions $d \geq 2$.

In Chapter 4 we prove rapid mixing for a lazy simple symmetric random walk on the set of Coxeter tournaments with a given score sequence. Coxeter tournaments serve as generalizations of standard tournaments, where in addition to competitive games, pairs of players can play collaborative games, and individual players can play solitary games. We obtain our main result by an intricate application of Bubley and Dyer's method of path coupling, using a re-weighting of the graph metric.

Contents

Chapter 1

Introduction

Given a random process on a graph, a natural line of inquiry is to investigate the asymptotics of certain properties of the process as the number of vertices of the underlying graph increases. This thesis focuses on three examples of random processes on graphs — that can be seen as graphs with additional temporal structure — namely discrete snakes, temporal graphs, and Markov chain Monte Carlo sampling. These examples are the respective subjects of Chapters 2, 3, and 4.

Sections 1.1, 1.2, and 1.3 discuss the relevant background, main results, and directions for further study of Chapters 2, 3, and 4, respectively.

1.1 Discrete snakes with globally centered displacements

If we simultaneously add vertices and shrink the lengths of the edges in a random graph appropriately, we might hope to achieve some notion of convergence, or in other words, prove a scaling limit. In recent years much research has been conducted on the topic of scaling limits of random graphs. Central to this study is the Brownian continuum random tree (CRT), introduced by Aldous in the early 1990s, [7–9]. Briefly, the CRT is a continuous tree-like structure which appears as the scaling limit of certain size-conditioned Bienaymé trees. Since the work of Aldous, the theory has been extended to encompass the scaling limits of various discrete tree-like structures, such as size-conditioned Bienaymé trees with heavy-tailed offspring distributions [39, 70], trees with prescribed degree sequences [23, 78], and random graphs [3, 35, 37]. There is increasing interest in models of random graphs with additional structure on the vertices, and their scaling limits. The project in Chapter 2 concerns a particular example of such models, called branching random walks. Informally, consider a population in which individuals are given a spatial position. Independently of one another, these individuals have random numbers of offspring, and each of their children are, in turn, endowed with a spatial location that is a random displacement away from their parent's location. By construction, the positions along each individual's (or rather vertex's) lineage form a random walk. In the scaling limit, the genealogy of the population is described by the scaling limit of the Bienaymé tree model, on top of which there is a Gaussian process capturing the continuum limit of the spatial motion. To prove such a scaling limit, it is convenient to encode the genealogy and spatial locations using a process called a *discrete snake*. To be brief, the discrete snake is a path-valued process which encodes the spatial trajectories of the individuals in the population in the order they are visited in a depth-first exploration of the underlying genealogical tree. In Chapter 2 we prove that under a global finite variance and a tail behaviour assumption on the displacements, any globally centered discrete snake on a Bienaymé tree whose offspring distribution is critical and admits a finite third moment has the Brownian snake driven by a normalised Brownian excursion as its scaling limit.

1.1.1 Background

In the course of our arguments in Chapter 2 we require a number of different random tree models. We will begin with a discussion of trees and their encodings, followed by a discussion of real trees, and the scaling limits of certain models of random trees. We will then introduce branching random walks and their encodings via discrete snakes, and the Brownian snake driven by a normalized Brownian excursion.

1.1.1.1 Trees

A *tree* is a simply connected acyclic graph $T = (v(T), e(T))$, where $v(T)$ is the *vertex set* of T, and $e(T)$ is the *edge set* of T. A *rooted tree* is a pair (T, ρ) where T is a tree and $\rho = \rho(T)$ is a distinguished vertex of T called the *root*. Given a rooted tree, for each $v \in v(T) \setminus \rho$, we write $p(v)$ for the neighbour of v on the unique path from v to ρ in T, and say that $p(v)$ is the *parent* of v. If $p(u) = v$, we say that u is a *child* of v. Furthermore, for $v \in v(T)$, we let $c(v, T) = |\{u \in v(T) \ : \ p(u) = v\}|$ be the number of children of v in T. If $c(v, T) = 0$ then we say that v is a *leaf* of T, and we denote the set of leaves of T by ∂T.

Letting $\mathbb{N}^0 = \{\emptyset\}$, the *Ulam–Harris tree* is the rooted tree with root $\emptyset$ and vertex set

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

in which, for each $v \in \mathcal{U}$, the set of children of v is $\{vi, i \in \mathbb{N}\}$. In this way, for $w = vi \in \mathcal{U}$, $p(w) = v$, and if $w' = vj$ with $j > i$ we say that w' is a *younger sibling* of w. For $w \in \mathcal{U} \cap \mathbb{N}^n$, we write $|w| = n$ for the *depth* of w. We will make use of the lexicographic order on $\mathcal{U}$, which is the total order in which each vertex precedes both its children and younger siblings.

An *ordered rooted tree* is a rooted tree T with $v(T) \subset \mathcal{U}$ such that $\emptyset \in v(T)$, and if $w = vj \in v(T)$ for some $j \in \mathbb{N}$ then $v \in v(T)$ and $vi \in v(T)$ for all $1 \le i < j$. The lexicographic order on $v(T)$ is the restriction of the lexicographic order on $\mathcal{U}$ to $v(T)$.

1.1.1.2 Depth-first encodings of trees

The definition of an ordered rooted tree does not lend itself easily to analysis in the context of proving scaling limits. As such, it is advantageous to encode such trees using continuous functions. We will define three such continuous functions, which in the setting of random trees will display their own advantages and disadvantages.

Let (T, ρ) be an ordered rooted tree and write $n = |v(T)|$. The *Łukasiewicz path* of T is the function $W_T : [0, n] \to \mathbb{R}$ defined as follows: Let $v_1, \ldots, v_n$ be the elements of $v(T)$ listed in lexicographical order; set $W_T(0) = 0$, and for $1 \le i \le n$ set $W_T(i) = \sum_{j=1}^{i} (c(v_i, T) - 1)$, and then extend the domain of W_T to $[0, n]$ by linear interpolation. The *height function* of T is the function $H_T : [0, n] \to \mathbb{R}_{\ge 0}$ defined as follows: For $0 \le i < n$ set $H_T(i) = |v_{i+1}|$ and set $H_T(n) = 0$; then extend the domain of H_T to $[0, n]$ by linear interpolation.

To define a third such encoding function, we first define the *contour order* of $v(T)$. This is the sequence $w_0, \ldots, w_{2(n-1)}$ of elements of $v(T)$ defined as follows. First, set $w_0 = \emptyset$. Then inductively for each $0 \le i < 2(n-1)$, if w_i has at least one child in T which does not appear in the sequence $w_0, \ldots, w_{i-1}$ then let w_{i+1} be the lexicographically least such child; otherwise let w_{i+1} be the parent of w_i in T. It is straightforward to verify that each vertex $v \in v(T)$ appears in the contour order of $v(T)$ exactly $1 + c(v, T)$ times. The *contour function* of T is the function $\widetilde{H}_T : [0, 2n] \to \mathbb{R}_{\ge 0}$ defined by setting $\widetilde{H}_T(i) = |w_i|$ for all $0 \le i \le 2(n-1)$, $\widetilde{H}_T(2n) = 0$, and extending the domain to $[0, 2n]$ by linear interpolation.

It is readily seen that a tree T can be recovered from the height and contour processes, and additionally since these encodings provide direct information about graph distances their utility is intuitively clear. The recovery of T from the Łukasiewicz path is less obvious. However, by Le Gall [42, Proposition 1.2] the height function H_T of T can be recovered from the Łukasiewicz path W_T using the formula

$$H_T(i) = \left| \left\{ j \in \{1, \ldots, i-1\} \ : \ W_T(j) = \inf_{j \le k \le i} W_T(k) \right\} \right|, \tag{1.1.1}$$

for all $i \in \{0, 1, \ldots, n-1\}$. The advantage of the Łukasiewicz path is clearer in the setting of random trees, and is discussed in the sequel.

1.1.1.3 Real trees, their encodings, and the Gromov-Hausdorff topology

We now describe the framework that we will use for the scaling limits of random trees.

A compact metric space $(\mathcal{T}, d)$ is a *real tree* if (i) for every $a, b \in \mathcal{T}$ there is a unique isometric map $f_{a,b} : [0, d(a, b)] \to \mathcal{T}$ such that $f_{a,b}(0) = a$ and $f_{a,b}(d(a, b)) = b$, and (ii) for any continuous injective map $g : [0, 1] \to \mathcal{T}$ such that $g(0) = a$ and $g(1) = b$ it holds that $g([0, 1]) = f_{a,b}([0, d(a, b)])$. A *rooted real tree* is a real tree $(\mathcal{T}, d)$ with a distinguished root vertex $\rho = \rho(\mathcal{T})$.

As for discrete trees, continuous functions may encode real trees. To this end, let $h : [0, \infty) \to [0, \infty)$ be a continuous function with compact support such that $h(0) = 0$. For $s, t \geq 0$, let

$$d_h(s, t) := h(s) + h(t) - 2 \inf_{u \in [s \wedge t, s \vee t]} h(u).$$

Furthermore, introduce the equivalence relation $s \sim t$ if and only if $d_h(s, t) = 0$. Then if $\mathcal{T}_h = [0, \infty)/\sim$, the metric space $(\mathcal{T}_h, d_h)$ is a real tree. Lastly, if $p_h : [0, \infty) \to \mathcal{T}_h$ is the canonical projection, then we can view $(\mathcal{T}_h, d_h)$ as a rooted real tree with root $\rho = p_h(0)$. We say that $\mathcal{T}_h$ is the real tree encoded by h.

Remark 1.1.1. *We note that we may think of trees $T \subseteq \mathcal{U}$ as real trees by simply viewing T as a union of line segments of unit length in the plane. For a tree $T \subseteq \mathcal{U}$, and a constant $c > 0$, we write cT for the tree obtained by rescaling the line segments of T by a factor of c.*

In the sequel, we will discuss scaling limits of certain size-conditioned random trees. As such, we require a notion of distance between two rooted real trees.

Definition 1.1.2 (Pointed Gromov–Hausdorff distance)**.** *Let (E, δ) be a metric space, and $\varepsilon > 0$. For $K \subseteq E$, denote by $B_\varepsilon(K; E)$ the ε-neighbourhood of K in E. Moreover, denote by $\delta_{\mathrm{Haus},E}(\cdot, \cdot)$ the Hausdorff distance: for $K, K' \subseteq E$*

$$d_{\mathrm{Haus},E}(K, K') := \inf\{\varepsilon > 0 \ : \ K \subseteq B_\varepsilon(K'; E), K' \subseteq B_\varepsilon(K; E)\}.$$

Let $((\mathcal{T}, d), \rho)$, and $((\mathcal{T}', d'), \rho')$ be rooted real trees. The pointed Gromov–Hausdorff distance between $((\mathcal{T}, d), \rho)$, and $((\mathcal{T}', d'), \rho')$, denoted by $d_{\mathrm{GH}}(\mathcal{T}, \mathcal{T}')$, is defined as

$$d_{\mathrm{GH}}(\mathcal{T}, \mathcal{T}') = \inf\{\delta_{\mathrm{Haus},E}(\phi(\mathcal{T}), \phi'(\mathcal{T}')) \vee \delta(\phi(\rho), \phi'(\rho'))\},$$

where the infimum is taken over all choices of metric space (E, δ) and all isometric embeddings $\phi : \mathcal{T} \to E$ and $\phi' : \mathcal{T}' \to E$.

The pointed Gromov–Hausdorff distance between two real rooted trees $((\mathcal{T}, d), \rho)$ and $((\mathcal{T}', d'), \rho')$ is then small if one can find isometric embeddings ϕ and ϕ' into a common metric space such that $\phi(\mathcal{T})$ and $\phi'(\mathcal{T}')$ have small Hausdorff distance, and $\phi(\rho)$ and $\phi'(\rho')$ are close in the common metric space. The following lemma states that for two continuous functions with compact support $h, h' : [0, \infty) \to [0, \infty)$ such that $h(0) = h'(0) = 0$, the pointed Gromov–Hausdorff distance between the real rooted trees $((\mathcal{T}_h, d_h), \rho)$ and $((\mathcal{T}_{h'}, d_{h'}), \rho')$ is small if $\|h - h'\|_\infty = \sup_{s \geq 0} |h(s) - h'(s)|$ is small.

Lemma 1.1.3 (Lemma 2.4 of Le Gall [42])**.** *For two continuous functions with compact support $h, h' : [0, \infty) \to [0, \infty)$ such that $h(0) = h'(0) = 0$,*

$$d_{\mathrm{GH}}(\mathcal{T}_h, \mathcal{T}_{h'}) \leq 2\|h - h'\|_\infty.$$

1.1.1.4 Bienaymé trees and their scaling limits

We now describe the model of random tree that will form the underlying genealogy of the random processes considered in Chapter 2.

Let $\mu = (\mu_k)_{k\geq 0}$ be an *offspring distribution*; that is a probability distribution on $\mathbb{Z}_+$. Recall the Ulam–Harris tree, with vertex set $\mathcal{U}$, and let $(\xi_v, v \in \mathcal{U})$ be IID random variables with distribution μ. A *Bienaymé tree* with offspring distribution μ, denoted by $\mathrm{T} = \mathrm{T}^{(\mu)}$, is a random subtree T with $v(T) \subseteq \mathcal{U}$, such that $\emptyset$ has $\xi_\emptyset$ children $\{1, \ldots, \xi_\emptyset\}$ and, iteratively, if $v \in v(T)$, then $c(v, T) := \xi_v$ so that v has children $\{v1, \ldots, v\xi_v\}$[1]. By Harris [53, Theorem 6.1], which was originally proved by Steffensen [103], if $\sum_{k\geq 1} k\mu_k \leq 1$, then $\mathbf{P}\left\{|\mathrm{T}^{(\mu)}| < \infty\right\} = 1$, and otherwise $\mathbf{P}\left\{|\mathrm{T}^{(\mu)}| < \infty\right\} = \gamma \in [0, 1)$, where γ is the smallest non-negative fixed point of the moment generating function of μ. We say that μ is *critical* if $\sum_{k\geq 1} k\mu_k = 1$.

In the sequel we will always denote by $\mathrm{T}_n^{(\mu)}$ a Bienaymé tree with offspring distribution μ conditioned to have n vertices. When the context is clear, we will omit the superscript and simply write T_n. Furthermore, we denote by H_n and $\widetilde{H}_n$ the height and contour functions of T_n respectively, and by W_n the Łukasiewicz path of T_n. In this setting, W_n has the same law as the first n steps of a random walk on $\mathbb{Z}$ with initial value 0, and jump distribution $\tilde{\mu}_k = \mu_{k+1}$ for all $k \geq -1$, that is conditioned to hit -1 for the first time at time n. By Kaigh [62], if μ is critical and has finite variance $\sigma^2 \in (0, \infty)$, then as $n \to \infty$,

$$\left(\frac{W_n(nt)}{\sqrt{n}}\right)_{0\leq t\leq 1} \xrightarrow{\mathrm{d}} \sigma(\mathbf{e}_t)_{0\leq t\leq 1} \tag{1.1.2}$$

[1]To avoid degenerate cases and technicalities, we shall assume that $\mu_0 + \mu_1 < 1$, and that the support of μ has greatest common divisor 1, so that the event that T has size $n \geq 0$ has strictly positive probability for all n large enough.

in $\mathbf{C}([0,1],\mathbb{R})$ where $(\mathbf{e}_t)_{0\le t\le 1}$ is a normalized Brownian excursion. Remarkably, Aldous [9] proved that the same limit holds for the contour function up to a constant scaling factor.

Theorem 1.1.4 (Aldous [9]). *Let $\mu = (\mu_k)_{k\ge 0}$ be a critical offspring distribution with finite variance $\sigma^2 \in (0,\infty)$ and a finite exponential moment. As $n \to \infty$,*

$$\left(\frac{\widetilde{H}_n(2nt)}{\sqrt{n}}\right)_{0\le t\le 1} \xrightarrow{\text{d}} \frac{2}{\sigma}(\mathbf{e}_t)_{0\le t\le 1}$$

in $\mathbf{C}([0,1],\mathbb{R})$ where $(\mathbf{e}_t)_{0\le t\le 1}$ is a normalized Brownian excursion.

Marckert and Mokkadem [82] later proved that if μ is critical with variance $\sigma^2 \in (0,\infty)$ and a finite exponential moment, then as $n \to \infty$,

$$\left(\frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{H_n(nt)}{\sqrt{n}}, \frac{W_n(nt)}{\sqrt{n}}\right)_{0\le t\le 1} \xrightarrow{\text{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \frac{2}{\sigma}\mathbf{e}_t, \sigma\mathbf{e}_t\right)_{0\le t\le 1} \tag{1.1.3}$$

in $\mathbf{C}([0,1],\mathbb{R}^3)$, where $(\mathbf{e}_t)_{0\le t\le 1}$ is a normalized Brownian excursion. (In fact, the finite exponential moment condition is unnecessary and may be removed; see Duquesne [39].)

By Lemma 1.1.3, it is natural to expect that if μ is critical with finite variance $\sigma^2 \in (0,\infty)$, then $\sigma n^{-1/2}\mathrm{T}_n$ converges in distribution in the Gromov–Hausdorff topology to a random real tree which is encoded by a normalized Brownian excursion, rescaled by a factor of 2. The real tree which is encoded by such a rescaled excursion, denoted by $\mathcal{T}_{2\mathbf{e}}$, is called the *Brownian Continuum Random Tree (CRT)*. This tree was first introduced and studied by Aldous in the series of papers [7–9], see Figure 1.1 for an illustration.

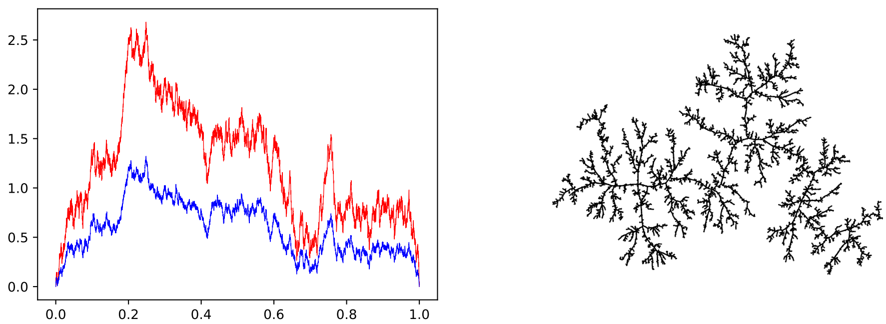

Figure 1.1: In red and blue (resp.), the height function and Łukasiewicz path (resp.) associated with the simulation of the CRT on the right.

Theorem 1.1.5 (Aldous [9], Le Gall [77]). *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$. Then as $n \to \infty$,*

$$\frac{\sigma}{\sqrt{n}} \mathrm{T}_n \xrightarrow{\mathrm{d}} \mathcal{T}_{2\mathbf{e}}$$

in the Gromov-Hausdorff topology, where $\mathcal{T}_{2\mathbf{e}}$ is the CRT.

The CRT can be constructed using a "line-breaking" construction. This construction will be particularly useful in the proofs in Chapter 2.

Line-breaking construction of the CRT

- Let $J_1, J_2, \ldots$ be the ordered jump times of a Poisson point process on $[0,\infty)$ with intensity tdt.
- For $i \geq 1$ sample attachment points $A_i \sim \mathrm{Uniform}([0, J_i])$, independently of $(A_j)_{j\neq i}$.
- Take the completion of each of the line segments $[0, J_1], (J_1, J_2], \ldots$ and for each $i \geq 1$ let J_i^* denote the limit point as $x \downarrow J_i$.
- Identify the points J_i^* and A_i, and denote the resulting random real tree by $\mathcal{T}$.

Heuristically, the best way to see that this construction leads to a tree $\mathcal{T} \stackrel{\mathrm{d}}{=} \mathcal{T}_{2\mathbf{e}}$, is through the Aldous–Broder algorithm [6, 21] which generates a uniform Cayley tree (that is, a uniform labelled rooted tree with vertex set $\{1, \ldots, n\}$).

Aldous–Broder algorithm

- Fix $n \geq 2$ and let $U_1, \ldots, U_n$ be IID Uniform($\{0, \ldots, n\}$) random variables.
- Construct a labelled rooted tree T as follows: let 1 be the root and for $2 \leq i \leq n$ let vertex i be the child of $V_i := \min(U_i, i-1)$.

By removing the labels and assigning a uniformly random planar order, the tree output by the Aldous–Broder algorithm has the same distribution as a Bienaymé tree $\mathrm{T}_n^{(\mu)}$ with μ being the Poisson(1) distribution. By Theorem 1.1.5, $n^{-1/2}\mathrm{T}_n^{(\mu)} \xrightarrow{\mathrm{d}} \mathcal{T}_{2\mathbf{e}}$ as $n \to \infty$ in the Gromov–Hausdorff topology. The Aldous–Broder algorithm makes the asymptotic behaviour of these random trees intuitively clear, and identifies the line-breaking construction as a method to construction the CRT. We provide a brief discussion of the proof following the presentation of Goldschmidt [47] as follows. Consider the length of the first branch constructed in the Aldous–Broder algorithm,

$J_1^n := \inf\{i \geq 2 \ : \ V_i \neq i-1\}$. Then for all $x \geq 0$,

$$\mathbf{P}\left\{J_1^n > x\sqrt{n}\right\} = \mathbf{P}\left\{J_1^n \geq \lfloor x\sqrt{n}\rfloor + 1\right\} = \prod_{i=1}^{\lfloor x\sqrt{n}\rfloor - 2}\left(1 - \frac{i}{n-1}\right),$$

and so by taking logarithms and Taylor expanding, we obtain that

$$\begin{aligned} -\log\left(\mathbf{P}\left\{J_1^n > x\sqrt{n}\right\}\right) &= -\sum_{i=1}^{\lfloor x\sqrt{n}\rfloor - 2}\log\left(1 - \frac{i}{n-1}\right) \\ &= \sum_{i=1}^{\lfloor x\sqrt{n}\rfloor - 2}\frac{i}{n-1} + o(1) \\ &= \frac{(\lfloor x\sqrt{n}\rfloor - 2)(\lfloor x\sqrt{n}\rfloor - 1)}{2(n-1)} + o(1) \\ &\to \frac{x^2}{2} \end{aligned}$$

as $n \to \infty$. Since $\mathbf{P}\{J_1 > x\} = \exp(-x^2/2)$ for all $x \geq 0$, it follows that $n^{-1/2}J_1^n \xrightarrow{\mathrm{d}} J_1$ as $n \to \infty$. Having built the first branch, subsequent branches are then constructed by selecting a uniform point along the one of the previous branches, and growing a second path starting from this uniform point. Indeed, if m_k^n is the k-th element of the set $\{i \geq 2 \ : \ V_i \neq i-1\}$, and $A_k^n = J_{m_k^n}$, then by [6, Theorem 8], for all $k \geq 1$, as $n \to \infty$

$$\left(\frac{1}{\sqrt{n}}(J_1^n, A_1^n), \frac{1}{\sqrt{n}}(J_2^n, A_2^n), \ldots, \frac{1}{\sqrt{n}}(J_k^n, A_k^n)\right) \xrightarrow{\mathrm{d}} ((J_1, A_1), (J_2, A_2), \ldots, (J_k, A_k)).$$

Convergence results have also been obtained for Bienaymé trees in settings where the offspring distribution no longer has finite variance. We will be brief in our description and refer the reader to Duquesne [39] for precise details. We say that a critical offspring distribution μ *belongs to the domain of attraction of a stable law with index* $\alpha \in (1, 2]$ if there exists an increasing sequence $(b_n)_{n\geq 1}$ such that if $(\xi_n)_{n\geq 1}$ is a sequence of IID random variables with common distribution μ, then

$$\frac{1}{b_n}\sum_{i=1}^{n}(\xi_i - 1) \xrightarrow{\mathrm{d}} X^{(\alpha)},$$

where $X^{(\alpha)}$ is an α-stable random variable whose law is given by the Laplace exponent:

$$\mathbf{E}\left[\exp\left(-\lambda X^{(\alpha)}\right)\right] = \exp(\lambda^\alpha) \quad \text{for every } \lambda \geq 0. \tag{1.1.4}$$

Let $(X_t^{(\alpha)})_{t\geq 0}$ be an α-stable Lévy process with no negative jumps, such that $X_1^{(\alpha)}$ has the same law as $X^{(\alpha)}$. Let $(\mathbf{x}_t)_{0\leq t\leq 1}$ be the normalized excursion of this process, defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Then the *associated height function*, $(\mathbf{h}_t)_{0\le t\le 1}$, is such that for $t \ge 0$

$$\mathbf{h}_t := \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\int_0^t \mathbf{1}_{[\mathbf{x}_s < \inf_{u\in[s,t]} \mathbf{x}_u + \varepsilon]} ds \tag{1.1.5}$$

where this limit exists $\mathbf{P}$-almost surely for a set of values of t of full Lebesgue measure on $[0,1]$. The existence of this limit is non-trivial, see Duquesne and Le Gall [40]. Note that $\mathbf{h}_t = 2^{-1/2}\mathbf{e}_t$ if $\alpha = 2$, and further that (1.1.5) can be thought of as the continuous analogue of (1.1.1).

Theorem 1.1.6 (Duquesne [39])**.** *Suppose that $\mu = (\mu_k)_{k\ge 0}$ is a critical offspring distribution that belongs to the domain of attraction of a stable law with index $\alpha \in (1,2]$. Then as $n \to \infty$,*

$$\left(\frac{b_n}{n}\widetilde{H}_n(2nt), \frac{b_n}{n}H_n(nt), \frac{1}{b_n}W_n(nt)\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} (\mathbf{h}_t, \mathbf{h}_t, \mathbf{x}_t)_{0\le t\le 1}$$

in $\mathbf{C}([0,1],\mathbb{R}^3)$ endowed with the uniform topology, where $(\mathbf{x}_t)_{0\le t\le 1}$ is the normalized excursion of the α-stable Lévy process with no negative jumps, and $(\mathbf{h}_t)_{0\le t\le 1}$ is the associated height function.

As a consequence of Lemma 1.1.3, when μ belongs to the domain of attraction of a stable law with index $\alpha \in (1,2]$, the tree $\mathrm{T}_n^{(\mu)}$ viewed as a metric space converges in distribution upon rescaling by (b_n/n) to the random real tree encoded by $(\mathbf{h}_t)_{0\le t\le 1}$ in the Gromov-Hausdorff topology. Such trees, called α-stable Lévy trees, serve as generalizations of the CRT (indeed when $\alpha = 2$ the tree is the CRT rescaled by $(2\sqrt{2})^{-1}$).

1.1.1.5 Discrete snakes

The fundamental object of study in Chapter 2 is the *branching random walk*, that is, a pair (T, Y), where T is an ordered rooted tree (possibly labelled) and $Y = (Y^{(v)}, v \in v(T)\setminus\partial T)$, where $Y^{(v)} = (Y_j^{(v)}, j \in [c(v,T)]) \in \mathbb{R}^{c(v,T)}$. We think of $Y^{(v)}$ as a vector of spatial displacements from vertex v to its children, so that $Y_j^{(v)}$ is the difference between the spatial locations of vertices v and vj. We assume that the displacements from individual vertices are independent, but allow for displacements of siblings to be dependent. In the sequel, for a vertex v with k children, the distribution of the vector of displacements from v to its children is denoted by ν_k for $k \ge 1$.

Similarly to trees, in order to prove a scaling limit for branching random walks, it is convenient to encode the branching random walk by an object that is more tractable. For this reason, we use *discrete snakes*, path-valued processes which encode branching random walks via depth-first traversals. Let (T, Y) be a branching random walk with $|T| = n$. For a vertex $v \in v(T)$, let $P_v = v_0, v_1, \ldots, v_k$ be the unique shortest path from $\emptyset$ to v. We associate a random walk trajectory with the path P_v, $\Phi_v = (\Phi_v(j))_{0 \le j \le |v|}$, where for each $j \in \{0, \ldots, |v|\}$,

$$\Phi_v(j) = \sum_{(u,ui) \in e(T): ui \preceq v_j} Y_i^{(u)}.$$

Let $w_0, \ldots, w_{2(n-1)}$ be the vertices of T in contour order. The *discrete snake* associated with T is the process $(\widetilde{H}_T(t), S_{\mathrm{T}}(t,s))_{0 \le t \le 2n, s \ge 0}$ where for each $t \in [0, 2n]$, $S_{\mathrm{T}}(t, \cdot)$ is a stopped continuous process defined as follows:

- For $k \in [[0, 2n-1]]$ and $s \in [0, \widetilde{H}_T(k)]$, let $S_{\mathrm{T}}(k, \cdot)$ be the process that linearly interpolates the random walk $\Phi_{w_k} = (\Phi_{w_k}(j))_{0 \le j \le |w_k|}$.
- For $t \in [0, 2n-1] \setminus \mathbb{Z}$ and $s \in [0, \widetilde{H}_T(t)]$,

$$S_{\mathrm{T}} = \begin{cases} S_{\mathrm{T}}(\lfloor t \rfloor, s) & \text{if } \widetilde{H}_T(\lceil t \rceil) < \widetilde{H}_T(\lfloor t \rfloor) \\ S_{\mathrm{T}}(\lceil t \rceil, s) & \text{if } \widetilde{H}_T(\lceil t \rceil) > \widetilde{H}_T(\lfloor t \rfloor) \end{cases}$$

- For $t \in [0, 2n]$ and $s \in (\widetilde{H}_T(t), \infty)$, $S_{\mathrm{T}}(t, s) = S_{\mathrm{T}}(t, \widetilde{H}_T(t))$.

In the sequel, we will see that in fact, a continuous function called the *head of the discrete snake* will prove sufficient to encode branching random walks for the purposes of proving scaling limits.

1.1.1.6 The Brownian snake driven by a normalized Brownian excursion

We now describe the framework that we will use for our limiting snakes.

Fix $d \in \mathbb{N}$ and $x \in \mathbb{R}_+$. Let $\mathcal{S}_x$ be the set of all finite paths in $\mathbb{R}^d$ started from x. An element $(\gamma, g) \in \mathcal{S}_x$ is a pair $\gamma \in \mathbb{R}_+$, and g is a continuous mapping from $\mathbb{R}_+$ into $\mathbb{R}^d$ with $g(0) = x$ that is constant over $[\gamma, \infty)$. We call γ the lifetime of g. The *Brownian snake driven by a normalized Brownian excursion (BSBE)* is a continuous strong Markov process with values in $\mathcal{S}_x$, denoted by $(\gamma_t, S_t)_{0 \le t \le 1}$, whose distribution is characterized by the following two conditions:

- The process $(\gamma_t)_{0 \le t \le 1}$ is a normalized Brownian excursion in $\mathbb{R}_+$.
- Conditionally on $(\gamma_t)_{0 \le t \le 1}$ the process $(S_t)_{0 \le t \le 1}$ is a time inhomogeneous Markov process with transition probabilities characterized as follows. For $t < t'$,
  - $S_{t'}(s) = S_t(s)$ for all $s \in [0, \inf_{r \in [t,t']} \gamma_r]$.

– $(S_{t'}(\inf_{r\in[t,t']}\gamma_r+s)-S_{t'}(\inf_{r\in[t,t']}\gamma_r))_{s\geq 0}$ is a standard Brownian motion in $\mathbb{R}^d$ stopped at $\gamma_{t'}-\inf_{r\in[t,t']}\gamma_r$, independent of S_t.

One can think of S_t as a Brownian path in $\mathbb{R}^d$ started at x, with a random lifetime γ_t. The lifetime γ_t evolves according to the law of a normalized Brownian excursion. When γ_t decreases, part of the path of S_t is "erased", while if γ_t increases, the path S_t is "extended" independently of the past. This is known as the *snake property.*

If we replace the normalized Brownian excursion, $(\gamma_t)_{0\leq t\leq 1}$, with a reflected Brownian motion $(\gamma_t)_{t\geq 0}$, obtaining a process $(\gamma_t, S_t)_{t\geq 0}$, we recover the *Brownian snake*, first defined and studied by Le Gall [75, 76]. In the sequel we will always take $x=0$, and $d=1$. For further details on the Brownian snake, its encoding, construction, and applications we refer the reader to [42].

1.1.1.7 The head of the discrete snake and the homeomorphism theorem

In what follows we will introduce a continuous function called the *head of the discrete snake*, which encodes the endpoints of the trajectories associated with the branching random walk (T, Y) in depth-first order. By the homeomorphism theorem (see Theorem 1.1.7 below), scaling limits of discrete snakes follow from those of their associated heads.

Let $\mathrm{T}=(T,Y)$ be a branching random walk with $|T|=n$ and $w_0,\ldots,w_{2(n-1)}$ be the contour order of T. Let $\widetilde{R}_{\mathrm{T}}:[0,2n]\to\mathbb{R}$ be the continuous function encoding the spatial positions of the vertices in T, defined as follows. Let $\widetilde{R}_{\mathrm{T}}(0)=\widetilde{R}_{\mathrm{T}}(2n)=0$, for $i\in[2(n-1)]$ let

$$\widetilde{R}_{\mathrm{T}}(i)=\sum_{(u,uj)\in e(T):uj\preceq w_i} Y_j^{(u)},$$

and extend the domain to $[0,2n]$ by linear interpolation. We also define the function $R_{\mathrm{T}}:[0,n]\to\mathbb{R}$ given by setting $R_{\mathrm{T}}(i)=\ell(v_{i+1},\mathrm{T})$, for $i\in\{0,\ldots,n-1\}$, $R_{\mathrm{T}}(n)=0$, and extending to $[0,n]$ by linear interpolation, where $v_1,\ldots,v_n$ are the elements of $v(T)$ listed in lexicographical order. When T is a size-conditioned Bienaymé tree with n vertices, we write $\widetilde{R}_n$ and R_n in place of $\widetilde{R}_{\mathrm{T}}$ and R_{T}.

The process

$$(\widetilde{H}_T(t),\widetilde{R}_{\mathrm{T}}(t))_{0\leq t\leq 2n}$$

encodes the endpoints of the discrete snake $(\widetilde{H}_T(t), S_{\mathrm{T}}(t,s))_{0\leq t\leq 2n, s\geq 0}$, and as such is called the *head of the discrete snake* associated with T. Analogously, in the continuum, the *head of the BSBE* is the process $(\gamma_t, S_{\gamma_t})_{0\leq t\leq 1}$. In what follows, we will use $(\mathbf{e}_t,\mathbf{r}_t)_{0\leq t\leq 1}$ to the denote the head of the BSBE, where $(\mathbf{e}_t)_{0\leq t\leq 1}$ is a normalised Brownian excursion, and conditionally on $(\mathbf{e}_t)_{0\le t\le 1}$, $(\mathbf{r}_t)_{0\le t\le 1}$ is a centred Gaussian process with covariance

$$\operatorname{cov}(\mathbf{r}_s, \mathbf{r}_t) = \inf_{u\in[s\wedge t, s\vee t]} \mathbf{e}_u. \tag{1.1.6}$$

The covariance in (1.1.6) follows from the "snake property" and can be interpreted as follows: for two vertices encoded by s and t in the limiting tree (that is, the CRT) with heights $\mathbf{e}_s$ and $\mathbf{e}_t$, the spatial motion associated with the vertices along their paths to the root evolve as a common Brownian motion up until their last common ancestor (which is at distance $\inf_{u\in[s\wedge t,s\vee t]} \mathbf{e}_t$), and then evolve as independent Brownian motions thereafter.

To conclude this subsection, we will outline the homeomorphism theorem, proved by Marckert and Mokkadem [83], which establishes a homeomorphism between the state space of the BSBE and the state space of the head of the BSBE. This implies that to prove a scaling limit for a discrete snake, it is sufficient to prove one for the associated head. To this end, let $\mathcal{S}$ be the subset of $\mathbf{C}([0,1], \mathbb{R}_{\ge 0}) \times \mathbf{C}([0,1], \mathbf{C}(\mathbb{R}_{\ge 0}, \mathbb{R}^d))$ of functions (γ, g) such that the following three conditions hold:

- $(\gamma_t, g_t(\cdot))$ is a stopped path in $\mathbb{R}^d$ for all $t \in [0,1]$;
- $\gamma_0 = \gamma_1 = 0$;
- for any $0 \le t \le t' \le 1$, $g_t(s) = g_{t'}(s)$ for all $s \le \min_{u\in[t,t']} \gamma_u$.

We endow the space $\mathcal{S}$ with the distance

$$d_{\mathcal{S}}((\gamma, g), (\gamma', g')) = \max\{\|g - g'\|_\infty, \|\gamma - \gamma'\|_\infty\},$$

where

$$\|g - g'\|_\infty = \sup_{(t,s)\in[0,1]\times[0,\infty)} |g_t(s) - g'_t(s)|.$$

Furthermore, let $\mathcal{H}$ be the subspace of $\mathbf{C}([0,1], \mathbb{R}_{\ge 0} \times \mathbb{R}^d)$ of functions (γ, f) that satisfy the following two conditions:

- $\gamma_0 = \gamma_1 = 0$;
- for any $0 \le t \le t' \le 1$, $f(t) = f(t')$ if $\gamma_t = \gamma_{t'}$ and $\min_{t\le u\le t'} \gamma_u = \gamma_t$.

We endow $\mathcal{H}$ with the distance

$$d_{\mathcal{H}}((\gamma, f), (\gamma', f')) = \max\{\|f - f'\|_\infty, \|\gamma - \gamma'\|_\infty\}.$$

where $\|f - f'\|_\infty = \sup_{t\in[0,1]} |f(t) - f'(t)|$.

Theorem 1.1.7 (The homeomorphism theorem, Marckert and Mokkadem [83])**.** *Let $\mathcal{F} : \mathcal{S} \to \mathcal{H}$ be the function such that for all $(\gamma, v) \in \mathcal{S}$, $\mathcal{F}(\gamma, g) = (\gamma, f)$, where*

$$f(t) = v_t(\gamma_t)) \quad \text{for all } t \in [0, 1].$$

Let $\bar{\mathcal{F}} : \mathcal{H} \to \mathcal{S}$ be the function such that for all $(\gamma, f) \in \mathcal{H}$, $\bar{\mathcal{F}}(\gamma, f) = (\gamma, g)$ where $g_t(s) = f(\rho(t, s, \gamma))$ for any $t \in [0, 1]$ and $s \geq 0$, with

$$\rho(t, s, \gamma) = \begin{cases} \sup\{\alpha \leq t : \gamma_\alpha = s\} & 0 \leq s \leq \gamma_t \\ t & s \geq \gamma_t. \end{cases}$$

Then $\mathcal{F}$ is a homeomorphism from $\mathcal{S}$ to $\mathcal{H}$ with inverse $\mathcal{F}^{-1} = \bar{\mathcal{F}}$.

1.1.2 Main results

In recent years there has been considerable research conducted on the convergence of the heads of discrete snakes. The key reason for this is that discrete snakes often play an important role in the study of planar maps, [1, 2, 32, 77, 81, 90].

For example, Addario-Berry and Albenque [1] prove a scaling limit for discrete snakes in order to prove a scaling limit of non-bipartite maps. In their proof, they study labelled multi-type trees with the property that vertices of the map are in correspondence with vertices of the tree, and the labels of the tree's vertices encode certain metric properties of the map. They subsequently show that the discrete snake associated with the labelled multi-type tree converges upon rescaling to the Brownian snake and build upon this result to prove a scaling limit for the non-bipartite maps. Their result concerning the convergence of discrete snakes is interesting in its own right, and we briefly describe it as follows: A multi-type Bienaymé tree with countable type space S is defined by a collection of probability distributions $(\mu^x, x \in S)$ on the set of all vectors of finite length with entries in S, denoted by S^{fin}. A multi-type branching random walk (T, Y) is a branching random walk where T is a multi-type tree with type space S, and the law of each displacement vector $Y^{(v)}$, is determined by the type of v together with the vector of types of its children. For $\mathrm{s} \in S^{\text{fin}}$ and $r \in S$, we write $Y^r_{\mathrm{s},i}$ for the displacement of the i-th child of a vertex of type r whose children have types given by s. Further, for $x \in S$, $\mathrm{s} \in S^{\text{fin}}$, we write $n_x(\mathrm{s}) = |\{1 \leq i \leq |\mathrm{s}| : s_i = x\}|$. The multi-type branching random walk is called *centered* if for any $k \in \mathbb{Z}_+$, for all $\mathrm{z} = (z_x, x \in S) \in \mathbb{Z}_+^S$ with $\sum_{x \in S} z_x = k$ and for all $r \in S$,

$$\sum_{\{\mathrm{s} \in S^k : (n_x(\mathrm{s}), x \in S) = \mathrm{z}\}} \sum_{i=1}^{k} \mathbf{E}\left[Y^r_{\mathrm{s},i}\right] = 0.$$

Addario-Berry and Albenque [1] show that the heads of discrete snakes for centered multi-type branching random walks converge upon rescaling to the head of the Brownian snake. To prove convergence they show that symmetrization (randomly permuting the order of each of the children of a vertex) turns such centered displacements into *locally centered* ones – displacements such that $\mathbf{E}\left[Y^r_{\mathrm{s},i}\right] = 0$ for all $r \in S$ and $\mathrm{s} \in S^{\mathrm{fin}}$ – and further that the convergence of such discrete snakes can be obtained via the convergence of their symmetrizations.

Returning now to branching random walks of a single type, we note that work where the offspring distribution is critical with finite exponential moment and the displacements are assumed to be IID, culminated in a paper by Janson and Marckert [60] in which the following result was proved.

Theorem 1.1.8 (Theorems 1 and 2 of Janson and Marckert [60])**.** *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$ such that μ has finite exponential moment. For each $k \geq 1$, let ν_k be the law of a vector of k IID copies of a random variable Y with $\mathbf{E}[Y] = 0$ and $\mathrm{var}(Y) = \beta^2$. Then,*

$$\left(\frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{\widetilde{R}_n(2nt)}{n^{1/4}}\right)_{0\leq t\leq 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\leq t\leq 1}$$

as $n \to \infty$, in the sense of finite-dimensional distributions. This convergence also holds in distribution in $\mathbf{C}([0,1],\mathbb{R}^2)$ endowed with the topology of uniform convergence if and only if

$$\mathbf{P}\{|Y| > y\} = o(y^{-4}) \text{ as } y \to \infty. \tag{1.1.7}$$

The finite exponential moment condition on the offspring distribution in Theorem 1.1.8 has subsequently been shown to be unnecessary, and may be weakened to a finite second moment assumption without changing the truth of the statement; see, for example, Marzouk [86].

The tail condition (1.1.7) is the most telling of the asymptotics. If this condition does not hold, it is possible that some of the positions of the branching random walk become very large in value, to the extent that they result in large spikes in $\widetilde{R}_n$. As such, in this setting, Janson and Marckert prove convergence for the head of the discrete snake in the space of non-empty compact subsets of $[0,1]\times\mathbb{R}$ equipped with the Hausdorff topology to an object which they call the *hairy tour*.

For a continuous function $f : [0,1] \to \mathbb{R}$ and a set $S \subset [0,1]\times\mathbb{R}^2_+ \setminus \{(0,0)\}$, write $U(f,S)$ for the union of the graph of f and the vertical line segments $[(t, f(t) - y), (t, f(t)+x)]$ for each $(t,x,y) \in S$.

Theorem 1.1.9 (Theorems 5 and 6 of Janson and Marckert [60]). *Let $\eta \in [0,4)$. Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$ such that μ has finite exponential moment. For each k, let ν_k be the law of a vector of k IID copies of a random variable Y with $\mathbf{E}[Y] = 0$, $\mathbf{E}[Y^2] = \beta < \infty$, and for $y > 0$*

$$\mathbf{P}\{Y \geq y\} = (a_+ + o(1))y^{-(4-\eta)} \text{ and } \mathbf{P}\{Y \leq -y\} = (a_- + o(1))y^{-(4-\eta)}$$

for some $a_+, a_- \geq 0$. Let Ξ be a Poisson process in $[0,1] \times (\mathbb{R} \setminus \{0\})$ with intensity $4a_+ y^{-(4-\eta)-1}dxdy$ for $y > 0$ and $4a_-|y|^{-(4-\eta)-1}dxdy$ for $y < 0$, which is independent of $(\mathbf{e}_t, \mathbf{r}_t)_{0\leq t\leq 1}$. Then if $\eta = 0$,

$$\left(\left(\frac{\widetilde{H}_n(2nt)}{\sqrt{n}}\right)_{0\leq t\leq 1}, U\left(\frac{\widetilde{R}_n(2n\cdot)}{n^{1/4}}, \emptyset\right)\right) \xrightarrow{\text{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\leq t\leq 1}, U\left(\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}, \Xi\right)\right),$$

and if $\eta \in (0,2)$,

$$\left(\left(\frac{\widetilde{H}_n(2nt)}{\sqrt{n}}\right)_{0\leq t\leq 1}, U\left(\frac{\widetilde{R}_n(2n\cdot)}{n^{1/(4-\eta)}}, \emptyset\right)\right) \xrightarrow{\text{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\leq t\leq 1}, U(0,\Xi)\right),$$

as $n \to \infty$, where the convergence in the first coordinate is in $\mathbf{C}([0,1],\mathbb{R})$ endowed with the topology of uniform convergence, and the convergence in the second is in the space of non-empty, compact subsets of $[0,1] \times \mathbb{R}$ endowed with the Hausdorff topology.

We note that by (1.1.3), Theorems 1.1.8 and 1.1.9 also hold if we replace the role of $\widetilde{H}_n$ by H_n, and $\widetilde{R}_n$ by R_n.

When the displacements are no longer assumed to be IID, of greatest relevance to our results in Chapter 2 is the work of Marckert [84]. Let ξ be a random variable with distribution μ and $\bar{\xi}$ be a size-biased version, that is, a random variable with distribution $\bar{\mu} = (\bar{\mu}_k)_{k\geq 1}$ where for all $k \geq 1$,

$$\bar{\mu}_k = \frac{k\mu_k}{\mathbf{E}[\xi]}.$$

Conditionally on $\bar{\xi}$, let $U_{\bar{\xi}}$ be a Uniform($\{1,\ldots,\bar{\xi}\}$) random variable. For each $k \geq 1$, let $Y_k := (Y_{k,1},\ldots,Y_{k,k})$ be a random vector with distribution ν_k. If

$$\mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\right] = 0,$$

we say that the discrete snake is *globally centered.* Furthermore, if

$$\beta^2 := \text{var}\left(Y_{\bar{\xi},U_{\bar{\xi}}}\right) < \infty,$$

we say that the discrete snake has *finite global variance.*

Theorem 1.1.10 (Theorem 1 of Marckert [84]). *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with bounded support (i.e., there exists $K > 0$ such that $\sum_{k=K+1}^{\infty} \mu_k = 0$). Suppose further that $(\nu_k)_{k\geq 1}$ is such that*

$$\mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\right] = 0 \quad \textit{and} \quad \beta^2 = \mathbf{E}\left[Y^2_{\bar{\xi},U_{\bar{\xi}}}\right] < \infty,$$

and there exists $p > 4$ such that

$$\sup_{1\leq j\leq k\leq K} \mathbf{E}\left[|Y_{k,j} - \mathbf{E}\left[Y_{k,j}\right]|^p\right] < \infty.$$

Then, as $n \to \infty$,

$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_n(nt)}{n^{1/4}}, \frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{\widetilde{R}_n(2nt)}{n^{1/4}}\right)_{0\leq t\leq 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t, \frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\leq t\leq 1},$$

in $\mathbf{C}([0,1],\mathbb{R}^4)$ endowed with the topology of uniform convergence.

This result allows one to obtain convergence results for a wide range of jump distributions, under the condition that the offspring distribution has bounded support. This is particularly powerful, as given the underlying tree, the jumps may be deterministic. To prove this result, Marckert tracks detailed information regarding the number of vertices of each possible different degree along a given branch of the tree, and shows that a record of this information converges upon appropriate rescaling to a Gaussian field. In so doing, Marckert yields a very fine understanding of the limiting object. In Chapter 2 we extend Theorem 1.1.10 to a class of offspring distributions which do not necessarily have bounded support. However, we do not achieve as fine information in the limit.

We use the following notion of convergence for a sequence of random elements $(f_n)_{n\geq 1}$ of $\mathbf{C}([0,1],\mathbb{R})$ such that $f_n(0) = f_n(1) = 0$ for all $n \geq 1$. Let $U_1, U_2, \ldots$ be IID Uniform$([0,1])$ random variables, independent of everything else. For $k \geq 1$, write $U^k_{(1)}, \ldots, U^k_{(k)}$ for the order statistics of $U_1, \ldots, U_k$. For another random element f of $\mathbf{C}([0,1],\mathbb{R})$ such that $f(0) = f(1) = 0$, se say that $f_n \xrightarrow{\mathrm{d}} f$ in the sense of *random finite-dimensional distributions* if, for every $k \geq 1$,

$$(f_n(U^k_{(1)}), \ldots f(U^k_{(k)})) \xrightarrow{\mathrm{d}} (f(U^k_{(1)}), \ldots, f(U^k_{(k)}))$$

as $n \to \infty$. In Chapter 2, we prove the following theorem.

Theorem 1.1.11. *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$. If $\nu = (\nu_k)_{k\geq 1}$ is such that*

[A1] $$\mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\right] = 0 \quad \textit{and} \quad \beta^2 = \mathbf{E}\left[Y^2_{\bar{\xi},U_{\bar{\xi}}}\right] < \infty,$$

then as $n \to \infty$ the following joint convergence holds in the sense of random finite-dimensional distributions:

$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_n(nt)}{n^{1/4}}, \frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{\widetilde{R}_n(2nt)}{n^{1/4}}\right)_{0\leq t\leq 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t, \frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\leq t\leq 1}. \tag{1.1.8}$$

The convergence (1.1.8) holds in distribution in $\mathbf{C}([0,1],\mathbb{R}^4)$ endowed with the topology of uniform convergence if, additionally,

[**A2**] $\qquad \mathbf{P}\left\{\max_{1\leq i\leq \xi} |Y_{\xi,i}| > y\right\} = o(y^{-4})$ *as $y \to \infty$ and* $\mathbf{E}\left[\xi^3\right] < \infty.$

The analogue of Theorem 1.1.11 also holds with $\mathbb{R}^d$-valued displacements for $d > 1$, and with essentially identical proofs to those in Chapter 2; the only change in the conclusion is that the limit $\mathbf{r}$ of the rescaled spatial displacements takes values in $\mathbb{R}^d$ rather than in $\mathbb{R}$, and that β^2 should be interpreted as the covariance matrix of $Y_{\bar{\xi},U_{\bar{\xi}}}$.

In Chapter 2 we also give two applications of Theorem 1.1.11. Firstly, we obtain a scaling limit for the difference between the height function and the Łukasiewicz path of a size-conditioned critical Bienaymé tree (see Corollary 2.1.8). Secondly, we obtain a scaling limit for the difference between the height function of a size-conditioned critical Bienaymé tree and the height function of its associated looptree (see Corollary 2.1.9). Similarly to Janson and Marckert [60], in Chapter 2 we also consider displacement distributions with heavier tails, for which we instead obtain convergence to a variant of the hairy snake. To state our convergence result in this case we introduce a further assumption, [**A3**] below. For $k \geq 1$ and $j \in [k]$ denote by

$$Y_{k,j}^+ := Y_{k,j} \vee 0 \text{ and } Y_{k,j}^- := (-Y_{k,j}) \vee 0,$$

the positive and negative displacements of the j-th child of a vertex with k children, respectively. Further let $Y_k^+ := (Y_{k,j}^+)_{j\in[k]}$ and $Y_k^- = (Y_{k,j}^-)_{j\in[k]}$.

[**A3**] Suppose that $\mathbf{E}[\xi^3] < \infty$. Furthermore, suppose that there exists a Borel measure π on $\mathbb{R}_+^2 \setminus \{(0,0)\}$ such that for any $\varepsilon > 0$, both $\pi(\mathbb{R}_+ \times (\varepsilon,\infty)) < \infty$ and $\pi((\varepsilon,\infty)\times\mathbb{R}_+) < \infty$, and there exists $\eta \in [0,2)$ such that for all Borel sets $A \subset \mathbb{R}_+^2 \setminus \{(0,0)\}$ for which $\pi(\partial A) = 0$,

$$r^{4-\eta}\mathbf{P}\left\{\frac{1}{r}\left(\max_{1\leq i\leq \xi} Y_{\xi,i}^+, \max_{1\leq i\leq \xi} Y_{\xi,i}^-\right) \in A\right\} \to \pi(A)$$

as $r \to \infty$.

Theorem 1.1.12. *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$, and let $\nu = (\nu_k)_{k\geq 1}$ be such that [**A1**] holds and [**A3**] holds for a given measure π with $\eta \in [0,2)$. Then, taking Ξ to be a Poisson process on $[0,1]\times\mathbb{R}_+^2 \setminus \{(0,0)\}$ with intensity $dt \otimes \pi(dx,dy)$, we have that if $\eta = 0$*

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}\right)_{0\leq t\leq 1}, \; U\left(\frac{R_n(n\cdot)}{n^{1/4}}, \emptyset\right)\right) \xrightarrow{\text{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\leq t\leq 1}, \; U\left(\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}, \Xi\right)\right), \quad (1.1.9)$$

and if $\eta \in (0,2)$,

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}\right)_{0\leq t\leq 1}, \; U\left(\frac{R_n(n\cdot)}{n^{1/(4-\eta)}}, \emptyset\right)\right) \xrightarrow{\text{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\leq t\leq 1}, \; U\left(0, \Xi\right)\right), \quad (1.1.10)$$

as $n \to \infty$, where the convergence in the first coordinate is in $\mathbf{C}([0,1],\mathbb{R})$ endowed with the topology of uniform convergence, and the convergence in the second is in the space of non-empty, compact subsets of $[0,1]\times\mathbb{R}$ endowed with the Hausdorff topology.

1.1.3 Further directions

We conclude this section with a list of possibilities for future study.

1. The scaling limit of a discrete snake depends on both the offspring distribution, and the displacement distributions of the branching random walk. An interesting avenue for future study would be to prove a scaling limit for discrete snakes of branching random walks with critical offspring distributions that are in the domain of attraction of α-stable laws, and have globally centered displacements. This would generalize the work of Marzouk [86] which proves a scaling limit for such discrete snakes when the displacements are IID and centered.

   Theorem 1.1.13 (Theorem 1.1 of Marzouk [86])**.** *Let $\alpha \in (1,2]$ and $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution in the domain of attraction of an α-stable law. For each $k \geq 1$, let ν_k be the law of k IID copies of a random variable Y with $\mathbf{E}[Y] = 0$ and $\mathrm{var}(Y) = \sigma^2$. Let $(b_n)_{n\geq 1}$ be as in (1.1.4). Then,*

   $$\left(\frac{b_n}{n}H_n(nt), \left(\frac{b_n}{n}\right)^{1/2} R_n(nt)\right)_{0\leq t\leq 1} \xrightarrow{\text{d}} (\mathbf{h}_t, \mathbf{r}_t)_{0\leq t\leq 1},$$

   in the sense of finite-dimensional distributions where $(\mathbf{h}_t)_{0\leq t\leq 1}$ is as in (1.1.5) and conditionally on $(\mathbf{h}_t)_{0\leq t\leq 1}$, $(\mathbf{r}_t)_{0\leq t\leq}$ is a centered Gaussian process with covariance function

   $$\mathrm{cov}(\mathbf{r}_s, \mathbf{r}_t) = \inf_{u\in[s\wedge t, s\vee t]} \mathbf{h}_u.$$

The convergence also holds in distribution in $\mathbf{C}([0,1],\mathbb{R}^2)$ *endowed with the topology of uniform convergence if and only if*

$$\mathbf{P}\left\{|Y| < (n/b_n)^{1/2}\right\} = o(n^{-1/2}).$$

Our methods for proving tightness in Theorem 1.1.11 do not immediately extend to this setting, as we rely heavily on the condition that $\mathbf{E}[\xi^3] < \infty$. However, it is likely that the convergence of the random-finite dimensional distributions holds given the convergence of a discrete line-breaking construction of T_n to a line-breaking construction of an α-stable tree. A stable tree analogue of Aldous' line-breaking construction was proved by Goldschmidt and Haas [48], and the convergence of a discrete line-breaking construction of T_n to said analogue is the topic of future work of Goldschmidt and Hill.

We note that this direction for future study is closely related to forthcoming work of Duquesne and Rebei [38] which proves limit theorems for discrete snakes whose jumps are centered, sibling independent and such that the underlying family tree converges to a Lévy tree.

2. Similarly to the above, it would be interesting to prove a scaling limit for the discrete snakes of branching random walks with a critical offspring distribution that has finite variance, and globally centered displacements that, in the limit, yield a stable process rather than a Brownian motion along any given lineage.

3. While the tail condition in [**A2**] is certainly necessary to achieve the convergence in Theorem 1.1.11 in $\mathbf{C}([0,1],\mathbb{R}^4)$, it is not clear to us whether the condition $\mathbf{E}[\xi^3] < \infty$ is necessary or just an artefact of our approach in proving tightness. Since Marzouk [86] proved that Theorem 1.1.8 could be extended to hold provided μ had finite variance, Theorem 1.1.11 does not cover the IID case completely. Furthermore, in the proof of Theorem 1.1.12, to prove convergence of the random finite-dimensional distributions, we require a local central limit theorem that assumes a finite third moment for the offspring distribution. It is possible that this local central limit theorem can be relaxed to only require a $2+\varepsilon$ moment for μ, but we do not pursue this in our work.

1.2 Temporal connectivity of random geometric graphs

Temporal graphs form another collection of models of random graphs with additional structure on the vertices. Briefly, temporal graphs are models for networks where chains of transmissions can only occur along interactions that are increasing in time. Typical examples of such real-world systems include the spread of disease or information in a community. More specifically, a temporal graph is an edge-labelled graph where the underlying graph is simple, and the edge labels are obtained by taking independent, uniform samples from $[0, 1]$. Much work has been conducted to study temporal connectivity – the property that there is a path composed of non-decreasing edge labels between all pairs of vertices in the graph – in the setting where the underlying graph is an Erdős–Rényi graph, see [14, 17, 22, 30]. However, to better model epidemiological processes, the underlying graph model must allow for the spatial closeness of individuals to affect interaction probabilities. As such, a stronger model for such networks is a temporal graph where the underlying graph is a random geometric graph. In Chapter 3, we prove asymptotic thresholds for temporal connectivity of a wide class of soft random geometric graphs (including hard random geometric graphs) in dimensions $d \geq 2$. More specifically, we show that temporal connectivity of random geometric graphs with n vertices only occurs when the average degree becomes of the order of $n^{1/(d+1)}$. This demonstrates that temporal connectivity of random geometric graphs occurs much later than simple connectivity, which occurs when the average degree becomes of the order $\log(n)$.

1.2.1 Background

In this section we provide a brief background on temporal random graphs. We will begin with a brief discussion of random simple temporal graphs, and their connectivity thresholds. Following this, we will introduce temporal random geometric graphs — the main object of study in Chapter 3.

1.2.1.1 Random simple temporal graphs

A *temporal graph* is an edge-labeled graph $\mathcal{G} = (G, \sigma)$ where the underlying graph $G = (V(G), E(G))$ is a finite simple graph and $\sigma : E(G) \to \{1, \dots, |E(G)|\}$ is a uniform ordering of the edges. Equivalently, one can generate a temporal graph using independent, uniform labels: each edge $e \in E(G)$ is assigned a label $\tau_e \sim \text{Unif}[0, 1]$. For $e \in E(G)$ we say that τ_e is the *time-stamp* of e, and for $f \in E(G) \setminus \{e\}$, we say that e *precedes* f if $\tau_e \leq \tau_f$. We use the latter construction in the sequel and denote the resulting graph by $\mathcal{G} = (G, (\tau_e)_{e \in E(G)})$.

For vertices $u, v \in V(G)$ we say that there is a *temporal path* from u to v, denoted $u \xrightarrow{\mathcal{G}} v$, if there exists a path $u = u_0, u_1, \dots, u_k = v$ from u to v in G with non-decreasing edge time-stamps, that is, $\tau_{(u_{i-1}, u_i)} \leq \tau_{(u_i, u_{i+1})}$ for all $i \in [k-1]$. We say that a vertex $u \in \mathcal{G}$ is a *temporal source* if there exist temporal paths $u \xrightarrow{\mathcal{G}} v$ for all $v \in \mathcal{G}$. Moreover, we say that $\mathcal{G}$ is *temporally connected* if every vertex of $\mathcal{G}$ is a temporal source.

The temporal graph obtained by taking the underlying graph to be an Erdős–Rényi graph, that is taking $G = G(n,p)$, is called a *random simple temporal graph.* Temporal structure aside, Erdős–Rényi graphs undergo a phase transition [41]: if $np \to c > 1$ as $n \to \infty$ then with high probability, $G(n,p)$ has a unique giant component, while if $np \to c < 1$ as $n \to \infty$ then the connected components of $G(n,p)$ have size $O(\log(n))$ with high probability. Furthermore, there is a sharp threshold for connectivity at $p \sim \log(n)/n$. The simple random temporal graph was first studied by Angel, Ferber, Sudakov and Tassion [12], who proved an asymptotically tight result for the length of the longest increasing path when $p = o(1)$. Casteigts, Raskin, Renken and Zamarev [31] later established a phase transition for temporal connectivity of simple random temporal graphs.

Theorem 1.2.1 (Theorem 5.4 of Casteigts et al., [31])**.** *The function* $\frac{3\log(n)}{n}$ *is a sharp threshold for temporal connectivity of simple random temporal graphs. More specifically, for any sufficiently large* n*, a simple random temporal graph* $\mathcal{G}_{n,p} = (G(n,p), (\tau_e)_{e \in E(G(n,p))})$

(i) is with high probability not temporally connected if $p \leq \frac{3\log(n)}{n} - \frac{6(\log(n))^{0.8}}{n}$;

(ii) is with high probability temporally connected, if $p \geq \frac{3\log(n)}{n} + \frac{3(\log(n))^{0.8}}{n}$.

Note that the threshold for temporal connectivity of simple random temporal graphs is a constant multiple of that for simple connectivity of Erdős–Rényi graphs. More recently Broutin, Kamčev and Lugosi [22] studied the shortest and longest increasing paths between typical vertices in $\mathcal{G}_{n,p}$, and Atamanchuk, Devroye and Lugosi [14] studied the size of the largest temporal clique.

1.2.1.2 Temporal random geometric graphs

While simple random temporal graphs capture the temporal structure of interactions, they do not account for the influence of spatial closeness between individuals on transmission. In Chapter 3 we initiate the study of *temporal random geometric graphs*. In this model, the vertices of the underlying graph G correspond to randomly drawn points in a Euclidean space, and pairs of points (u, v) are connected by an edge with a probability dependent on the Euclidean distance between u and v.

We first define a *random geometric graph*. Fix $n \geq 1$ and $d \geq 2$. Let $K : \mathbb{R}_+ \to [0,1]$, $r_n \in \mathbb{R}_+$, and let $\|\cdot\|$ denote the Euclidean distance on the torus $[0,1]^d$. A *soft random geometric graph* with n vertices in dimension d, denoted by $G_n = (\mathcal{X}_n, K, r_n)$, is a graph on n uniform points $\mathcal{X}_n$ of the unit torus $[0,1]^d$, such that for every pair of vertices $u, v \in \mathcal{X}_n$, $u \neq v$, u and v are neighbours in $\mathcal{G}_n$ with probability

$$\mathbf{P}\left\{(u,v) \in E(G_n) \mid \mathcal{X}_n\right\} = K\left(\frac{\|u-v\|}{r_n}\right).$$

The edges are assumed to be conditionally independent given $\mathcal{X}_n$. When $K(x) = \mathbf{1}_{[x\leq 1]}$, G_n is called a *hard random geometric graph.*

Penrose [93] proved that as $n \to \infty$ the probability of connectivity of soft random geometric graphs is governed by the property of having no isolated vertices, and moreover that random geometric graphs become connected when the average degree of the graph is of the order $\log(n)$. We refer the reader to Penrose [93] for a precise statement of this result.

A *temporal random geometric graph* with n vertices in dimension d is a temporal graph $\mathcal{G}_n = (G_n, (\tau_e)_{e\in E(G_n)})$ where $G_n = (\mathcal{X}_n, K, r_n)$ is a soft random geometric graph on n vertices on the torus $[0,1]^d$.

This model is better suited to model epidemiological processes than simple random temporal graphs as the geometry allows for spatial closeness of individuals to affect interaction probabilities.

1.2.2 Main results

In Chapter 3 we prove the following theorem establishing an asymptotic threshold for temporal connectivity of temporal random geometric graphs.

Theorem 1.2.2. *Let $\mathcal{G}_n = (G_n, (\tau_e)_{e\in E(G_n)})$ be a temporal random geometric graph with $G_n = (\mathcal{X}_n, K, r_n)$ satisfying the following:*

[**A1**] *For $x > 1$, $K(x) \leq \beta x^{-d} e^{-2(x+1)\log(x+1)}$ for some $\beta > 0$.*

[**A2**] *There exists $\alpha > 0$ such that $K(x) \geq \alpha$ for all $x \leq 1$.*

Then, there exist constants $c_d, C_d > 0$ such that

(i) if $r_n \leq c_d n^{-1/(d+1)}$ for all n sufficiently large, then with high probability $\mathcal{G}_n$ is temporally disconnected,

(ii) if $r_n \geq C_d n^{-1/(d+1)}$ for all n sufficiently large, then with high probability $\mathcal{G}_n$ is temporally connected.

The assumptions [**A1**] and [**A2**] ensure respectively that there are not too many long edges in the soft random geometric graph, and that there are sufficiently many vertices at graph distance $O(r_n)$ to any vertex.

We note that the by monotonicity of temporal connectivity, Theorem 1.2.2 (i) holds under just assumption [**A1**] and Theorem 1.2.2 (ii) holds under just assumption [**A2**]. Moreover, the choice of $K(x) = \mathbf{1}_{[x\leq 1]}$ satisfies both [**A1**] and [**A2**], and therefore Theorem 1.2.2 yields an asymptotic threshold for temporal connectivity of temporal hard random geometric graphs.

As discussed in the previous subsection, random geometric graphs become connected when the average degree of the graph is of the order $\log(n)$, [93]. Our results show that temporal connectivity only occurs when the average degree becomes $n^{1/(d+1)}$. As such, the threshold for temporal connectivity of temporal random geometric graph is markedly different than that of simple random temporal graphs where temporal connectivity and simple connectivity both occur when the average degree is of the order $\log(n)$.

1.2.3 Future directions

We conclude this section with a list of possibilities for future study.

1. Since the existence of temporal paths in $\mathcal{G}_n$ is neither a transitive nor a symmetric property, it would be interesting to study thresholds for weaker forms of temporal connectivity. For example, the threshold for the existence of a temporal path between any to fixed vertices and the threshold for a typical vertex to be a temporal source. This would build upon previous work of Becker et al. [17] which proves thresholds for such weaker forms of temporal connectivity for simple random temporal graphs.

2. Our results determine the order of magnitude of the temporal connectivity threshold. It is natural to conjecture that, similarly to the case of Erdős–Rényi random graphs, there is a sharp threshold for temporal connectivity of random geometric graphs. More specifically, that there exists a constant $\kappa_{K,d}$ such that for all $\varepsilon > 0$, if $r_n \leq (\kappa_{K,d} - \varepsilon)n^{-1/(d+1)}$, then with high probability $\mathcal{G}_n$ is temporally disconnected, while for $r_n \geq (\kappa_{K,d} + \varepsilon)n^{-1/(d+1)}$, with high probability $\mathcal{G}_n$ is temporally connected. This conjecture is the topic of future work by Angel, Bassan, Kerriou, and Jorritsma.

1.3 Random walks on Coxeter interchange graphs

In Chapter 4 we change focus and study the asymptotic behaviour of random processes on deterministic graphs. More specifically, we show that there is an efficient Markov chain Monte Carlo algorithm to sample from a set of generalised tournaments called

Coxeter tournaments. Tournaments are orientations of the complete graph on n vertices. Each directed edge in the oriented graph symbolizes a "competitive" game which is played between the vertices at its endpoints, and is oriented away from the winner. Each tournament is associated with a *score sequence* $s = (s_1, \ldots, s_n)$, which specifies the number of points earned by each player (or rather vertex) in the tournament. In precise terms, $s_i = (\deg^+(i) - \deg^-(i))/2$, where $\deg^{\pm}(i)$ is the out/in degree of vertex i. The set of all score sequences was studied by Landau [73] and is well-understood. However, the set of tournaments with a given score sequence is combinatorially rich in general, and appears difficult to describe in precise terms. Since this set is difficult to understand precisely, many have instead tried to understand how to approximately sample a uniform tournament with a given score sequence. Kennan, Tetali and Vempala [64] used simple random walks as a means of Markov chain Monte Carlo sampling from the set of tournaments with a given score sequence, and upper bounded the time it takes the random walk to reach stationarity in the case when the score sequence is close to regular (that is, when all players win approximately the same number of games). McShine [89] later built upon their result, using Bubley and Dyer's method of path coupling [25] to approximately sample from the set of tournaments with a given score sequence, for any score sequence. Loosely speaking, Bubley and Dyer's method of path coupling shows that for an aperiodic, irreducible discrete-time Markov chain on the vertices of a connected graph, if there exists a contracting coupling of the Markov chain (i.e., in expectation the distance between some coupling of copies of the Markov chains decreases in one time step), then we can obtain an upper bound for the time it takes for the Markov chain to reach stationarity as a function of the size of the state space. If the mixing time grows polynomially with the size of the state space, we say that the Markov chain is *rapidly mixing*.

We discuss McShine's application [89] of path coupling with regard to tournaments in subsection 1.3.1.2. In Chapter 4, we use the same framework to approximately sample from the set of Coxeter tournaments. Briefly, Coxeter tournaments are tournaments where, in addition to competitive games, each player plays a collaborative game, which contributes to the scores of both players, and in some instances individual players all play solitary games, which each contribute solely to the score of their respective player. While the general framework that we employ in Chapter 4 is the same as that used in [89], the addition of collaborative and solitary games results in a much more intricate analysis.

In certain models where the state spaces of the Markov chains have an inductive structure (for example any sub-tournament of a tournament is also an (incomplete) tournament), establishing rapid mixing leads to a way to approximately count the number of elements in the state space. We note that while the results of McShine [89] and the results of Chapter 4 establish rapid mixing for a sequence of Markov chains, the models studied lack the necessary structure to immediately obtain an approximate count of the number of tournaments with a given score sequence. We leave this as an interesting direction for future study.

1.3.1 Background

In this section we provide the necessary background on tournaments and their Coxeter analogues, the latter being the main object of study in Chapter 4. We begin with a discussion of standard tournaments, followed by a discussion of previous work on sampling tournaments with a given score sequence using path coupling. We then introduce signed graphs, Coxeter tournaments, and Coxeter interchange graphs.

1.3.1.1 Tournaments

We begin with some definitions. A *tournament* T is an orientation of the complete graph on n vertices, K_n. We encode a tournament $T = \{w_{ij}, i > j\}$ using values $w_{ij} \in \{0, 1\}$. If $w_{ij} = 1$ we orient the edge $\{i, j\}$ as $i \to j$, and otherwise $i \leftarrow j$ if $w_{ij} = 0$. We think of vertices as players, and edges as competitive games, directed away from the winner.

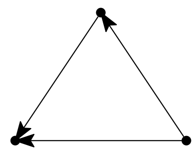

Figure 1.2: A tournament T on $n = 3$ vertices with $\mathbf{w}(T) = (0, 1, 2)$.

The *win sequence* of a tournament T,

$$\mathbf{w}(T) = \sum_{i>j} [w_{ij}\mathbf{e}_i + (1 - w_{ij})\mathbf{e}_j],$$

lists the total number of wins by each player, where $\mathbf{e}_i \in \mathbb{Z}^n$ are the standard basis vectors. See Figure 1.2 for an example. The *standard win sequence*, corresponding to the *transitive* (acyclic) tournament, in which all $w_{ij} = 1$, and so player i wins against all players $j < i$ of smaller index, is given by

$$\mathbf{w}_n = (0, 1, \ldots, n - 1).$$

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$, we say that $\mathbf{x}$ is *majorized* by $\mathbf{y}$, and write $\mathbf{x} \preceq \mathbf{y}$, if and only if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ and $\sum_{i=1}^k (\mathbf{x}^{\downarrow})_i \leq \sum_{i=1}^k (\mathbf{y}^{\downarrow})_i$, for all $1 \leq k \leq n$, where for $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z}^{\downarrow}$ is the (weakly) decreasing rearrangement of $\mathbf{z}$. We also write $\mathbf{z}_{\uparrow}$ for the (weakly) increasing rearrangement of $\mathbf{z}$. Landau [73] showed that $\mathbf{w} \in \mathbb{Z}^n$ is a win sequence if and only if $\mathbf{w} \preceq \mathbf{w}_n$. That is, the set of win sequences is

$$\mathrm{Win}(n) = \{\mathbf{w} \in \mathbb{Z}^n \ : \ \mathbf{w} \preceq \mathbf{w}_n\}.$$

More explicitly, $\mathbf{w} = (w_1, \ldots, w_n)$ is a win sequence if and only if it satisfies $\sum_{i=1}^n w_i = \binom{n}{2}$ and all partial sums $\sum_{i=1}^k (\mathbf{w}_{\uparrow})_i \geq \binom{k}{2}$. These conditions are clearly necessary, and while there are many proofs in the literature that they are sufficient, this is perhaps most easily seen through the Havel-Hakimi algorithm [51, 54] and standard techniques from majorization theory.

In Chapter 4 it will be convenient to make a linear shift, and consider instead the *score sequence* of a tournament T, $\mathbf{s}(T) := \mathbf{w}(T) - ((n-1)/2)\mathbf{1}_n$ where $\mathbf{1}_n = (1, \ldots, 1) \in \mathbb{Z}^n$. We note that

$$\mathbf{s}(T) = \sum_{i>j} (w_{ij} - 1/2)(\mathbf{e}_i - \mathbf{e}_j).$$

We denote by $s_n = w_n - ((n-1)/2)\mathbf{1}_n$ the *standard score sequence*, and if $\mathbf{s}(T) = \mathbf{0}_n = (0, \ldots, 0) \in \mathbb{Z}^n$, we say that T is *neutral*. Furthermore, we let $\mathrm{Score}(n)$ denote the set of all possible score sequences for tournaments with n players.

Although by Landau [73] the set of score sequences $\mathrm{Score}(n)$ is well understood, it appears that the set of all tournaments with a given score sequence $\mathbf{s} = (s_1, \ldots, s_n)$, which we henceforth denote by $\mathrm{Tour}(n, \mathbf{s})$, is combinatorially rich in general, and difficult to describe in simple precise terms. For instance, the enumerative information in the literature, beginning with Spencer [101] and then McKay et al. [56, 87, 88], requires to varying degrees that $\mathbf{s}$ be sufficiently close to $\mathbf{0}_n$. Kannan, Tetali and Vempala [64] were the first to use simple random walks as a means of Markov chain Monte Carlo sampling from $\mathrm{Tour}(n, \mathbf{s})$. McShine [89] subsequently used similar methods to generalize their result and approximately sample tournaments with a given score sequence. We discuss this work in the next subsection.

1.3.1.2 Sampling tournaments with a given score sequence

In this subsection, we will describe the work of McShine [89] which uses lazy simple random walks to Markov chain Monte Carlo sample from $\mathrm{Tour}(n, \mathbf{s})$. We begin with a brief discussion of mixing times and path coupling, a standard technique used for bounding mixing times.

Recall that an aperiodic, irreducible discrete-time Markov chain $(X_n)_{n \geq 0}$ on a finite state space Ω has a unique *equilibrium* distribution π on S such that for all

$x, y \in \Omega$,

$$p_n(x,y) := \mathbf{P}\{X_n = y \mid X_0 = x\} \to \pi(y),$$

as $n \to \infty$. We note that $\pi(y)$ is also the asymptotic proportion of time spent by the Markov chain at state $y \in \Omega$. The maximal total variation distance of the distribution of the walk at time n from π,

$$\tau(n) := \max_{x\in\Omega} \frac{1}{2} \sum_{y\in S} |p_n(x,y) - \pi(y)|$$

is non-increasing. The *mixing time* is then defined as

$$t_{\text{mix}} = \inf\{n \geq 0 \ : \ \tau(n) \leq 1/4\}.$$

A Markov chain is said to be *rapidly mixing* if t_{mix} is bounded by a polynomial in $\log|\Omega|$. Path coupling was introduced by Bubley and Dyer [25]. See for example Aldous and Fill [10, Section 12.1.12] or Levin et al. [79, Section 14.2] for reformulations of the original result that are closer in appearance to the following.

Theorem 1.3.1 (Bubley and Dyer [25])**.** *Consider a Markov chain $(X_n)_{n\geq 0}$ on the vertices of a connected graph $G = (V, E)$. Let δ denote the graph distance, which assigns weight $\delta(u,v) = 1$ to edges $\{u,v\} \in E$. Let w be a re-weighting, assigning some $w(u,v) \geq 1$ to $\{u,v\} \in E$. Let $w(x,y)$ be the re-weighted distance between x and y (that is, the minimal sum of the weight of the edges along a path between them in G). Let*

$$D_w = \max_{x,y\in V} w(x,y)$$

denote the re-weighted diameter of G. Suppose that for some $\alpha > 0$, for each $\{x', x''\} \in E$ there is a coupling $(X'_n, X''_n)_{n\geq 0}$ of two copies of the Markov chain $(X_n)_{n\geq 0}$ with $(X'_0, X''_0) = (x', x'')$ so that

$$\mathbf{E}\left[w(X'_1, X''_1)\right] \leq (1-\alpha) w(x', x'').$$

Then $(X_n)_{n\geq 0}$ mixes in time $t_{\text{mix}} = O(\alpha^{-1} \log D_w)$.

Before we discuss McShine's [89] application of path coupling, we will define the underlying graph whose vertex set forms the state space of the Markov chain studied in the work, and is called the *interchange graph.* For a score sequence $\mathbf{s} \in \text{Score}(n)$, Brualdi and Li [24] introduced the interchange graph $\text{IntGr}(n, \mathbf{s})$, which encodes the combinatorics of $\text{Tour}(n, \mathbf{s})$. In this graph, there is a vertex $v(T)$ for each $T \in \text{Tour}(n, \mathbf{s})$, and vertices $v(T_1), v(T_2)$ are joined by an edge if T_1 can be obtained by reversing the orientation of all directed edges in a single copy of a cyclic triangle $\Delta \subset T_2$, see Figure 1.3 for an example. To see that the set of tournaments with a given score sequence is closed under cyclic triangle reversals note that the cyclic triangle, denoted by Δ_c, is the smallest non-trivial neutral tournament. More specifically, each of the three players wins exactly one game against the other two, so $\mathbf{s}(\Delta_c) = \mathbf{0}_3$, and for a tournament T with a copy $\Delta \subset T$ of Δ_c, if we let $T * \Delta$ denote the tournament obtained from T by reversing the orientation of all directed edges in Δ, then since Δ is neutral, $\mathbf{s}(T * \Delta) = \mathbf{s}(T)$.

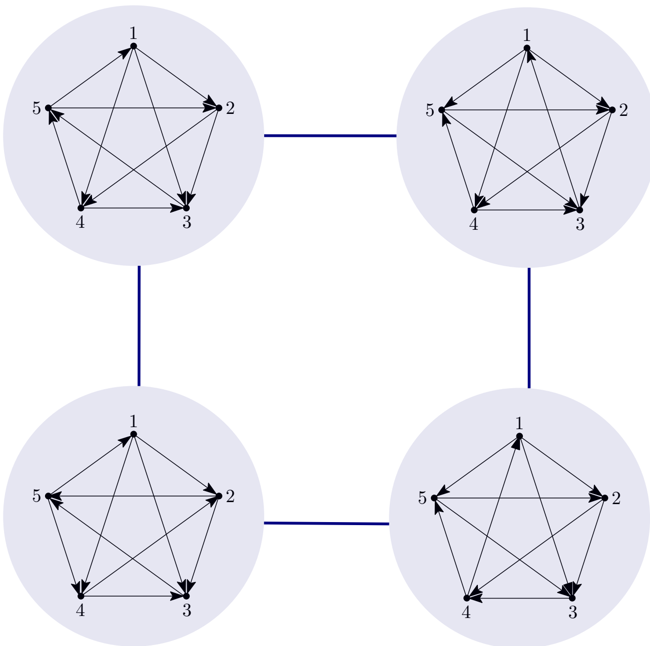

Figure 1.3: Interchange graph, $\mathrm{IntGr}(5, \mathbf{s})$ for $\mathbf{s} = (1, 0, -1, 0, 0)$.

In [89], McShine proves two fundamental properties of $\mathrm{IntGr}(n, \mathbf{s})$. The first is that for any n and any score sequence $\mathbf{s} \in \mathbb{R}^n$, $\mathrm{IntGr}(n, \mathbf{s})$ is connected and has diameter $O(n^2)$. The second fundamental property of $\mathrm{IntGr}(n, \mathbf{s})$ is that it is regular; in other words, any two tournaments with the same score sequence have the same number of cyclic triangles. This fact can be seen by counting non-cyclic triangles. Note that vertices $i, j, k \in T$ *do not* induce a copy of Δ_c if and only if there is a (unique) player that wins against the other two. As such, all tournaments with win sequence $\mathbf{w}$ have exactly

$$\binom{n}{3} - \sum_{i=1}^{n} \binom{w_i}{2}$$

cyclic triangles. Recalling that $\mathbf{s} = \mathbf{w} - ((n-1)/2)\mathbf{1}_n$, we find that $\mathrm{IntGr}(n, \mathbf{s})$ is regular, with degree

$$d(n, \mathbf{s}) = \frac{\|\mathbf{s}_n\|^2 - \|\mathbf{s}\|^2}{2} = O(n^3), \tag{1.3.1}$$

where $\|\cdot\|$ is the ℓ^2 norm.

Consequently, to apply path coupling to approximately sample from $\mathrm{Tour}(n, \mathbf{s})$, a natural Markov chain candidate for $(X_n)_{n\geq 0}$ is the lazy simple random walk on $\mathrm{IntGr}(n, \mathbf{s})$. Indeed, since the interchange graph is connected and $\mathrm{Tour}(n, \mathbf{s})$ is closed under triangle reversals, the lazy simple random walk is aperiodic, and irreducible, hence admits the uniform distribution as its stationary distribution. McShine [89] proves that there exists a contractive coupling $(X'_n, X'')_{n\geq 0}$ of copies of the lazy simple random walk on $\mathrm{IntGr}(n, \mathbf{s})$ and subsequently applies Theorem 1.3.1 with $w = \delta$ to obtain the following result.

Theorem 1.3.2 (McShine [89])**.** *Fix* $\mathbf{s} \in \mathrm{Score}(n)$*. Then the lazy simple random walk* $(X_n)_{n\geq 0}$ *on* $\mathrm{IntGr}(n, \mathbf{s})$ *mixes in time* $t_{\mathrm{mix}} = O(n^3 \log(n))$.

To define the contractive coupling, $(X'_n, X''_n)_{n\geq 0}$, it suffices to prove a third fundamental property of the interchange graph $\mathrm{IntGr}(n, \mathbf{s})$: For any two tournaments $v(T_1), v(T_2)$ such that $\delta(v(T_1), v(T_2)) = 1$, for every copy $\Delta_1 \subset T_1$ of Δ_c such that $T_1 * \Delta_1 \neq T_2$, there exists a unique copy of a triangle $\Delta_2 \subset T_2$ such that

$$\delta(v(T_1 * \Delta_1), v(T_2 * \Delta_2)) = 1. \tag{1.3.2}$$

Now, suppose that $\delta(X'_0, X''_0) = 1$. To couple (X'_1, X''_1), let $e' = \{X'_0, X'\}$ be a uniformly random element of the subset of the edge set of $\mathrm{IntGr}(n, \mathbf{s})$ comprised of the edges that are incident to X'_0, and let r be a Bernoulli(1/2) random variable. If $X''_0 = X'$, then let $X'_1 = X''_1 = X'_0$ if $r = 0$, and $X'_1 = X''_1 = X''_0$ otherwise. On the other hand, if $X''_0 \neq X'$, then as a consequence of (1.3.2) there exists a unique edge $e'' = \{X''_0, X''\}$ in $\mathrm{IntGr}(n, \mathbf{s})$ such that $\delta(X', X'') = 1$. Therefore, in this case we let $X'_1 = X'_0$ and $X''_1 = X''_0$ if $r = 0$, and $X'_1 = X'$, $X''_1 = X''$ otherwise. It follows that

$$\mathbf{E}\left[\delta(X'_1, X''_1)\right] = 1 - 1/d(n, \mathbf{s}).$$

In Chapter 4 we prove an analogue of Theorem 1.3.2 using the same framework, but for tournaments which allow for pairs of players to play collaborative games, and single players to play solitary games. The remainder of this section is dedicated to defining such tournaments, which are called *Coxeter tournaments*. More formally, Coxeter tournaments are orientations of signed graphs, which we define in the sequel.

1.3.1.3 Signed graphs and Coxeter tournaments

A signed graph is a graph in which every edge is endowed with a sign. Before we can define the class of signed graphs that we will use, we require a brief summary of root systems. We refer to Kolesnik and Sanchez [69, Section 2] for detailed information about root systems, and here provide only what is needed for the purposes of Chapter

4. Let $V = \mathbb{R}^n$ be a Euclidean vector space with inner product $\langle \cdot, \cdot \rangle$. For a non-zero vector $\mathbf{v} \in V$, let $s_{\mathbf{v}} : V \to V$ denote the automorphism of V which reflects through the hyperplane perpendicular to $\mathbf{v}$. That is, for $\mathbf{x} \in V$,

$$s_{\mathbf{v}}(\mathbf{x}) := \mathbf{x} - 2\frac{\langle \mathbf{v}, \mathbf{x} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v}.$$

A subset $\Phi \subset V$ of non-zero vectors that spans V is called a *root system* if Φ is closed under $s_{\mathbf{a}}$, for all $\mathbf{a} \in \Phi$, and the only other multiple of any $\mathbf{a} \in \Phi$ is $-\mathbf{a}$. If $2\langle \mathbf{a}, \mathbf{b} \rangle / \langle \mathbf{a}, \mathbf{a} \rangle \in \mathbb{Z}$ for all $\mathbf{a}, \mathbf{b} \in \Phi$, we say that Φ is *crystallographic.* If Φ is not the direct sum of two root systems we say that it is *irreducible.*

In this work, we will focus on the irreducible and crystallographic infinite families of root systems of types $\Phi = A_{n-1}, B_n, C_n$, and D_n given by

$$\begin{aligned}
A_{n-1} &= \{\mathbf{e}_i - \mathbf{e}_j \ : \ i \neq j \in [n]\}, \\
B_n &= \{\pm\mathbf{e}_i \pm \mathbf{e}_j \ : \ i \neq j \in [n]\} \cup \{\pm\mathbf{e}_i \ : \ i \in [n]\}, \\
C_n &= \{\pm\mathbf{e}_i \pm \mathbf{e}_j \ : \ i \neq j \in [n]\} \cup \{\pm 2\mathbf{e}_i \ : \ i \in [n]\}, \\
D_n &= \{\pm\mathbf{e}_i \pm \mathbf{e}_j \ : \ i \neq j \in [n]\}.
\end{aligned}$$

The associated *positive root systems* are

$$\begin{aligned}
A_{n-1}^+ &= \{\mathbf{e}_i - \mathbf{e}_j \ : \ i > j \in [n]\}, \\
B_n^+ &= \{\mathbf{e}_i \pm \mathbf{e}_j \ : \ i > j \in [n]\} \cup \{\mathbf{e}_i \ : \ i \in [n]\}, \\
C_n^+ &= \{\mathbf{e}_i \pm \mathbf{e}_j \ : \ i > j \in [n]\} \cup \{2\mathbf{e}_i \ : \ i \in [n]\}, \\
D_n^+ &= \{\mathbf{e}_i \pm \mathbf{e}_j \ : \ i > j \in [n]\}.
\end{aligned}$$

For any of the above positive root systems, Φ^+, to each $S \subset \Phi^+$ we associate a *signed graph of type* Φ, $\mathcal{S}$, with vertex set $[n]$ and signed edge set $E(\mathcal{S})$ which includes a

- negative edge e_{ij}^-, with endpoints i and j, for each $\mathbf{e}_{ij}^- := \mathbf{e}_i - \mathbf{e}_j \in S$,
- positive edge e_{ij}^+, with endpoints i and j, for each $\mathbf{e}_{ij}^+ := \mathbf{e}_i + \mathbf{e}_j \in S$,
- half edge e_i^h, with endpoint i, for each $\mathbf{e}_i \in S$,
- loop edge e_i^ℓ, with endpoint i, for each $2\mathbf{e}_i \in S$.

The geometric theory of signed graphs was pioneered by Zaslavsky [106–108]. The signed edges in $\mathcal{S}$ and vectors in $\mathcal{S}$ are in bijective correspondence, and so we will denote the vector corresponding to a signed edge e by $\mathbf{e}$.

Note that signed graphs of types B_n, C_n and D_n can have negative and positive edges, however, only signed graphs of type B_n can have half edges, and only signed graphs of type C_n can have loop edges. Classical graphs with vertex set $[n]$ correspond to A_{n-1} type signed graphs with only negative edges.

A *complete* Φ*-graph* associated with $S = \Phi^+$ is denoted by $\mathcal{K}_\Phi$. The complete graph K_n corresponds to $\mathcal{K}_{A_{n-1}}$. Recall that a tournament is an orientation of K_n. This notion can be extended to the setting of signed graphs in the following way, see [108].

Definition 1.3.3 (Coxeter tournament). *A Coxeter tournament $\mathcal{T}$ on a signed graph $\mathcal{S}$ is an orientation of the signed edges $e \in E(\mathcal{S})$, which, in our context, we refer to as games. The orientation of e, called its outcome, is encoded by $w_e \in \{0, 1\}$. We say that e is won if $w_e = 1$ and lost if $w_e = 0$.*

The score sequence of a Coxeter tournament $\mathcal{T}$ on a signed graph $\mathcal{S}$ is given by

$$\mathbf{s}(\mathcal{T}) = \sum_{e \in E(\mathcal{S})} (w_e - 1/2)\mathbf{e}.$$

In other words, points are awarded to players in the following way:

- Each negative edge e_{ij}^- is a *competitive game*. One player wins and the other loses a 1/2 point, contributing $(w_{ij}^- - 1/2)\mathbf{e}_{ij}^-$ to $\mathbf{s}$.
- Each positive edge e_{ij}^+ is a *collaborative game*. Both players win or lose a 1/2 point, contributing $(w_{ij}^+ - 1/2)\mathbf{e}_{ij}^+$ to $\mathbf{s}$.
- Each half edge e_i^h is a *(half edge) solitary game*. One player wins or loses 1/2 a point, contributing $(w_i^h - 1/2)\mathbf{e}_i^h$ to $\mathbf{s}$.
- Each loop e_i^ℓ is a *(loop) solitary game*. One player wins or loses a point, contributing $(w_i^\ell - 1/2)\mathbf{e}_i^\ell$ to $\mathbf{s}$.

The *standard score sequence* corresponding to the Coxeter tournament on $\mathcal{K}_\Phi$ in which all $w_e = 1$, is denoted by

$$s_\Phi = \sum_{e \in \Phi^+} \mathbf{e}/2.$$

When depicting Coxeter tournaments, we draw competitive games as directed edges, directed away from the winner, and we draw collaborative wins/losses as solid/dotted (undirected) lines. We draw solitary games as half edges (with only one endpoint), directed away/towards their endpoint if won/lost. In type C_n tournaments, we draw solitary wins/losses as solid/dotted (undirected) loops. See Figure 1.4 for examples.

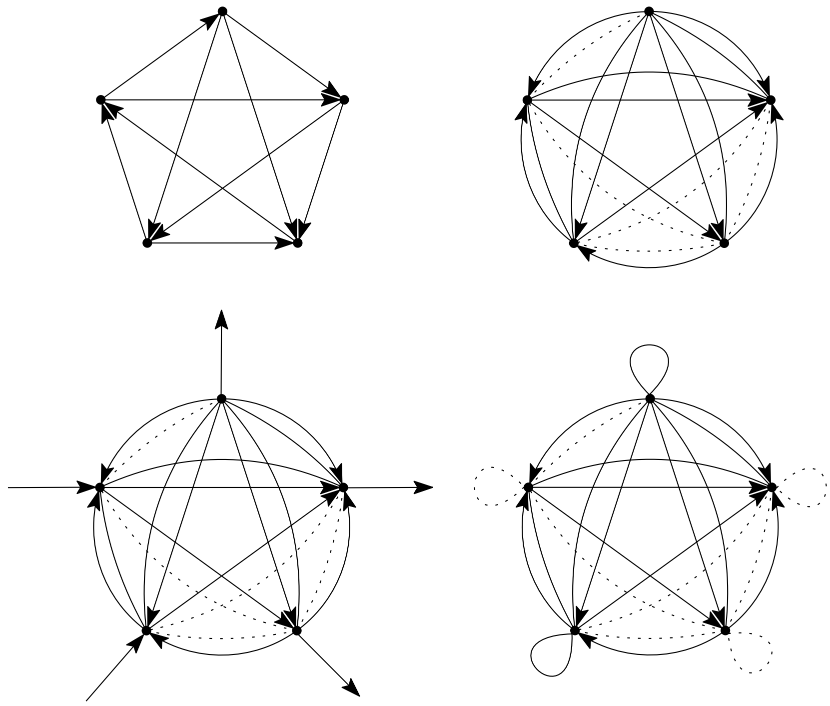

Figure 1.4: Top left: a standard tournament of type A_4; top right: a tournament of type D_5; bottom left: a tournament of type B_5; and bottom right: a tournament of type C_5.

In [68] (not included in this thesis) we classify the set of score sequences of Coxeter tournaments, generalising the result of Landau [73], see Theorem 1.3.4 below. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we say that $\mathbf{x}$ is *weakly sub-majorized* by $\mathbf{y}$, and write $\mathbf{x} \preceq_w \mathbf{y}$ if and only if $\sum_{i=1}^{k} (\mathbf{x}^{\downarrow})_i \leq \sum_{i=1}^{k} (\mathbf{y}^{\downarrow})_i$ for all $1 \leq k \leq n$. For $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we let $|\mathbf{x}| = (|x_1|, \ldots, |x_n|)$.

Theorem 1.3.4 (Theorem 4, [68]). *Let Φ be a root system of type B_n, C_n, or D_n. Then $\mathbf{s} \in \mathbb{R}^n$ is a score sequence of a Coxeter tournament on the complete Φ-graph $\mathcal{K}_\Phi$ if and only if $|\mathbf{s}| \preceq_w \mathbf{s}_\Phi$ and*

- *in B_n: $\mathbf{s} \in (\mathbb{Z} + 1/2)^n$,*
- *in C_n: $\mathbf{s} \in \mathbb{Z}^n$ and $\sum_{i=1}^{n} s_i \equiv \binom{n}{2} + n \mod 2$,*
- *in D_n: $\mathbf{s} \in Z^n$ and $\sum_{i=1}^{n} s_i \equiv \binom{n}{2} \mod 2$.*

The quantities $\binom{n}{2}+n$ and $\binom{n}{2}$ are equal to $\sum_i (\mathbf{s}_\Phi)_i$ in $\Phi = C_n$ and D_n respectively. There is no parity condition in B_n due to the 1/2 point (half edge) solitary games. We refer the reader to [68] for the proof of this result, and just note here that the proof is constructive, in that it shows how to *find* a Coxeter tournament $\mathcal{T}$ with given score sequence $\mathbf{s} \in \text{Score}(\Phi)$. In Chapter 4 this is particularly useful, as it allows us to initialise the random walk that we use to Markov chain Monte Carlo sample from the set of Coxeter tournaments with a fixed score sequence.

1.3.2 Main results

In [68] we also identify (see Figure 1.5)) the Coxeter tournaments which, in a sense, play the same role as the directed triangle in the standard tournament setting, and further obtain a formula (see Theorem 1.3.6 below) for the degree of the associated *Coxeter interchange graphs*, which we formally define below. Coxeter interchange graphs encode the combinatorial structure of the set of Coxeter tournaments with a given score sequence, and generalise the interchange graphs introduced by Brualdi and Li [24] discussed in subsection 1.3.1.2.

Recall that in the classical tournament setting, the set $\mathrm{Tour}(n, \mathbf{s})$ is generated by the cyclic triangle Δ_c. The generating sets in the Coxeter settings of types $\Phi = B_n, C_n$, and D_n contain the smallest, non-trivial neutral Coxeter tournaments.

In D_n, in addition to Δ_c, we require the *balanced triangle* Δ_b, involving a collaborative win, a collaborative loss, and a competitive game directed towards the collaborative win, as in Figure 1.5.

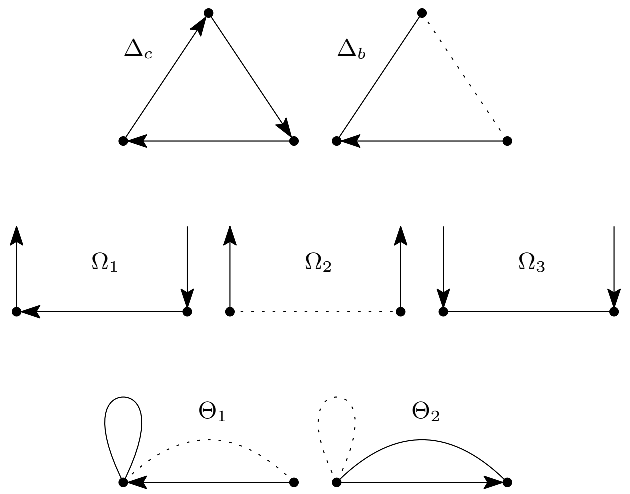

Figure 1.5: From top row to bottom row: the neutral tournaments Δ_c, Δ_b of types $D_n/B_n/C_n$; the neutral pairs $\Omega_1, \Omega_2, \Omega_3$ of type B_n; and the neutral clovers Θ_1, Θ_2 of type C_n. Note that Θ_1 and Θ_2 are reversals of one another.

In B_n and C_n, we require Δ_c, Δ_b and more. In B_n, there are three *neutral pairs* Ω_1, Ω_2, and Ω_3, as in Figure 1.5. In C_n, there are two *neutral clovers* Θ_1 and Θ_2, as in Figure 1.5. We call each of $\Delta_c, \Delta_b, (\Omega_i)_{i=1,2,3}, (\Theta_i)_{i=1,2}$ a *generator*.

The reversal $\mathcal{T}^*$ of a Coxeter tournament $\mathcal{T}$ on a signed graph $\mathcal{S}$ is obtained by reversing the outcome of all games (oriented signed edges) in $\mathcal{T}$. That is, $\mathcal{T}^* = \{w_e^* :$

$e \in E(\mathcal{S})\}$, where $w_e^* = 1 - w_e$. More generally, for $\mathcal{X} \subset \mathcal{T}$ we let $\mathcal{T} * \mathcal{X}$ denote the Coxeter tournament obtained from $\mathcal{T}$ by reversing the outcome of all the games in $\mathcal{X}$. Note that $\mathcal{T}^* = \mathcal{T} * \mathcal{T}$. See Figure 1.5 for an illustration of the reversal of a clover.

We say that $\mathcal{T}$ is *neutral* if $\mathbf{s}(\mathcal{T}) = \mathbf{0}_n$. Note that all generators are neutral. As such, reversing a copy of a generator in a Coxeter tournament does not change the tournament's score sequence.

Definition 1.3.5. *The Coxeter* Φ*-interchange graph* $\mathrm{IntGr}(\Phi, \mathbf{s})$ *has a vertex* $v(\mathcal{T})$ *for each* $\mathcal{T} \in \mathrm{Tour}(n, \mathbf{s})$*. Vertices* $v(\mathcal{T}_1)$*,* $v(\mathcal{T}_2)$ *are neighbours if* $\mathcal{T}_2 = \mathcal{T}_1 * \mathcal{G}$ *for some copy of a type* Φ *generator* $\mathcal{G}$ *in* $\mathcal{T}_1$*. When* $\mathcal{G}$ *is a clover, there is a double edge between neighbours, and a single edge otherwise.*

The double edges, associated with clover reversals, are added to maintain regularity. This should perhaps not be unexpected, since clovers are the only generators that include loop games, worth ± 1 point rather than $\pm 1/2$. In Chapter 4 we show that $\mathrm{IntGr}(\Phi, \mathbf{s})$ is connected and we bound its diameter. The following theorem from [68] extends extends the previous result of McShine [89] (as started in (1.3.1)).

Theorem 1.3.6 (Theorem 5 of [68])**.** *Let* $\Phi = B_n, C_n$ *or* D_n *and* $\mathbf{s} \in \mathrm{Score}(\Phi)$*. Then the Coxeter interchange graph* $\mathrm{IntGr}(\Phi, \mathbf{s})$ *is regular, with degree*

$$d(\Phi, \mathbf{s}) = \frac{\|\mathbf{s}_\Phi\|^2 - \|\mathbf{s}\|^2}{2},$$

where $\mathbf{s}_\Phi$ *is the standard score sequence in* Φ*.*

This result will be applied in Chapter 4 analogously to its application by McShine [89] to prove that there exists a contractive coupling of the lazy simple random walk on the interchange graph $\mathrm{IntGr}(n, \mathbf{s})$.

More specifically, in Chapter 4 we prove the following theorem.

Theorem 1.3.7. *Let* $\Phi = B_n, C_n$ *or* D_n*. Fix any* $\mathbf{s} \in \mathrm{Score}(\Phi)$*. Then the lazy simple random walk,* $(\mathcal{T}_n : n \geq 0)$ *on* $\mathrm{IntGr}(\Phi, \mathbf{s})$ *mixes in time* $t_{\mathrm{mix}} = O(n^3 \log(n))$ *if* $\Phi = B_n$ *or* D_n*, and in time* $t_{\mathrm{mix}} = O(n^4 \log(n))$ *if* $\Phi = C_n$*.*

We note it is possible to initialize this random walk since the classification of $\mathrm{Score}(\Phi)$ in Theorem 1.3.4 is constructive.

In Chapter 4 we will actually prove sharper bounds than those written in Theorem 1.3.7. More specifically, in types B_n and D_n we will show that $t_{\mathrm{mix}} = O(d \log(n))$, where d is the degree of the interchange graph $\mathrm{IntGr}(\Phi, \mathbf{s})$. The result above follows since by Theorem 1.3.6, $d = O(n^3)$. In type, C_n, we will show that $t_{\mathrm{mix}} = O(\gamma d \log(n))$ where $\gamma > 0$ is a certain quantity, formally defined in Chapter 4 satisfying $\gamma \leq$

$\min\{d, 2n\}$. We call γ the *maximal crystal degree* of the interchange graph as roughly speaking, it is the maximal number of *crystals* in $\mathrm{IntGr}(C_n, \mathbf{s})$ – these are copies of a special subgraph that can be found in interchange graphs of type C_n. The quantity γ is related to the re-weighting w we will use in applying Theorem 1.3.1 to prove Theorem 1.3.7 in type C_n.

In proving Theorem 1.3.7 we also show that all Coxeter interchange graphs are connected and we bound their diameter.

Theorem 1.3.8. *Let* $\Phi = B_n, C_n$ *or* D_n*. Fix any* $\mathbf{s} \in \mathrm{Score}(\Phi)$*. The Coxeter interchange graph* $\mathrm{IntGr}(\Phi, \mathbf{s})$ *is connected and its diameter* $D = O(n^2)$.

Further, in constructing our random walk couplings to apply Theorem 1.3.1 in the proof of Theorem 1.3.7, we uncover various other fine, structural properties of the graphs $\mathrm{IntGr}(\Phi, \mathbf{s})$, and hence the sets $\mathrm{Tour}(\Phi, \mathbf{s})$ which might be of independent (algebraic, geometric, etc.) interest.

1.3.3 Future directions

We conclude with a list of possibilities for future study as they appear in [27].

1. It remains open to find lower bounds for the mixing time, and to determine whether our bounds are sharp. In types A_{n-1}, B_n and D_n, we might conjecture so, at least up to logarithmic factors.

   Recall that, in these types, we have shown that $t_{\mathrm{mix}} = O(d \log n)$, for *any* score sequence, where d is the degree of the interchange graph. In type A_{n-1}, Sarkar [99] has shown that $t_{\mathrm{mix}} = \Omega(n^3)$ for a special class of score sequences with $d = \Theta(n^3)$ and a "bottleneck" that is simple to analyze. A similar argument also works in the other types B_n, C_n and D_n. Perhaps at least $t_{\mathrm{mix}} = \Omega(d)$ can be shown to hold in general.

   As discussed, in type A_n, Chen, Chang and Wang [33] have shown that the interchange graph is the hypercube, for some very specific score sequences. This shows, at least in some cases, that the bound $t_{\mathrm{mix}} = O(d \log n)$ is sharp, *with* the logarithmic factor.

   In type C_n, on the other hand, we have used a re-weighting of the metric to show that $t_{\mathrm{mix}} = O(\gamma d \log n)$, where γ is the maximal crystal degree. Perhaps other techniques can lead to an improvement. However, we think that crystals in type C_n are a genuine obstacle, so it might be surprising if, in fact, $t_{\mathrm{mix}} = O(d \log n)$ also in this type.

2. Recall that Theorem 4.4.1 shows that the interchange graphs are connected with diameter $D = O(n^2)$. It might be of theoretical interest to find a precise formula for D, or at least good bounds, as a function of $\mathbf{s}$. Note that Theorem 4.2.1 above (proved in [68]) gives such a formula for the degree d. We also note that in [24] some results are proved about the diameter D in type A_{n-1}.

3. Recall that each edge in the interchange graph corresponds to a generator reversal. Generators are the smallest neutral structures. It might be interesting to consider a generalization, in which neutral structures up to a given size can be reversed in a single step, and to quantify the decrease in the mixing time.

   Related to this, Gioan [45] has studied cycle and cocycle reversing systems, and these have been generalized by Backman [15, 16]. One might pursue Coxeter analogues of these results.

4. A *graphical zonotope* is a polytope obtained as a Minkowski sum of line segments, where the sum is indexed by the edges of the graph (see, e.g., Ziegler [109]). The permutahedron

$$\Pi_{n-1} = \mathbf{w}_n + \sum_{1 \leq i < j \leq n} [\mathbf{0}_n, \mathbf{e}_j - \mathbf{e}_i]$$

   is the graphical zonotope of the complete graph K_n. Likewise, the Coxeter versions Π_Φ are obtained as sums indexed by the edges in the complete signed graphs $\mathcal{K}_\Phi$. It could be interesting to study random walks on the fibers of other graphical zonotopes.

   That being said, our current arguments take full advantage of the symmetry of $\mathcal{K}_\Phi$. Once some edges become unavailable, can be more challenging (or even impossible) to show connectivity (and bound the diameter) of the interchange graph, and to devise a path coupling (which we have accomplished, via a non-trivial edge pairing argument).

5. Rapid mixing can be a starting point for approximate counting. It would be interesting if our result could help with counting the number of vertices in interchange graphs, for a general score sequence. As already discussed, these have been approximated (see [56, 87, 88, 101]) only in type A_{n-1} and when $\mathbf{s}$ is close to the center of Π_{n-1}.

6. In this work, we have studied random walks on interchange graphs associated with score sequences. However, one could also, quite naturally, try to study random walks on the set of score sequences itself. In type A_{n-1}, all lattice

points are score sequences. In types B_n, C_n and D_n the score sequences are more complicated sets of points, characterized in [68].

7. Finally, we note that the interchange graphs in Figures 4.3 and 4.4 are Cartesian products. Further, in type A_n, some interchange graphs are the hypercube [33], which are a simple example of a product graph. It might be enlightening to investigate the product structure of interchange graphs more generally.

Chapter 2

Discrete snakes with globally centered displacements

This chapter is based on joint work with Louigi Addario-Berry (McGill University), Serte Donderwinkel (University of Groningen) and Christina Goldschmidt (University of Oxford) which appears in the preprint [4].

2.1 Introduction

We consider a branching random walk whose genealogy is given by the family tree of a Bienaymé branching process (which we refer to as a *Bienaymé tree*) conditioned to have n vertices. We assume that the offspring distribution $\mu = (\mu_k)_{k\geq 0}$ is critical and has finite, non-zero variance, so that the genealogical tree has the Brownian continuum random tree as its scaling limit.[1] Each vertex of the tree is endowed with a spatial location in $\mathbb{R}$: the root is fixed to be at 0; for every other vertex, its location is obtained via a random displacement away from the location of its parent. The random displacements of children of distinct vertices will always be independent but, in general, the displacements of siblings may be dependent and may, moreover, depend on the vertex degree. For a vertex v with k children, the distribution of the vector of displacements from v to its ordered children is denoted by ν_k. In the sequel, $Y_k = (Y_{k,1}, \ldots, Y_{k,k})$ always denotes a random vector with law ν_k. In this paper, we explore conditions on μ and $\nu = (\nu_k)_{k\geq 1}$ such that the whole object converges to a Brownian motion indexed by the Brownian tree.

A convenient formulation is via the notion of a *discrete snake*. We imagine exploring the vertices of the tree one by one in depth-first order (we shall give precise definitions in Section 2.2 below) and record a list of the spatial locations of the

[1] To avoid technicalities, we shall also assume that the support of μ has greatest common divisor 1, so that the event that the tree has size n has strictly positive probability for all n large enough.

ancestors of the vertex we are currently visiting. In other words, the snake is a process taking values in the set of finite random walk paths (one should imagine it wiggling around as we explore the tree!). In fact, it turns out to be sufficient for many purposes to track the spatial location of the vertex that we are visiting only: this gives the so-called *head of the discrete snake*, which is our primary object of interest. We aim to prove convergence, after an appropriate rescaling, of the head of the discrete snake to the head of the Brownian snake driven by a normalised Brownian excursion (BSBE), first introduced by Le Gall [75, 76]. This is a stochastic process $(\mathbf{e}, \mathbf{r}) = (\mathbf{e}_t, \mathbf{r}_t)_{0\leq t\leq 1}$ taking values in $\mathbb{R}_+ \times \mathbb{R}$, such that $\mathbf{e}$ is a normalised Brownian excursion and, conditionally on $\mathbf{e}$, the second coordinate $\mathbf{r}$ is a centered Gaussian process taking values in $\mathbb{R}$ with covariance function

$$\operatorname{cov}(\mathbf{r}_s, \mathbf{r}_t) = \min_{u\in[s\wedge t, s\vee t]} \mathbf{e}_u. \tag{2.1.1}$$

Let us give some interpretation. For any pair of vertices in the Brownian tree, encoded by $s, t \in [0, 1]$, having heights $\mathbf{e}_s$ and $\mathbf{e}_t$, the spatial locations along their genealogical paths evolve as a common Brownian motion until their most recent common ancestor (which lies at distance $\min_{u\in[s\wedge t, s\vee t]} \mathbf{e}_u$ from the root) is reached, and they evolve as independent Brownian motions thereafter.

The problem of proving convergence of rescaled discrete snakes to the BSBE has been studied by a number of authors, under a wide range of different conditions on μ and $(\nu_k)_{k\geq 1}$. We shall give a review of the literature after we state our main results.

In order to obtain a Brownian limit for the displacements along a lineage, we require appropriate centering and moment conditions, which we now explain. Let ξ be a random variable with distribution μ and let $\bar{\xi}$ be a size-biased version, that is, having distribution $\bar{\mu} := (\bar{\mu}_k)_{k\geq 1}$, where for all $k \geq 1$,

$$\bar{\mu}_k = \frac{k\mu_k}{\mathbf{E}[\xi]} = k\mu_k.$$

(Recall that the offspring distribution μ is assumed to be critical, so that $\mathbf{E}[\xi] = 1$.)

Conditionally on $\bar{\xi}$, let $Y_{\bar{\xi}} = (Y_{\bar{\xi},1}, \ldots, Y_{\bar{\xi},\bar{\xi}})$ be $\nu_{\bar{\xi}}$-distributed and, independently, let $U_{\bar{\xi}}$ be a Uniform($[\bar{\xi}]$) random variable (where $[m] := \{1, 2, \ldots, m\}$). Then we say that the discrete snake is *globally centered* if

$$\mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\right] = 0.$$

In other words, the expected displacement of a uniform child of a vertex with a size-biased number of offspring is 0. We define the *global variance* to be

$$\beta^2 := \mathbf{E}\left[Y^2_{\bar{\xi},U_{\bar{\xi}}}\right],$$

and will prove our results under the condition that $\beta^2 < \infty$. Since distances in the tree scale as $n^{1/2}$, the spatial displacements along a lineage will then scale as $n^{1/4}$.

2.1.1 Main result

Denote by T_n a Bienaymé tree with offspring distribution μ, conditioned to have n vertices. Write $v(\mathrm{T}_n)$ for the vertex set of T_n and $\partial\mathrm{T}_n$ for its set of leaves. Conditionally given T_n, let $Y = (Y^{(v)}, v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n)$ be independent random vectors, such that if $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ has k children then $Y^{(v)}$ has distribution ν_k. Endow the vertices of T_n with spatial locations using the displacement vectors $Y^{(v)}$ as described above. We call the pair $\mathbf{T}_n = (\mathrm{T}_n, Y)$ a *(μ,ν)-branching random walk conditioned to have n vertices*, or simply a (μ,ν)-branching random walk.

Let $H_n = (H_n(i))_{0\le i\le n}$ and $\widetilde{H}_n = (\widetilde{H}_n(i))_{0\le i\le 2n}$ be the height and contour processes of T_n, respectively. Let $R_n(i)$ be the spatial location of the i-th vertex visited in a depth-first exploration of $\mathbf{T}_n$. We call the process (H_n, R_n) the *head of the discrete snake* (see Section 2.2 for a careful description of this). We may alternatively encode the endpoints of the random walk trajectories using the process $(\widetilde{H}_n, \widetilde{R}_n)$, where $\widetilde{R}_n(i)$ is the spatial location of the i-th vertex visited in a contour exploration of $\mathbf{T}_n$. (Compared to (H_n, R_n), this simply revisits some vertices.) We interpolate all of these functions linearly between integer times, which turns H_n and R_n into elements of $\mathbf{C}([0,n],\mathbb{R})$ and turns $\widetilde{H}_n$ and $\widetilde{R}_n$ into elements of $\mathbf{C}([0,2n],\mathbb{R})$.

We use two different notions of convergence for a sequence of random elements $(f_n)_{n\ge1}$ of $\mathbf{C}([0,1],\mathbb{R})$ such that $f_n(0) = f_n(1) = 0$ for all $n \ge 1$. Let $U_1, U_2, \ldots$ be IID Uniform([0,1]) random variables, independent of everything else. For $k \ge 1$, write $U^k_{(1)}, \ldots, U^k_{(k)}$ for the order statistics of $U_1, \ldots, U_k$. For another random element f of $\mathbf{C}([0,1],\mathbb{R})$ such that $f(0) = f(1) = 0$, we say that $f_n \xrightarrow{\mathrm{d}} f$ in the sense of *random finite-dimensional distributions* if, for every $k \ge 1$,

$$(f_n(U^k_{(1)}), \ldots, f_n(U^k_{(k)})) \xrightarrow{\mathrm{d}} (f(U^k_{(1)}), \ldots, f(U^k_{(k)}))$$

as $n \to \infty$. (We will discuss our choice of this notion of convergence in more detail below.) We will also use the stronger notion of convergence with respect to the topology generated by the uniform norm.

Theorem 2.1.1. *Let $\mu = (\mu_k)_{k\ge0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$. If $\nu = (\nu_k)_{k\ge1}$ is such that*

[A1] $$\mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\right] = 0 \quad \textit{and} \quad \beta^2 = \mathbf{E}\left[(Y_{\bar{\xi},U_{\bar{\xi}}})^2\right] < \infty,$$

then as $n \to \infty$ the following joint convergence holds in the sense of random finite-dimensional distributions:

$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_n(nt)}{n^{1/4}}, \frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{\widetilde{R}_n(2nt)}{n^{1/4}}\right)_{0\le t\le 1} \xrightarrow{\text{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t, \frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\le t\le 1}. \tag{2.1.2}$$

The convergence (2.1.2) holds in distribution in $\mathbf{C}([0,1],\mathbb{R}^4)$ *endowed with the topology of uniform convergence if, additionally,*

[**A2**] $\qquad \mathbf{P}\left\{\max_{1\le i\le \xi} |Y_{\xi,i}| > y\right\} = o(y^{-4})$ *as* $y \to \infty$ *and* $\mathbf{E}\left[\xi^3\right] < \infty$.

Theorem 2.1.1 follows immediately from Corollary 2.4.2 and Proposition 2.5.1 below. The analogue of Theorem 2.1.1 holds with $\mathbb{R}^d$-valued displacements for $d > 1$, and with essentially identical proofs to those in the current work; the only change in the conclusion is that the limit $\mathbf{r}$ of the rescaled spatial displacements takes values in $\mathbb{R}^d$ rather than in $\mathbb{R}$, and that β^2 should be interpreted as the covariance matrix of $Y_{\bar{\xi},U_{\bar{\xi}}}$.

Let $\Phi_n(i)$ be the random walk trajectory associated with path from the root to the i-th vertex visited in the contour exploration of $\mathbf{T}_n$, for $0 \le i \le 2n$. Then $(\widetilde{H}_n, \Phi_n)$ is the *discrete snake driven by* $\widetilde{H}_n$. By the homeomorphism theorem of Marckert and Mokkadem (Theorem 2.1 of [83]), Theorem 2.1.1 entails also that $(\widetilde{H}_n, \Phi_n)$ has the BSBE as its scaling limit; see Figure 2.1 for an illustration.

By [40, Corollary 2.5.1] and [82] it turns out that the convergence of the parametrisations of the head of the snake via the height and contour processes are essentially equivalent. In particular, in order to prove Theorem 2.1.1, it suffices to show that, under assumption [**A1**], we have

$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_n(nt)}{n^{1/4}}\right)_{0\le t\le 1} \xrightarrow{\text{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\le t\le 1}, \tag{2.1.3}$$

as $n \to \infty$ in the sense of random finite-dimensional distributions, and in $\mathbf{C}([0,1],\mathbb{R}^2)$ endowed with the topology of uniform convergence under the additional assumption [**A2**].

It is not clear to us whether the requirement that $\mathbf{E}\left[\xi^3\right] < \infty$ in [**A2**] is necessary or just an artefact of our approach to proving tightness. We shall see in the next subsection that the tail condition in [**A2**] *is* necessary.

2.1.2 Necessity of the tail condition

If we adjust assumption [**A2**] to allow for heavier tails, displacements start to appear near the leaves which are not negligible on the scale $n^{1/4}$. In this case, one can no

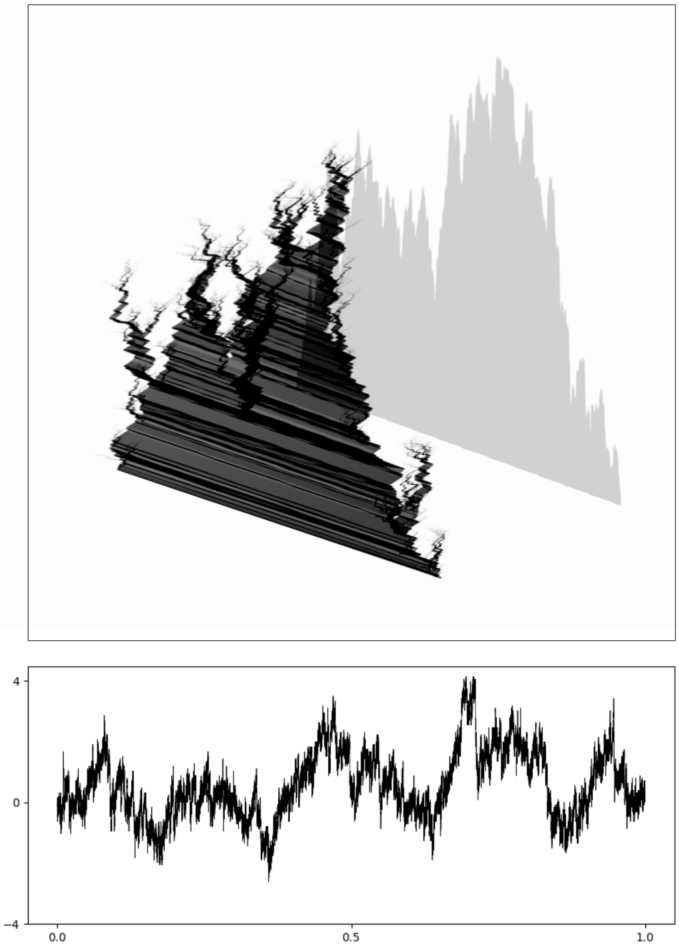

Figure 2.1: Top: a (rescaled) discrete snake; bottom: its head. The underlying tree is a size-conditioned Poisson(1) Bienaymé tree with $n = 25000$ vertices and deterministic displacement distributions given by (2.1.5), below. The area under the contour process of the underlying tree is illustrated by the gray shaded region.

longer expect a continuous limit process. This is a consequence of the following proposition on the largest displacement in the tree, which we prove in Section 2.8.2.

Proposition 2.1.2. *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$. If $\limsup_{y\to\infty} y^4 \mathbf{P}\left\{\max_{1\leq i\leq \xi} |Y_{\xi,i}| > y\right\} > 0$, then there exists a $\delta > 0$ such that*

$$\limsup_{n\to\infty} \mathbf{P}\left\{\max_{v\in v(\mathrm{T}_n)} \max_{j\geq 1} |Y_j^{(v)}| > \delta n^{1/4}\right\} > 0.$$

We may however still obtain a global convergence result on the appropriate scale, under an additional regularity assumption; see [**A3**] below. Since the large displacements from a vertex with k children need not be independent, in this setting the limit depends on the joint distribution of the displacements from a vertex to its children. For $k \geq 1$ and $j \in [k]$ denote by

$$Y_{k,j}^+ := Y_{k,j} \vee 0 \text{ and } Y_{k,j}^- := (-Y_{k,j}) \vee 0,$$

the positive and negative displacements of the j-th child of a vertex with k children, respectively. Further let $Y_k^+ := (Y_{k,j}^+)_{j\in[k]}$ and $Y_k^- = (Y_{k,j}^-)_{j\in[k]}$.

[**A3**] Suppose that $\mathbf{E}[\xi^3] < \infty$. Furthermore, suppose that there exists a Borel measure π on $\mathbb{R}_+^2 \setminus \{(0,0)\}$ such that for any $\varepsilon > 0$, both $\pi(\mathbb{R}_+ \times (\varepsilon, \infty)) < \infty$ and $\pi((\varepsilon,\infty) \times \mathbb{R}_+) < \infty$, and there exists $\eta \in [0,2)$ such that for all Borel sets $A \subset \mathbb{R}_+^2 \setminus \{(0,0)\}$ for which $\pi(\partial A) = 0$,

$$r^{4-\eta}\mathbf{P}\left\{\frac{1}{r}\left(\max_{1\leq i\leq \xi} Y_{\xi,i}^+, \max_{1\leq i\leq \xi} Y_{\xi,i}^-\right) \in A\right\} \to \pi(A)$$

as $r \to \infty$.

We note the following lemma, whose proof may be found in Section 2.8.2.

Lemma 2.1.3. *[**A3**] implies that the projection of π onto either of its coordinates has no atom in $(0,\infty)$.*

Under assumption [**A3**] we prove convergence results for the head of the discrete snake in the space of non-empty compact subsets of $[0,1] \times \mathbb{R}$ equipped with the Hausdorff topology. In what follows, for a continuous function $f : [0,1] \to \mathbb{R}$ and a set $S \subset [0,1] \times \mathbb{R}_+^2 \setminus \{(0,0)\}$, write $U(f,S)$ for the union of the graph of f and the vertical line segments $[(t, f(t) - y), (t, f(t) + x)]$ for each $(t,x,y) \in S$. The next theorem relates to the case $\eta = 0$ in [**A3**].

Theorem 2.1.4. *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$, and let $\nu = (\nu_k)_{k\geq 1}$ be such that [**A1**] holds and [**A3**] holds for a given measure π with $\eta = 0$. Then, taking Ξ to be a Poisson process on $[0,1] \times \mathbb{R}_+^2 \setminus \{(0,0)\}$ with intensity $dt \otimes \pi(dx,dy)$, we have*

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}\right)_{0\leq t\leq 1}, U\left(\frac{R_n(n\cdot)}{n^{1/4}}, \emptyset\right)\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\leq t\leq 1}, U\left(\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}, \Xi\right)\right), \quad (2.1.4)$$

as $n \to \infty$, where the convergence in the first coordinate is in $\mathbf{C}([0,1],\mathbb{R})$ endowed with the topology of uniform convergence, and the convergence in the second is in the space of non-empty, compact subsets of $[0,1] \times \mathbb{R}$ endowed with the Hausdorff topology.

We refer to the object on the right-hand side of (2.1.4) as the *hairy tour*, in keeping with the previous work of Janson and Marckert [60].

When $\eta \in (0, 2)$, the large jumps dominate the smaller ones to such an extent that, in the limit, we obtain a pure jump process.

Theorem 2.1.5. *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0, \infty)$, and let $\nu = (\nu_k)_{k\geq 1}$ be such that [**A1**] holds and [**A3**] holds for a given measure π with $\eta \in (0, 2)$. Then, taking Ξ to be a Poisson process on $[0, 1] \times \mathbb{R}_+^2 \setminus \{(0, 0)\}$ with intensity $dt \otimes \pi(dx, dy)$, we have*

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}\right)_{0\leq t\leq 1}, \ U\left(\frac{R_n(n\cdot)}{n^{1/(4-\eta)}}, \emptyset\right)\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\leq t\leq 1}, \ U(0, \Xi)\right),$$

as $n \to \infty$, where the convergence in the first coordinate is in $\mathbf{C}([0, 1], \mathbb{R})$ endowed with the topology of uniform convergence, and the convergence in the second is in the space of non-empty, compact subsets of $[0, 1] \times \mathbb{R}$ endowed with the Hausdorff topology.

In contrast to Theorem 2.1.1, in Theorems 2.1.4 and 2.1.5 we need the condition $\mathbf{E}[\xi^3] < \infty$ not just for tightness but also for the convergence of the random finite-dimensional distributions. The reason for this is that we apply a quantitative local central limit theorem which requires a third moment on the offspring distribution. (See Theorem 2.8.3 for the precise statement.)

The fact that we obtain a continuous function decorated by intervals in both Theorems 2.1.4 and 2.1.5 is really an artefact of the choice to interpolate R_n linearly between integer times. Indeed, the endpoints of the intervals capture the asymptotic behaviour of the two extremities of the displacements away from vertices, but tell us nothing about how the "point process" of displacements in between behaves. If we instead consider the graph of $\left(\frac{R_n(\lfloor nt\rfloor)}{n^{1/4}}\right)_{0\leq t\leq 1}$ in the case where we do not have $\mathbf{P}\{\max_{1\leq i\leq \xi} |Y_{\xi,i}| > y\} = o(y^{-4})$ there are, in fact, many possible behaviours. We will not undertake any sort of exhaustive classification here, but let us give a couple of illustrative examples.

Suppose first that the displacements are simply IID copies of a random variable Y such that, for some Borel measure π on $\mathbb{R}_+ \setminus \{0\}$ such that for any $\varepsilon > 0$, $\pi((\varepsilon, \infty)) < \infty$, we have $r^4\mathbf{P}\{Y \in rA\} \to \pi(A)$ as $r \to \infty$ for every Borel set $A \subset \mathbb{R} \setminus \{0\}$ such that $\pi(\partial A) = 0$. Then we will not, in the limit, observe two or more $\Theta(n^{1/4})$ displacements away from the same vertex of T_n (nor, indeed, from vertices at distance $o(n^{1/2})$ from one another), and so we just obtain the graph of $\mathbf{r}$ decorated by isolated points which occur as a Poisson process of intensity $dt \otimes \pi(dy)$ on $[0, 1] \times \mathbb{R} \setminus \{0\}$.

On the other hand, suppose that we have the following deterministic displacements:

$$Y_{k,j} = \sigma - \frac{2}{\sigma}(k-j) \text{ for } \quad 1 \le j \le k. \tag{2.1.5}$$

These displacements have a particular significance, which we will discuss in the next subsection. For the moment, let us just observe that it is straightforward to check that they are globally centered and of finite global variance whenever the offspring distribution is critical and admits a finite third moment. Suppose that there exists a Borel measure π_1 on $(0, \infty)$ with $\pi_1((\varepsilon, \infty)) < \infty$ for all $\varepsilon > 0$, such that $r^4 \mathbf{P}\{\xi \in rA\} \to \pi_1(A)$ as $r \to \infty$ for any Borel set $A \subset (0, \infty)$ with $\pi_1(\partial A) = 0$. Then all of the children of a vertex with $\Theta(n^{1/4})$ children will have $\Theta(n^{1/4})$ displacements which are regularly spaced with spacing $2/\sigma$. Again, with high probability, we will not see two vertices of degree $\Theta(n^{1/4})$ within distance $o(n^{1/2})$ in T_n. So in the limit for the graph of $\left(\frac{R_n(\lfloor nt \rfloor)}{n^{1/4}}\right)_{0 \le t \le 1}$ we will see decorations driven by a Poisson process on $[0,1] \times \mathbb{R}_+$ with intensity $dt \otimes \pi(dx)$ such that when we observe a point (t, x) of the Poisson process, we attach the whole interval $[-2x/\sigma, 0]$ to the graph of $\mathbf{r}$ at t.

2.1.3 Related work

As mentioned earlier, versions of the topic studied in this paper have received extensive attention in the literature. One reason for this is that discrete snakes play a crucial role in the study of random planar maps; see [1, 2, 32, 77, 81, 90].

The earliest discrete snake convergence results were proved in models with a fixed offspring distribution. Chassaing and Schaeffer [32] treated the setting of a Geometric(1/2) offspring distribution (which results in uniformly random planar trees) with IID displacements uniform on $\{-1, 0, 1\}$. Marckert and Mokkadem [83] treated the same offspring distribution, but where the displacements away from a vertex all have the same centered marginal distribution (but may depend on one another) with a $6+\varepsilon$ moment. Gittenberger [46] later generalised these results to critical, finite variance offspring distributions with centered (but not necessarily IID) displacements having finite $8+\varepsilon$ moment.

Work on the IID displacement case culminated in a paper of Janson and Marckert [60] which established the following result.

Theorem 2.1.6 ([60], Theorems 1 and 2)**.** *Let $\mu = (\mu_k)_{k \ge 0}$ be a critical offspring distribution with variance $\sigma^2 \in (0, \infty)$ such that μ has a finite exponential moment.*

For each $k \geq 1$, let ν_k be the law of a vector of k IID *copies of a random variable Y with $\mathbf{E}[Y] = 0$ and $\mathbf{E}[Y^2] = \beta^2 \in (0, \infty)$. Then*

$$\left(\frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{\widetilde{R}_n(2nt)}{n^{1/4}}\right)_{0 \leq t \leq 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0 \leq t \leq 1} \tag{2.1.6}$$

as $n \to \infty$, in the sense of finite-dimensional distributions. The convergence also holds in distribution in $\mathbf{C}([0,1], \mathbb{R}^2)$ endowed with the topology of uniform convergence if and only if

$$\mathbf{P}\{|Y| > y\} = o(y^{-4}) \text{ as } y \to \infty. \tag{2.1.7}$$

The finite exponential moment condition on the offspring distribution has subsequently been shown to be unnecessary, and may be weakened to a finite second moment assumption; see, for example, Marzouk [86]. Our Theorem 2.1.1 recovers this theorem under the additional assumption of a finite third moment for μ (and replacing convergence in the sense of finite-dimensional distributions in the first statement with random finite-dimensional distributions).

Janson and Marckert [60] also considered what happens in some of the "heavy-tailed" cases for which the tail condition $\mathbf{P}\{|Y| > y\} = o(y^{-4})$ fails. In particular, they considered the setting in which

$$\mathbf{P}\{Y \geq y\} \sim a_+ y^{-q}, \quad \mathbf{P}\{Y \leq -y\} \sim a_- y^{-q} \quad \text{as } y \to \infty$$

for some constants $a_+, a_- \geq 0$ and $q \in (2, 4]$, and prove analogues of Theorems 2.1.4 and 2.1.5 in such cases. They call the limiting object in this setting the *hairy tour*, and the associated snake the *jumping snake*. Their results were an important inspiration for Theorems 2.1.4 and 2.1.5.

Marzouk [86] later extended Janson and Marckert's results in [60] to the situation where the offspring distribution is in the domain of attraction of a stable law, and the displacements are IID.

Returning now to non-IID displacements, there are several notions of centering and finite variance which have been imposed in order to obtain convergence to the BSBE. Marckert and Miermont [81] worked under the "local centering" assumption that $\mathbf{E}[Y_{k,j}] = 0$ for all $1 \leq j \leq k$. For multi-type Bienaymé trees, [2] establishes convergence of discrete snakes under assumptions that impose in particular that the displacements away from vertices of each type are centered.

Most closely related to our results is a paper of Marckert [84], which proves the following theorem.

Theorem 2.1.7 ([84], Theorem 1). *Let $\mu = (\mu_k)_{k\geq 0}$ be a critical offspring distribution with $\mu_0 + \mu_1 < 1$ and with bounded support. Suppose further that $\nu = (\nu_k)_{k\geq 1}$ is such that*

$$\mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\right] = 0 \quad \text{and} \quad \beta^2 = \mathbf{E}\left[(Y_{\bar{\xi},U_{\bar{\xi}}})^2\right] < \infty,$$

and that there exists $p > 4$ such that

$$\sup_{1\leq j\leq k\leq K} \mathbf{E}\left[|Y_{k,j} - \mathbf{E}\left[Y_{k,j}\right]|^p\right] < \infty.$$

where μ is supported by $\{0, \ldots, K\}$. Then, as $n \to \infty$,

$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_n(nt)}{n^{1/4}}, \frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{\widetilde{R}_n(2nt)}{n^{1/4}}\right)_{0\leq t\leq 1} \xrightarrow{\text{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t, \frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\leq t\leq 1} \tag{2.1.8}$$

in $\mathbf{C}([0,1],\mathbb{R}^4)$ endowed with the topology of uniform convergence.

The boundedness condition is a necessary requirement of Marckert's proof technique, which is a *tour de force* involving tracking very detailed information about the number of vertices of each possible different degree along a lineage, which converge on appropriate rescaling to a Gaussian field. Our approach removes the boundedness requirement, but we do not obtain such fine information on the limit object.

Finally, we mention a forthcoming work of Duquesne and Rebei [38], which proves limit theorems for snakes whose jumps are centered and sibling-independent and such that the underlying family tree converges to a Lévy tree. Our understanding is that the results and technique of [38] are rather different from those of the current work.

2.1.4 A first application

One nice consequence of Theorem 2.1.1 is a strengthening of a result of Marckert and Mokkadem [82] concerning the difference between the height process, H_n, and the Łukasiewicz path, here denoted by W_n (and formally defined in Section 2.2) of T_n. It is proved in [82] that if ξ is critical with variance $\sigma^2 \in (0, \infty)$ and has a finite exponential moment, then

$$\left(\frac{\widetilde{H}_n(2nt)}{\sqrt{n}}, \frac{H_n(nt)}{\sqrt{n}}, \frac{W_n(nt)}{\sqrt{n}}\right)_{0\leq t\leq 1} \xrightarrow{\text{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \frac{2}{\sigma}\mathbf{e}_t, \sigma\mathbf{e}_t\right)_{0\leq t\leq 1}$$

as $n \to \infty$ in $\mathbf{C}([0,1],\mathbb{R}^3)$. (As mentioned after Theorem 2.1.6, the finite exponential moment condition is unnecessary and may be removed; see Duquesne [39] for this result in the context of trees rather than snakes.)

Moreover, under the same assumptions, [82] establishes that, for any $\varepsilon > 0$, there exists $\gamma > 0$ such that for $n > 0$ sufficiently large

$$\mathbf{P}\left\{\sup_{0\le i\le n}\left|\sigma H_n(i) - 2\sigma^{-1}W_n(i)\right| \ge n^{1/4+\varepsilon}\right\} \le \exp\left(-\gamma n^{\varepsilon}\right).$$

It is natural to conjecture that, under suitable conditions, the difference varies precisely on the order of $n^{1/4}$. We are able to prove this conjecture in a large degree of generality. It turns out that the difference $(\sigma H_n(i) - 2\sigma^{-1}W_n(i), 0 \le i \le n)$ evolves precisely as the head of a discrete snake (see Lemma 2.2.1 for a proof of this fact). The relevant displacements are given by $Y_{k,j} = \sigma - (2/\sigma)(k-j)$; this formula already appeared at (2.1.5). We have

$$\sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\mathbf{E}\left[Y_{k,j}\right] = \sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\left(\sigma - \frac{2}{\sigma}(k-j)\right) = \sum_{k=1}^{\infty}\mu_k\left(\sigma k - \frac{k(k-1)}{\sigma}\right) = \sigma - \frac{\sigma^2}{\sigma} = 0,$$

so that the associated discrete snake is globally centered. Moreover, the global variance is

$$\begin{aligned}\sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\mathbf{E}\left[Y_{k,j}^2\right] &= \sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\left(\sigma - \frac{2}{\sigma}(k-j)\right)^2 \\ &= \sum_{k=1}^{\infty}\mu_k\left(\sigma^2 k - 2k(k-1) + \frac{2}{3\sigma^2}k(k-1)(2k-1)\right) \\ &= \frac{4}{3\sigma^2}(\mathbf{E}\left[\xi^3\right] - 1) - (\sigma^2 + 2),\end{aligned}$$

which is finite provided that $\mathbf{E}\left[\xi^3\right] < \infty$. Also,

$$\mathbf{P}\left\{\max_{1\le i\le \xi}|Y_{\xi,i}| > y\right\} = \mathbf{P}\left\{\left|\sigma - \frac{2}{\sigma}(\xi-1)\right| \vee \sigma > y\right\} = o(y^{-4})$$

as $y \to \infty$ if and only if $\mathbf{P}\{\xi > y\} = o(y^{-4})$ as $y \to \infty$; moreover, the latter condition implies $\mathbf{E}\left[\xi^3\right] < \infty$. We obtain the following corollary of Theorem 2.1.1.

Corollary 2.1.8. *Let $\mu = (\mu_k)_{k\ge1}$ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$. Let $\beta^2 = \frac{4}{3\sigma^2}(\mathbf{E}\left[\xi^3\right] - 1) - (\sigma^2+2)$. Then*

$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{\sigma H_n(nt) - 2\sigma^{-1}W_n(nt)}{n^{1/4}}\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\le t\le 1},$$

as $n \to \infty$ in $\mathbf{C}([0,1],\mathbb{R}^2)$ if and only if $\mathbf{P}\{\xi > y\} = o(y^{-4})$ as $y \to \infty$.

Let us observe that, while Corollary 2.1.8 concerns the difference between the height process and the Łukasiewicz path, the joint convergence in Theorem 2.1.1 can be used to prove an analogous result for the difference between the Łukasiewicz path and the contour process encoding of the head of the same discrete snake. (We leave the details of this statement to the reader.)

In the case where ξ is bounded, Marckert's result (Theorem 2.1.7) applies, so the corollary is new only in the case of unbounded offspring distributions. In an earlier paper [80], Marckert had already observed that the difference between the left and right pathlengths (also known as the *imbalance*) of a size-conditioned Bienaymé tree with offspring distribution $\mu_0 = \mu_2 = 1/2$ converges in distribution after rescaling to $2^{1/4}S$, where $S = \int_0^1 \mathbf{r}_t dt$. We note that such trees are binary, and recall that the left pathlength (resp. right pathlength) of a vertex v is the number of vertices in its ancestral lineage who precede (resp. succeed) their siblings in the lexicographical order. The left (resp. right) pathlength of binary trees is then the sum of the left (resp. right) path lengths over all vertices in the tree. Janson [57] later used the method of moments to give an alternate proof of this convergence in distribution.

It can also be the case that the sequence $(n^{-1/4} \max_{0\le i\le n} |\sigma H_n(i) - 2\sigma^{-1} W_n(i)|)_{n\ge 1}$ is tight *without* converging in distribution to the maximum modulus of the head of the BSBE; indeed, by Theorem 2.1.4, if $r^4 \mathbf{P}\{\xi \in rA\} \to \pi_1(A)$ as $r \to \infty$ for all Borel sets A such that $\pi_1(\partial A) = 0$ and a Borel measure π_1 on $\mathbb{R}_+ \setminus \{0\}$ such that for any $\varepsilon > 0$, $\pi_1((\varepsilon, \infty)) < \infty$, then it is possible to prove that $n^{-1/4} \max_{0\le i\le n} |\sigma H_n(i) - 2\sigma^{-1} W_n(i)|$ converges in distribution to the maximum modulus of the appropriate hairy tour. If, on the other hand, we have $r^{4-\eta} \mathbf{P}\{\xi \in rA\} \to \pi_1(A)$ as $r \to \infty$ for some $\eta \in (0, 2)$, then Theorem 2.1.5 yields the convergence

$$n^{-1/(4-\eta)} \max_{0\le i\le n} |\sigma H_n(i) - 2\sigma^{-1} W_n(i)| \xrightarrow{\mathrm{d}} L,$$

where $\mathbf{P}\{L \le \ell\} = \exp(-\int_\ell^\infty \pi_1(x) dx)$ is the probability that no point of a Poisson point process of intensity $dt\pi_1(dx)$ on $[0, 1] \times \mathbb{R}_+$ has second co-ordinate greater than ℓ.

2.1.5 A second application

A second consequence of Theorem 2.1.1 concerns the difference between the height process of T_n and the height process of the corresponding *looptree*. The looptree corresponding to T_n, denoted by T_n°, is the connected multigraph obtained by replacing the edges from a vertex to its children by a cycle going through the parent and all of its children in order (whose length, therefore, equals its number of children plus one). See Figure 2.2 for an illustration. (It turns out that it is possible to make sense of a continuum analogue of this notion, as proved by Curien and Kortchemski [36].)

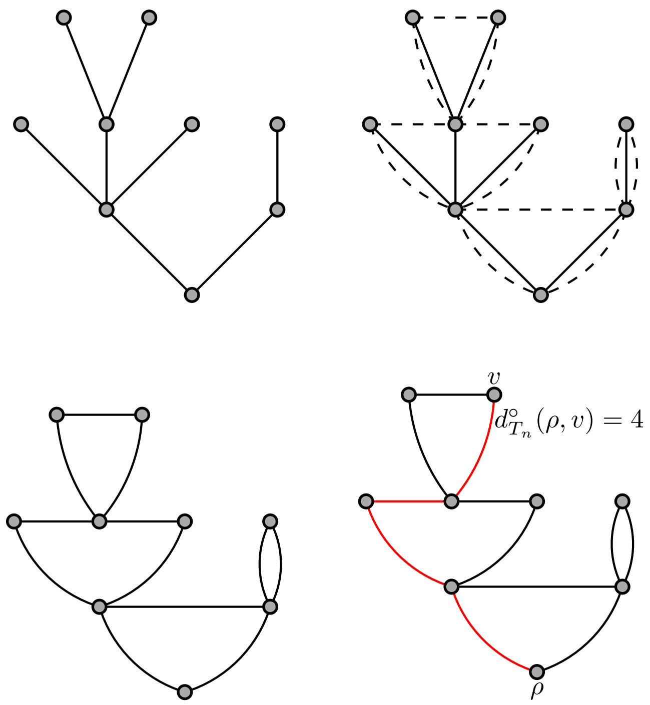

Figure 2.2: In the top left figure, a tree, and in the bottom left figure its corresponding looptree. The top-right figure serves to aid in understanding the construction, and the bottom-right figure illustrates how distances are calculated in the loop-tree.

Vertices in the original tree naturally correspond to vertices in the looptree. Let $v_1, \ldots, v_n$ be the vertices of T_n listed in lexicographical order. We define the height function of the looptree, denoted $H_n^\circ : [0, n] \to \mathbb{R}$, to give the graph distance between the root and each of the vertices in the looptree, visited in the order $v_1, \ldots, v_n$. This is the height process (in the usual sense) of the spanning tree of the looptree made up of the union of the geodesic paths from each of its vertices to the root. Formally, using the Ulam–Harris notation (see Section 2.2 for details) for $0 \leq i \leq n-1$ let

$$H_n^\circ(i) := \sum_{(u,uj) \in e(\mathrm{T}_n)\ :\ uj \preceq v_{i+1}} \min\{j, c(u, \mathrm{T}_n) + 1 - j\},$$

where for $u \in v(\mathrm{T}_n)$, $c(u, \mathrm{T}_n)$ denotes the number of children of u in T_n. Finally let $H_n^\circ(n) = 0$, and extend the domain to $[0, n]$ by linear interpolation. For $c \in \mathbb{R}$, it is readily seen that the difference $(cH_n(i) - H_n^\circ(i), 0 \leq i \leq n)$ evolves as the head of a discrete snake whose displacements are given by

$$Y_{k,j} = c - \min\{j, k + 1 - j\}.$$

Moreover, if we fix $c = \frac{1}{4}\mathbf{E}\left[\xi^3\right] + \frac{1}{2} + \frac{1}{4}\mathbf{P}\left\{\xi \in 2\mathbb{Z}+1\right\}$, then

$$\begin{aligned}
\sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\mathbf{E}\left[Y_{k,j}\right] &= \sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\left(c-\min\{j,k+1-j\}\right)\\
&= \sum_{k=1}^{\infty}\mu_k\left(ck-2\sum_{i=1}^{\lfloor k/2\rfloor} i - \left\lceil\frac{k}{2}\right\rceil \mathbf{1}_{[k\in 2\mathbb{Z}+1]}\right)\\
&= \sum_{k=1}^{\infty}\mu_k\left(ck-\frac{k}{2}\left(\frac{k}{2}+1\right)-\frac{1}{4}\mathbf{1}_{[k\in 2\mathbb{Z}+1]}\right)\\
&= c-\frac{1}{4}\mathbf{E}\left[\xi^3\right]-\frac{1}{2}-\frac{1}{4}\mathbf{P}\left\{\xi\in 2\mathbb{Z}+1\right\}\\
&= 0,
\end{aligned}$$

so that the associated discrete snake is globally centered. Moreover, the global variance is

$$\begin{aligned}
&\sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\mathbf{E}\left[Y_{k,j}\right] \qquad (2.1.9)\\
&\quad= \sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\left(c-\min\{j,k+1-j\}\right)^2\\
&\quad= \sum_{k=1}^{\infty}\mu_k\left(c^2k-ck\left(\frac{k}{2}+1\right)-\frac{c}{2}\mathbf{1}_{[k\in 2\mathbb{Z}+1]}+2\sum_{i=1}^{\lfloor k/2\rfloor}i^2+\left\lceil\frac{k}{2}\right\rceil^2\mathbf{1}_{[k\in 2\mathbb{Z}+1]}\right)\\
&\quad= \sum_{k=1}^{\infty}\mu_k\left(c^2k+k\left(\frac{k}{2}+1\right)\left(\frac{k}{6}+\frac{1}{6}-c\right)+\left(\frac{k^2}{12}+\frac{k}{3}+\frac{1}{4}-\frac{c}{2}\right)\mathbf{1}_{[k\in 2\mathbb{Z}+1]}\right)\\
&\quad= c^2+\frac{\mathbf{E}\left[\xi^3\right]}{12}+\left(\frac{1}{4}-\frac{c}{2}\right)\mathbf{E}\left[\xi^2\right]+\frac{1}{6}-c+\mathbf{E}\left[\left(\frac{\xi^2}{12}+\frac{\xi}{3}+\frac{1}{4}-\frac{c}{2}\right)\mathbf{1}_{[\xi\in 2\mathbb{Z}+1]}\right]\\
&=: \beta^2, \qquad (2.1.10)
\end{aligned}$$

which is finite provided $\mathbf{E}\left[\xi^3\right] < \infty$.

Finally,

$$\mathbf{P}\left\{\max_{1\le i\le \xi}|Y_{\xi,i}|>y\right\} = \mathbf{P}\left\{|c-\lceil \xi/2\rceil| \vee |c-1| > y\right\} = o(y^{-4})$$

as $y\to\infty$ if and only if $\mathbf{P}\left\{\xi>y\right\} = o(y^{-4})$ as $y\to\infty$ and, moreover, the latter condition implies $\mathbf{E}\left[\xi^3\right]<\infty$. We obtain the following corollary of Theorem 2.1.1.

Corollary 2.1.9. *Let $\mu = (\mu_k)_{k\ge 1}$ be a critical offspring distribution with variance $\sigma^2\in(0,\infty)$, and let ξ be a random variable with distribution μ. Let $c = \frac{1}{4}\mathbf{E}\left[\xi^2\right] + \frac{1}{2} + \frac{1}{4}\mathbf{P}\left\{\xi\in 2\mathbb{Z}+1\right\}$ and β^2 be as in (2.1.9). Then,*

$$\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{H_n^{\circ}(nt)}{c\sqrt{n}}, \frac{cH_n(nt)-H_n^{\circ}(nt)}{n^{1/4}}\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_t, \frac{2}{\sigma}\mathbf{e}_t, \sqrt{\frac{2}{\sigma}}\beta\mathbf{r}_t\right)_{0\le t\le 1}, \qquad (2.1.11)$$

as $n \to \infty$ in $\mathbf{C}([0,1],\mathbb{R}^3)$ endowed with the topology of uniform convergence if and only if $\mathbf{P}\{\xi > y\} = o(y^{-4})$ as $y \to \infty$.

Analogues of this result also hold in the settings of Theorems 2.1.4 and 2.1.5. Even the functional convergence for the height process of looptrees of Bienaymé trees, expressed in the second coordinate of (2.1.11), is new, although pointwise convergence was proved by Kortchemski and Marzouk [71]. (The convergence of looptrees in the Gromov–Hausdorff topology was proved in [72] via spinal decomposition – see Theorem 1.2 and the generic case in Corollary 1.4 for the application to maps – and convergence in the Gromov–Hausdorff–Prokhorov topology was shown in [65, Theorem 15].)

2.1.6 Overview of the proofs

We will prove weak convergence of the head of the snake by making use of the following variant of the usual formulation of weak convergence for a sequence of random continuous functions. (This formulation is inspired by Theorem 20 of [9], and can be proved by essentially the same method as the second proof of Theorem 7.5 of [18].)

Proposition 2.1.10. *Let $(f_n)_{n\geq 1}$ and f be random elements of $\mathbf{C}([0,1],\mathbb{R})$ such that $f_n(0) = f_n(1) = 0$ for every $n \geq 1$ and $f(0) = f(1) = 0$. Let $U_1, U_2, \ldots$ be IID $U[0,1]$ random variables, independent of $(f_n)_{n\geq 1}$ and f. For $k \geq 1$, write $U^k_{(1)}, U^k_{(2)}, \ldots, U^k_{(k)}$ for the values of $U_1, U_2, \ldots, U_k$ written in increasing order, and set $U^k_{(0)} = 0$ and $U^k_{(k+1)} = 1$.*

Suppose that for each $k \geq 1$ we have

$$\left(f_n(U^k_{(1)}), \ldots, f_n(U^k_{(k)})\right) \xrightarrow{\mathrm{d}} \left(f(U^k_{(1)}), \ldots, f(U^k_{(k)})\right), \tag{2.1.12}$$

as $n \to \infty$, and that for any $\varepsilon > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\leq i\leq k} \sup_{s,t\in[U^k_{(i)}, U^k_{(i+1)}]} |f_n(s) - f_n(t)| > \varepsilon \right\} = 0. \tag{2.1.13}$$

Then $f_n \xrightarrow{\mathrm{d}} f$ as $n \to \infty$, for the topology generated by the uniform norm on $\mathbf{C}([0,1],\mathbb{R})$.

In the sequel we will refer to assumption (2.1.12) as the *convergence of random finite-dimensional distributions* and to (2.1.13) as *tightness*. Observe that (2.1.12) is weaker than the usual convergence of finite-dimensional distributions. However, it is more natural in the context of random trees, and indeed plays a key role in Aldous' theory of continuum random trees as developed in [9]. (See the appendix of [11] for a discussion and for further references.)

Let $\mathcal{T}_{2\mathbf{e}}$ be the real tree encoded by $2\mathbf{e}$, where $\mathbf{e}$ is a normalised Brownian excursion. (We refer to the survey of Le Gall [42] for standard definitions concerning random real trees.) Fix $k \geq 1$ and let $U_1, \ldots, U_k$ be IID Uniform$([0,1])$ random variables. Furthermore, let $\mathcal{T}_{2\mathbf{e}}^k$ be the subtree of $\mathcal{T}_{2\mathbf{e}}$ spanned by the images of 0 and of $U_1, \ldots, U_k$ in $\mathcal{T}_{2\mathbf{e}}$. Formally, it is useful to think of this as an ordered rooted tree with leaves labeled by $1, 2, \ldots, k$ and edge-lengths, where we use the relative ordering of $U_1, U_2, \ldots, U_k$ to determine the planar ordering of the leaves. Using Aldous' line-breaking construction [9], we may construct a tree which is equal in distribution to $\mathcal{T}_{2\mathbf{e}}^k$ as follows.

Let $J_1, \ldots, J_k$ be the first k jump times of a Poisson point process on $[0, \infty)$ with intensity tdt at time t. For $i = 1, \ldots, k-1$, sample an *attachment point* $A_i \sim$ Uniform$([0, J_i])$, independent of $(A_j)_{j \neq i}$. Take the completion of each of the line segments $[0, J_1], (J_1, J_2], \ldots,$ $(J_{k-1}, J_k]$, and for each $i \in \{1, \ldots, k-1\}$ let J_i^* denote the limit point as $x \downarrow J_{i-1}$. Identify the points J_i^* and A_i, and think of the line-segment as being attached to the left side of the branch containing A_i with probability 1/2 and to the right side with probability 1/2. Denote the resulting rooted ordered tree with leaf-labels and edge-lengths by $\mathcal{T}^k$. Then, $\mathcal{T}_{2\mathbf{e}}^k \stackrel{\mathrm{d}}{=} \mathcal{T}^k$; see [9, p. 279].

The proof of Theorem 2.1.1 (and similarly Theorems 2.1.4 and 2.1.5) relies on proving that a certain discrete line-breaking construction of T_n, described formally in Section 2.2.2, converges to Aldous' line-breaking construction upon rescaling. The discrete construction builds a tree on $[n]$ by first constructing paths $P^{(1)}, \ldots, P^{(\ell^*)}$ and then attaching them to one another by identifying one endpoint of each path $P^{(i)}$ with a point in $(P^{(j)})_{j<i}$. The proof of convergence of the random finite-dimensional distributions relies on the observation that, along each path, the sequence of partial sums of the displacements is essentially a random walk trajectory with IID steps with the same distribution as $Y_{\bar{\xi}, U_{\bar{\xi}}}$ and, moreover, that random displacements appearing at branch points do not contribute to the displacements of the discrete snake on the "macroscopic" spatial scale of $\Theta(n^{1/4})$.

For the proofs of tightness, we adapt a method of Haas and Miermont [50] used to prove tightness for the height process of a Markov branching tree. (Note that size-conditioned Bienaymé trees are examples of Markov branching trees.) Let T_n^k be a subtree of T_n spanned by its root and k uniform vertices. The difference $\mathrm{T}_n \setminus \mathrm{T}_n^k$ is a forest F_n^k, and to prove tightness we bound the maximum modulus of the spatial locations in each tree in F_n^k. Following Haas and Miermont, we reduce this bound to an expression involving only a size-biased pick among the trees in F_n^k. The proof of tightness then reduces to proving an explicit tail bound for the maximum modulus of the spatial location of a vertex in T_n when rescaled by $n^{-1/4}$. As a key part of our argument, we require a strong control on the total variation distance between the laws of $\bar{\xi}$ and of the number of children of the root of T_n, which we denote by $\widehat{D}_1^n$. For $k \in [n]$, by Kemperman's formula [96, Chapter 6],

$$\mathbf{P}\left\{\widehat{D}_1^n = k\right\} = \left(\frac{n}{n-1}\right) \frac{\mathbf{P}\left\{S_{n-1} = n-1-k\right\}}{\mathbf{P}\left\{S_n = n-1\right\}} \mathbf{P}\left\{\bar{\xi} = k\right\}, \tag{2.1.14}$$

where $(S_n)_{n\geq 1}$ is a random walk with IID μ-distributed increments. In order to control this total variation distance, we use a version of the local central limit theorem ([94, Theorem 13, Chapter VII] which, for completeness, we also state below in Theorem 2.8.2) which holds whenever $\mathbf{E}[\xi^3] < \infty$; this is the origin of the third moment condition in our main theorem.

2.1.7 Asymptotic notation

We will use the following notation related to the asymptotics of random variables $(X_n)_{n\geq 1} \in \mathbb{R}$. (See Janson [58].) For $(y_n)_{n\geq 1} \in \mathbb{R}_{>0}$,

- $X_n = o_{\mathbf{P}}(y_n)$ means that $X_n/y_n \xrightarrow{\text{p}} 0$ as $n \to \infty$;
- $X_n = \omega_{\mathbf{P}}(y_n)$ means that $X_n/y_n \xrightarrow{\text{p}} \infty$ as $n \to \infty$;
- $X_n = O_{\mathbf{P}}(y_n)$ means that for all $\varepsilon > 0$, there exist constants n_ε, $C_\varepsilon > 0$ such that for all $n \geq n_\varepsilon$,
$$\mathbf{P}\left\{X_n \leq C_\varepsilon y_n\right\} \geq 1 - \varepsilon;$$
- $X_n = \Omega_{\mathbf{P}}(y_n)$ means that for all $\varepsilon > 0$ there exist constants n_ε, $C_\varepsilon > 0$ such that for all $n \geq n_\varepsilon$,
$$\mathbf{P}\left\{X_n \geq C_\varepsilon y_n\right\} \geq 1 - \varepsilon;$$
- $X_n = \Theta_{\mathbf{P}}(y_n)$ means that $X_n = O_{\mathbf{P}}(y_n)$ and $X_n = \Omega_{\mathbf{P}}(y_n)$.
- Lastly, "with high probability" always means "with probability tending to 1 as $n \to \infty$.

2.2 Trees, branching random walks, and their encodings

We require a number of different tree models, which we now define.

First, a *tree* is simply a connected acyclic graph $T = (v(T), e(T))$. A *rooted tree* consists of a tree together with a distinguished root vertex $\rho = \rho(T) \in v(T)$. Given a rooted tree T and a vertex v of T write $C(v, T)$ for the set of children of v in T and $c(v, T) = |C(v, T)|$; vertex v is a *leaf* of T if $c(v, T) = 0$. We write ∂T for the set of leaves of T. Also, for a non-root vertex v we write $p(v) = p(v, T)$ for the parent of v in T. For vertices $v, w \in v(T)$ we write $v \prec w$ if v is an ancestor of w, and for an edge e we also write $e \prec v$ if at least one endpoint of e is an ancestor of v. For $S \subset v(T)$, the *subtree of T spanned by S* is the minimal subtree of T containing all elements of S.

Letting $\mathbb{N}^0 := \{\emptyset\}$, the *Ulam–Harris tree* is the rooted tree with root $\emptyset$ and vertex set

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

in which, for each $v \in \mathcal{U}$, the set of children of v is $\{vi, i \in \mathbb{N}\}$. (Here, and in the sequel, for a string $v = (v_1, \ldots, v_k)$ we write $vi := (v_1, \ldots, v_k, i)$.) We say w is a younger sibling of u if $w = vj$, $u = vi$ and $j > i$. We will make use of the usual lexicographic order on $\mathcal{U}$, which is the total order in which each vertex precedes all of its descendants and all of its younger siblings. Also, for $v \in \mathbb{N}^n \subset \mathcal{U}$ we write $|v| = n$ for the depth of v in $\mathcal{U}$.

The definitions of the coming paragraph are illustrated in Figure 2.3. An *ordered rooted tree* is a tree T with $v(T) \subset \mathcal{U}$ and the following properties: (i) $\emptyset \in v(T)$; (ii) if $v \in v(T)$ then $p(v, \mathcal{U}) \in v(T)$; (iii) if $vi \in v(T)$ then $vj \in v(T)$ for all $1 \leq j \leq i$. Note that the edge set of an ordered rooted tree may be recovered from its vertex set, and we will often identify ordered rooted trees with their vertex sets. The lexicographic order on $v(T)$ is simply the restriction of the lexicographic order on $\mathcal{U}$ to $v(T)$.

A *labeled ordered rooted tree* is a finite rooted tree $T = (v(T), e(T))$ with $v(T) = [n]$ in which, for each non-leaf vertex of T, the set of children is endowed with a total order $\sigma_v = \sigma_{v,T} : C(v, T) \to [c(v, T)]$. We will sometimes abuse notation by writing $vi = \sigma_v^{-1}(i)$ for the i-th child of v under this total order. This abuse of notation is justified by the observation that the ordering of the children of each non-leaf induces an injection $\varphi : v(T) \to \mathcal{U}$ defined inductively by $\varphi(\rho(T)) = \emptyset$ and $\varphi(vi) = \varphi(\sigma_v^{-1}(i)) = \varphi(v)i$ for $i \in [c(v, T)]$; and $\varphi(v(T))$ is indeed (the vertex set of) an ordered rooted tree. As such, a labeled ordered rooted tree could equivalently be represented as a pair (T, f) where $T \subset \mathcal{U}$ is a finite ordered rooted tree and

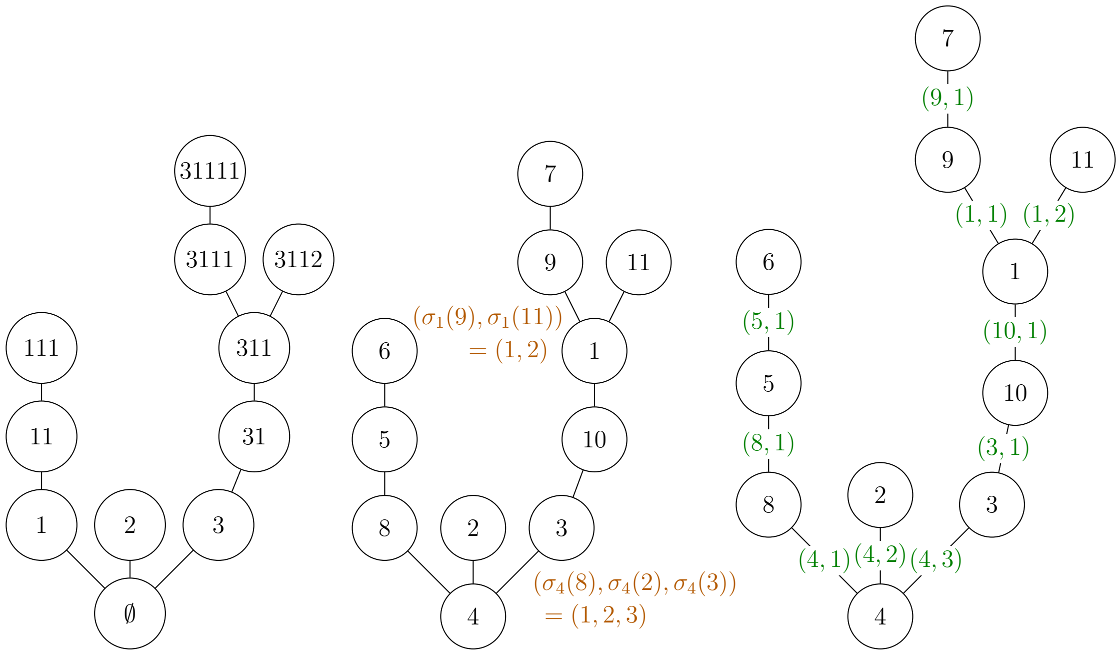

Figure 2.3: Left: an ordered rooted tree. Center: a labeled ordered rooted tree, with the functions σ_v indicated for $v \in \{1, 4\}$. Right: the edge labeling of T, introduced in Section 2.2.2.

$f : T \to [n]$ is a bijection (so $n = |T|$). However, the first representation is more natural in the context of the methods we shall shortly use for constructing random labeled ordered rooted trees. Moreover, the second representation would be confusing, as it is very similar to our representations of *branching random walks* and of *spatial trees*, which we now describe.

2.2.1 Branching random walks, Łukasiewicz path, contour and height processes

A *branching random walk* is a pair $\mathrm{T} = (T, Y)$, where T is an ordered rooted tree (possibly labeled) and $Y = (Y^{(v)}, v \in v(T) \setminus \partial T)$, where $Y^{(v)} = (Y_j^{(v)}, j \in [c(v, T)]) \in \mathbb{R}^{c(v,T)}$. We think of $Y^{(v)}$ as a set of spatial displacements from vertex v to its children, so $Y_j^{(v)}$ is the difference in the spatial locations of vertices v and vj. The spatial *location* of $u \in v(T)$ is then given by the sum of displacements along u's ancestral path:

$$\ell(u) = \ell(u, \mathrm{T}) := \sum_{\{(v,vj) \in e(T) : vj \preceq u\}} Y_j^{(v)}.$$

We refer to the pair (T, ℓ) as a *spatial tree*. The branching random walk (T, Y) can clearly be recovered from the spatial tree (T, ℓ), and vice versa.

Let $T = (v(T), e(T))$ be a finite ordered rooted tree and write $n = |T|$. The *Lukasiewicz path* of T is the function $W_T : [0, n] \to \mathbb{R}$ defined as follows. List the elements of $v(T)$ in lexicographic order as $v_1, \ldots, v_n$. Set $W_T(0) = 0$. For $1 \le i \le n$, set $W_T(i) = \sum_{j=1}^{i}(c(v_i, T) - 1)$, and then extend the domain of W_T to $[0, n]$ by linear interpolation.

The *height process* of T is the function $H_T : [0, n] \to \mathbb{R}_{\ge 0}$ defined as follows. For $0 \le i < n$ set $H_T(i) = |v_{i+1}|$ and set $H_T(n) = 0$; then extend the domain of H_T to $[0, n]$ by linear interpolation.

The *contour order* of $v(T)$ is the sequence $w_0, \ldots, w_{2(n-1)}$ of elements of $v(T)$ defined as follows. First, $w_0 = \emptyset$ is the root of T. Inductively, for each $0 \le i < 2(n-1)$, if w_i has at least one child in T which does not appear in the sequence $w_0, \ldots, w_{i-1}$, then let w_{i+1} be the lexicographically least such child. Otherwise, let $w_{i+1} = p(w_i, T)$. It is straightforward to verify that each vertex v of T appears in the resulting sequence exactly $1 + c(v, T)$ times. The *contour process* of T is the function $\widetilde{H}_T : [0, 2(n-1)] \to \mathbb{R}_{\ge 0}$ defined by setting $\widetilde{H}_T(i) = |w_i|$ for integers i with $0 \le i \le 2(n-1)$, letting $\widetilde{H}_T(2n) = 0$, and extending to $[0, 2n]$ by linear interpolation.

If $\mathrm{T} = (T, Y)$ is a branching random walk with underlying tree T then we encode the spatial locations by a function $R_{\mathrm{T}} : [0, n] \to \mathbb{R}$ given by setting $R_{\mathrm{T}}(i) = \ell(v_{i+1}, \mathrm{T})$, for $i \in \{0, \ldots, n-1\}$, $R_{\mathrm{T}}(n) = 0$, and extending to $[0, n]$ by linear interpolation. We also define a process $\widetilde{R}_{\mathrm{T}} : [0, 2n] \to \mathbb{R}$ by setting $\widetilde{R}_{\mathrm{T}}(i) = \ell(w_i, \mathrm{T})$ for integers i with $0 \le i \le 2(n-1)$, $\widetilde{R}_{\mathrm{T}}(2n) = 0$, and extending to $[0, 2n]$ by linear interpolation.

The following result appears somewhat implicitly in Section 3 of [23]. For completeness we give a proof.

Lemma 2.2.1. *Fix $\alpha_1, \alpha_2 \neq 0$. Let $\mathrm{T} = (T, Y)$ be the branching random walk with $Y = (Y^{(v)}, v \in v(T) \setminus \partial T)$ such that $Y^{(v)} = (\alpha_1 - \frac{2}{\alpha_2}(c(v, T) - j), j \in [c(v, T)])$, $v \in v(T)$. Let R_{T} be the function encoding the spatial locations of* T*. Then for all $t \in [0, n]$,*

$$R_{\mathrm{T}}(t) = \alpha_1 H_T(t) - \frac{2}{\alpha_2} W_T(t).$$

Proof. It is sufficient to prove that

$$R_{\mathrm{T}}(i) = \alpha_1 H_T(i) - \frac{2}{\alpha_2} W_T(i)$$

for $i \in \{0, 1, \ldots, n-1\}$. Let $i \in \{0, \ldots, n-1\}$. Then $H_T(i)$ is the number of ancestors of v_{i+1} in T. Further, $W_T(i)$ is the number of younger siblings of ancestors of v_{i+1}. It follows that

$$\begin{aligned}\alpha_1 H_T(i) - \frac{2}{\alpha_2} W_T(i) &= \alpha_1 \cdot \left(\sum_{(u,uj)\in e(T):uj\preceq v_{i+1}} 1 \right) - \frac{2}{\alpha_2} \sum_{(u,uj)\in e(T):uj\preceq v_{i+1}} (c(u,T) - j) \\ &= \sum_{(u,uj)\in e(T):uj\preceq v_{i+1}} \left(\alpha_1 - \frac{2}{\alpha_2} (c(u,T) - j) \right) \\ &= \sum_{(u,uj)\in e(T):uj\preceq v_{i+1}} Y_j^{(u)} \\ &= R_{\mathrm{T}}(i),\end{aligned}$$

and the result follows. □

2.2.2 Sequential encodings of labeled ordered rooted trees

Given a labeled ordered rooted tree $T = ([n], e(T))$, we assign labels to the *edges* of T as follows. For $v \in [n]$ and $i \in [c(v,T)]$, assign label (v,i) to the edge $\{v, vi\} = \{v, \sigma_v^{-1}(i)\}$. The set of all edge labels is then $L(T) = \{(v,i) : v \in v(T), i \in [c(v,T)]\}$. Given any path $P = v_0 v_1 \dots v_k$ from a vertex v_0 of T to one of its descendants, let π_P be the sequence of edge labels along the path from v_0 to v_k: formally, $\pi_P = \pi_P(T) = ((v_0, c_0), \dots, (v_{k-1}, c_{k-1}))$, where $c_0, \dots, c_{k-1}$ are such that $v_j = v_{j-1} c_{j-1}$ for each $j \in [k]$.

We say a sequence $\mathrm{d} = (d_1, \dots, d_n)$ of non-negative integers is a *degree sequence* if $\sum_{v\in[n]} d_v = n - 1$. We say a labeled tree T with $v(T) = [n]$ has degree sequence d if $c(v,T) = d_v$ for all $v \in [n]$. Write $\mathcal{L}_{\mathrm{d}}$ for the set of labeled ordered rooted trees with degree sequence d. For any tree $T \in \mathcal{L}_{\mathrm{d}}$, it is the case that $L(T) = \{(v,c) : v \in [n], c \in [d_v]\}$. Write $\mathcal{P}_{\mathrm{d}}$ for the set of permutations of $\{(v,c) : v \in [n], c \in [d_v]\}$; this set has size $(n-1)!$. For a fixed degree sequence d, we will make extensive use of a bijection $B : \mathcal{P}_{\mathrm{d}} \to \mathcal{L}_{\mathrm{d}}$ for $\mathrm{d} = (d_1, \dots, d_n)$ a degree sequence of length $n \geq 2$, which we give below. We first describe B^{-1}, as it is slightly simpler.

The bijection $B^{-1} : \mathcal{L}_{\mathrm{d}} \to \mathcal{P}_{\mathrm{d}}$. **Input**: $T \in \mathcal{L}_{\mathrm{d}}$.

- Let $T^{(0)}$ be the subtree of T consisting of the root alone.
- For $\ell \geq 1$, if $T^{(\ell-1)} \neq T$ then let $y^{(\ell)}$ be the smallest label of a vertex in T which is not in $T^{(\ell-1)}$, let $P^{(\ell)}$ be the path in T from $T^{(\ell-1)}$ to $y^{(\ell)}$, and let $T^{(\ell)}$ be the subtree of T spanned by $\{P^{(\ell)}, y^{(1)}, \dots, y^{(\ell-1)}\}$.
- Let ℓ^* be the first value for which $T^{(\ell^*)} = T$.
- Let π_T be the concatenation of the sequences $\pi_{P^{(1)}}, \dots, \pi_{P^{(\ell^*)}}$, and set $B^{-1}(T) = \pi_T$.

In the example of Figure 2.3, $\ell^* = 6$ and the paths are $P^{(1)} = 4, 3, 10$, $P^{(2)} = 4$, $P^{(3)} = 4, 8$, $P^{(4)} = 5$, $P^{(5)} = 1, 9$ and $P^{(6)} = 1$, so

$$\pi_T = ((4,3),(3,1),(10,1),(4,2),(4,1),(8,1),(5,1),(1,1),(9,1),(1,2))\,. \qquad (2.2.1)$$

We next describe B; for this we make use of the fact that to specify a labeled ordered rooted tree T with vertex set $[n]$ it suffices to specify the set $C(v,t)$ and the total orderings $\sigma_v : C(v,T) \to [c(v,T)]$ for each $v \in [n]$.

Informally, this construction can be thought of as a discrete analog of the continuous line-breaking construction from the second paragraph of Section 2.1.6. More specifically, given $\pi = ((v_1, c_1), \ldots, (v_{n-1}, c_{n-1})) \in \mathcal{P}_{\mathrm{d}}$, certain substrings of π will correspond to paths in the tree $B(\pi)$. We will list these paths as $P^{(1)}, \ldots, P^{(\ell^*)}$. As in the continuous line-breaking construction, for each $i \geq 2$ we will identify one endpoint of the path $P^{(i)}$, with a vertex in $(P^{(j)})_{j<i}$. In the following formal description we denote the i-th identified vertex by v_{j_i}. When we identify the endpoint of path $P^{(i)}$ with vertex v_{j_i}, we use the second coordinate of the pair (v_{j_i}, c_{j_i}) to determine the position of the unique child of v_{j_i} belonging to $P^{(i)}$ among the children of v_{j_i}.

The bijection $B : \mathcal{P}_{\mathrm{d}} \to \mathcal{L}_{\mathrm{d}}$. Input: $\pi = ((v_1, c_1), \ldots, (v_{n-1}, c_{n-1})) \in \mathcal{P}_{\mathrm{d}}$.

- Set $m_1 = \min\{m \in \mathbb{N} : m \neq v_1\}$ and let

$$j_1 = \inf\{j > 1 : v_j \in \{m_1, v_1, \ldots, v_{j-1}\}\} \wedge n\,.$$

- For $i \geq 1$, if $j_i < n$ then:
  - set $m_{i+1} = \min\{m > m_i : m \notin \{v_1, \ldots, v_{j_i}\}\}$;
  - let

$$j_{i+1} = \inf\{j > j_i : v_j \in \{m_1, \ldots, m_{i+1}, v_1, \ldots, v_{j-1}\}\} \wedge n.$$

- Let $\ell^* = \min\{i \geq 1 : j_i = n\}$.
- Define a labeled ordered rooted tree $T \in \mathcal{L}_{\mathrm{d}}$ as follows. For $1 \leq i \leq n-1$, if $i+1 \notin \{j_1, \ldots, j_{\ell^*}\}$ then set $v_i c_i = \sigma_{v_i}^{-1}(c_i) := v_{i+1}$. If $i+1 = j_k$ for some $1 \leq k \leq \ell^*$ then set $\sigma_{v_i}^{-1}(c_i) = m_k$.
- Set $B(\pi) = T$.

The rightmost tree in Figure 2.3 is the tree $B(\pi)$ where π is equal to π_T from (2.2.1).

When needed, we will emphasise the dependence of the quantities m_i, j_i and ℓ^* on π by writing $m_i(\pi)$, $j_i(\pi)$ and $\ell^*(\pi)$. Setting $j_0 = 1$ for convenience, we may think of $T = B(\pi)$ as the union of the paths $P^{(1)}, \ldots, P^{(\ell^*)}$, where $P_i = v_{j_{i-1}} \ldots v_{j_i - 1} m_i$ is the path in T from $v_{j_{i-1}}$ to m_i. Note that since $m_i \geq i$ for all $i \in [\ell^*]$, vertices $1, \ldots, k$ are contained within the union of paths $P^{(1)}, \ldots, P^{(k \wedge \ell^*)}$ for all $k \in [n]$.

Recall from Section 2.1.1 that T_n denotes a Bienaymé tree with offspring distribution μ conditioned to have n vertices. Suppose now that $D^n = (D_1^n, \ldots, D_n^n)$ is a sequence of IID μ-distributed random variables conditioned to have total sum $\sum_{i=1}^n D_i^n = n-1$, and let $\Pi_{D^n} \in_{\mathcal{U}} \mathcal{P}_{D^n}$. Then the tree T_n has the same law as $B(\Pi_{D^n})$. Furthermore, $\mathbf{T}_n = (\mathrm{T}_n, Y)$, which we refer to as a (μ, ν)-branching random walk (conditioned to have size n), has the same law as $(B(\Pi_{D^n}), Y)$. (Here, conditionally on the underlying tree T, $Y = (Y^{(v)}, v \in v(T) \setminus \partial T)$ are independent random vectors such that if $c(v, T) = k$ then $Y^{(v)}$ has distribution ν_k). The associated spatial tree $(B(\Pi_{D^n}), \ell)$ is such that $\ell(v_1) = 0$, and for $0 \leq i \leq n-1$ if $i + 1 \notin \{j_1, \ldots, j_{\ell^*}\}$,

$$\ell(v_{i+1}) = \ell(v_i) + Y_{c_i}^{(v_i)},$$

and if $i + 1 = j_k$ for some $1 \leq k \leq \ell^*$, then

$$\ell(m_k) = \ell(v_i) + Y_{c_i}^{(v_i)}.$$

In Section 2.3 we study the above bijective construction of uniform trees with a given deterministic degree sequence d; that is, for $T = B(\Pi_{\mathrm{d}})$ for $\Pi_{\mathrm{d}} \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$. We note however, that by conditioning on D^n, all results in Section 2.3 also apply to T_n and, consequently, to the underlying tree of $\mathbf{T}_n = (\mathrm{T}_n, Y)$.

2.3 Sampling from $\mathcal{L}_{\mathrm{d}}$

Fix a degree sequence $\mathrm{d} = (d_1, \ldots, d_n)$. The bijection B applied to a uniform element $\Pi_{\mathrm{d}} \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$ yields a uniform element $T = B(\Pi_{\mathrm{d}})$ of $\mathcal{L}_{\mathrm{d}}$. We can think of the bijection as constructing T from Π_{d} by adding vertices one at a time in order of their first appearance in a pair (V, C) of Π_{d}. Below, we use this perspective to study properties of T, in particular the law of the sequence of vertices ordered by first appearance in a pair (V, C) of Π_{d}, and the law of the number of vertices contained in the union of the paths $P^{(1)}, .., P^{(k)}$, for given $k \geq 1$.

2.3.1 Size-biased random re-ordering

For $n \geq 1$ let $\mathcal{S}_n$ denote the set of permutations of $[n]$. For $(k_1, \ldots, k_n) \in \mathbb{N}^n$, let $\Sigma = \Sigma_{(k_1, \ldots, k_n)}$ be the random permutation of $[n]$ with law given by

$$\mathbf{P}\{\Sigma = \sigma\} = \prod_{i=1}^{n} \frac{k_{\sigma(i)}}{\sum_{j=i}^{n} k_{\sigma(j)}}, \qquad \text{for } \sigma \in \mathcal{S}_n.$$

We call $(k_{\Sigma(1)}, \dots, k_{\Sigma(n)})$ the *size-biased random re-ordering* of $(k_1, \dots, k_n)$.

For a degree sequence d, let $N_{\mathrm{d}} = |\{i \in [n] : d_i > 0\}|$. For $\pi = ((v_1, c_1), \dots, (v_n, c_n)) \in \mathcal{P}_{\mathrm{d}}$ we let $\hat{v}_1(\pi), \dots, \hat{v}_{N_{\mathrm{d}}}(\pi)$ denote the internal vertices in $T = B(\pi)$ ordered by their first appearance in a pair (v, c) in π. When $\pi = \Pi_{\mathrm{d}} = ((V_1, C_1), \dots, (V_{n-1}, C_{n-1})) \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$ is random, we write $\widehat{V}_i(\Pi_{\mathrm{d}}) = \hat{v}_i(\Pi_{\mathrm{d}})$ to reinforce the fact that the order of the vertices is random. The next lemma states that $(\widehat{V}_1(\Pi_{\mathrm{d}}), \dots, \widehat{V}_{N_{\mathrm{d}}}(\Pi_{\mathrm{d}}))$ are the vertices corresponding to a size-biased random reordering of $\{d_i \ : \ d_i > 0, i \in [n]\}$.

Lemma 2.3.1. *Fix a degree sequence* $\mathrm{d} = (d_1, \dots, d_n)$ *and let* $\Pi_{\mathrm{d}} \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$. *Then for any permutation* $(i_1, \dots, i_{N_{\mathrm{d}}})$ *of* $\{i \in [n] : d_i > 0\}$,

$$\mathbf{P}\left\{(\widehat{V}_1(\Pi_{\mathrm{d}}), \dots, \widehat{V}_{N_{\mathrm{d}}}(\Pi_{\mathrm{d}})) = (i_1, \dots, i_{N_{\mathrm{d}}})\right\} = \frac{d_{i_1}}{n-1} \frac{d_{i_2}}{n-1-d_{i_1}} \cdots \frac{d_{i_{N_{\mathrm{d}}}}}{n-1-\sum_{j=1}^{N_{\mathrm{d}}-1} d_{i_j}}$$

Consequently, the size-biased random reordering of the positive entries of d *is equal in distribution to* $(d_{\widehat{V}_1(\Pi_{\mathrm{d}})}, \dots, d_{\widehat{V}_{N_{\mathrm{d}}}(\Pi_{\mathrm{d}})})$.

Proof. We show the statement by induction on N_{d}. For $N_{\mathrm{d}} = 1$, the statement is immediate for all n and for all degree sequences of length n with $|\{i : d_i > 0\}| = 1$ since if $N_{\mathrm{d}} = 1$ there is a single vertex of positive degree and $\widehat{V}_1(\Pi_{\mathrm{d}}) = i_1$.

Next, fix $\ell \in \mathbb{N}$ and suppose the statement holds for all degree sequences d with $N_{\mathrm{d}} \leq \ell$. Then fix any degree sequence $\mathrm{d} = (d_1, \dots, d_n)$ with $N_{\mathrm{d}} = \ell + 1$, and any permutation $(i_1, \dots, i_{N_{\mathrm{d}}})$ of $\{i \in [n] : d_i > 0\}$. To specify an element of $\{\pi \in \mathcal{P}_{\mathrm{d}} : (\hat{v}_1(\pi), \dots, \hat{v}_{\ell+1}(\pi)) = (i_1, \dots, i_{\ell+1})\}$, it is necessary and sufficient to specify

1. $\pi_1 = (v_1, c_1) \in \{(i_1, c) : c \in [d_{i_1}]\}$;
2. The $d_{i_1} - 1$ values $j \in \{2, 3, \dots, n-1\}$ for which $\pi_j = (i_1, c)$ for some $1 \leq c \leq d_{i_1}$;
3. The order of the $d_{i_1} - 1$ elements of $\{(i_1, c), 1 \leq c \leq d_{i_j}\} \backslash \{\pi_1\}$ in π;
4. The order of the elements of $\{(i_j, c), 2 \leq j \leq \ell + 1, 1 \leq c \leq d_{i_j}\}$ in π, which must ensure that $(\hat{v}_2(\pi), \dots, \hat{v}_{\ell+1}(\pi)) = (i_2, \dots, i_{\ell+1})$.

By the induction hypothesis applied to $(d_{i_2}, \dots, d_{i_{\ell+1}}, 0, \dots, 0) \in \mathbb{Z}_{\geq 0}^{n-d_{i_1}}$ this implies that

$$\begin{aligned}
&|\{\pi \in \mathcal{P}_{\mathrm{d}} : (\hat{v}_1(\pi), \dots, \hat{v}_{\ell+1}(\pi)) = (i_1, \dots, i_{\ell+1})\}| \\
&\quad = d_{i_1} \binom{n-2}{d_{i_1} - 1} (d_{i_1} - 1)!(n - 1 - d_{i_1})! \frac{d_{i_2}}{n-1-d_{i_1}} \cdots \frac{d_{i_{\ell+1}}}{n-1-\sum_{j=1}^{\ell} d_{i_j}} \\
&\quad = (n-1)! \frac{d_{i_1}}{n-1} \frac{d_{i_2}}{n-1-d_{i_1}} \cdots \frac{d_{i_{\ell+1}}}{n-1-\sum_{j=1}^{\ell} d_{i_j}} ; \qquad (2.3.1)
\end{aligned}$$

since $|\mathcal{P}_{\mathrm{d}}| = (n-1)!$, the claim follows. □

2.3.2 Repeats in Π_{d}

Let $\Pi_{\mathrm{d}} = ((V_1, C_1), \ldots, (V_{n-1}, C_{n-1})) \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$. Recall that $\widehat{V}_1(\Pi_{\mathrm{d}}), \ldots, \widehat{V}_{N_{\mathrm{d}}}(\Pi_{\mathrm{d}})$ are the internal vertices in $B(\Pi_{\mathrm{d}})$ ordered by their first appearance in a pair (V, C) in Π_{d}.

For $i \in [\ell^*(\Pi_{\mathrm{d}})]$ let $M_i^{\mathrm{d}} = m_i^{\mathrm{d}}(\Pi_{\mathrm{d}})$ and $J_i^{\mathrm{d}} = j_i^{\mathrm{d}}(\Pi_{\mathrm{d}})$. We introduce this notation to emphasise that $M_1^{\mathrm{d}}, \ldots, M_{\ell^*(\Pi_{\mathrm{d}})}^{\mathrm{d}}$ and $J_1^{\mathrm{d}}, \ldots, J_{\ell^*(\Pi_{\mathrm{d}})}^{\mathrm{d}}$ are random variables. We will see later that for the random degree sequences $D^n = (D_1^n, \ldots, D_n^n)$ arising in this paper, for $k \geq 1$ fixed and for n large, $\{V_1, \ldots, V_{J_k^{D^n}}\} \cap [k] = \emptyset$ with high probability. In this case, for each $i \in [k]$ the first coordinate of the pair $(V_{J_i^{D^n}}, C_{J_i^{D^n}}) \in \Pi_{D^n}$, corresponds to a *repeated* first coordinate of Π_{D^n}. It is therefore convenient to define a second set of indices which correspond to the indices of Π_{d} for which the first coordinate is a repeat. Specifically, let $\widetilde{J}_1^{\mathrm{d}} = \inf\{j > 1 : V_j \in \{V_1, \ldots, V_{j-1}\}\} \wedge n$, and for $i \geq 1$, let

$$\widetilde{J}_{i+1}^{\mathrm{d}} = \inf\{j > \widetilde{J}_i^{\mathrm{d}} : V_j \in \{V_1, \ldots, V_{j-1}\}\} \wedge n.$$

The next two lemmas describe the laws of $\widetilde{J}_1^{\mathrm{d}}$ and $(\widetilde{J}_i^{\mathrm{d}}, i \geq 2)$, respectively.

Lemma 2.3.2. *Fix an integer $n \geq 2$ and a degree sequence $\mathrm{d} = (d_1, \ldots, d_n)$ and let $\Pi_{\mathrm{d}} \in_{\mathcal{U}} \mathcal{P}_{\mathrm{d}}$. Then for $1 \leq k \leq N_{\mathrm{d}}$,*

$$\mathbf{P}\left\{\widetilde{J}_1^{\mathrm{d}} > k \;\middle|\; \widehat{V}_1(\Pi_{\mathrm{d}}), \ldots, \widehat{V}_{N_d}(\Pi_{\mathrm{d}})\right\} = \prod_{j=1}^{k}\left(1 - \frac{\sum_{i=1}^{j}(d_{\widehat{V}_i(\Pi_{\mathrm{d}})} - 1)}{n-j}\right),$$

and, for $k > N_{\mathrm{d}}$,

$$\mathbf{P}\left\{\widetilde{J}_1^{\mathrm{d}} > k \;\middle|\; \widehat{V}_1(\Pi_{\mathrm{d}}), \ldots, \widehat{V}_{N_d}(\Pi_{\mathrm{d}})\right\} = 0.$$

Proof. Observe that $\widetilde{J}_1^{\mathrm{d}} \leq N_{\mathrm{d}} + 1$ deterministically, so the statement for $k > N_{\mathrm{d}}$ is immediate. To prove the statement for $1 \leq k \leq N_{\mathrm{d}}$, fix any ordering $i_1, \ldots, i_{N_{\mathrm{d}}}$ of $\{i \in [n] : d_i > 0\}$. Then using Bayes' formula and the fact that $|\mathcal{P}_{\mathrm{d}}| = (n-1)!$, the probability

$$\mathbf{P}\left\{\widetilde{J}_1^{\mathrm{d}} > k \;\middle|\; (\widehat{V}_1(\Pi_{\mathrm{d}}), \ldots, \widehat{V}_{N_d}(\Pi_{\mathrm{d}})) = (i_1, \ldots, i_{N_{\mathrm{d}}})\right\}$$

may be expressed as a ratio with denominator

$$|\{\pi = ((v_i, c_i), i \in [n]) \in \mathcal{P}_{\mathrm{d}} : (\hat{v}_j(\pi))_{1 \leq j \leq N_d} = (i_j)_{1 \leq j \leq N_d}\}|$$

and numerator

$$|\{\pi = ((v_i, c_i), i \in [n]) \in \mathcal{P}_{\mathrm{d}} : (\hat{v}_j(\pi))_{1 \leq j \leq N_d} = (i_j)_{1 \leq j \leq N_d}, (v_j)_{1 \leq j \leq k} = (i_j)_{1 \leq j \leq k}\}|\,.$$

Equation (2.3.1) directly yields a formula for the denominator. Also, letting d' be the degree sequence $(d_{i_k+1}, \ldots, d_{N_d}, 0, \ldots, 0) \in \mathbb{Z}_{\geq 0}^{n-d_{i_1}-\ldots-d_{i_k}}$, then the numerator is

$$\prod_{j=1}^{k} d_{i_j} \cdot (n-1-k)_{d_{i_1}+\ldots+d_{i_k}-k} \cdot \left|\{\pi' \in \mathcal{P}_{\mathrm{d}'} : (\hat{v}_1(\pi'), \ldots, \hat{v}_{N_\mathrm{d}-k}(\pi')) = (i_{k+1}, \ldots, i_{N_d})\}\right|.$$

The first term selects $c_j \in [d_{i_j}]$ for each $j \in [k]$; the second, falling factorial term selects the locations of the remaining entries of π whose first coordinate belongs to $\{i_1, \ldots, i_k\}$; and the third term specifies the order of the remaining entries of $\pi \in \mathcal{P}_\mathrm{d}$. Equation (2.3.1) also gives a formula for this final term, and the lemma then follows by routine algebra. □

Lemma 2.3.3. *Fix a degree sequence* $\mathrm{d} = (d_1, \ldots, d_n)$ *and let* $\Pi_\mathrm{d} \in_\mathcal{U} \mathcal{P}_\mathrm{d}$*. Let* $i \geq 1$*. Then for* $n \geq 2$ *and* k *such that* $\widetilde{J}_i^\mathrm{d} + k \in [N_\mathrm{d}]$,

$$\begin{aligned}
\mathbf{P}&\left\{\widetilde{J}_{i+1}^\mathrm{d} > \widetilde{J}_i^\mathrm{d} + k \;\middle|\; \widetilde{J}_1^\mathrm{d}, \ldots, \widetilde{J}_i^\mathrm{d}, \widehat{V}_1(\Pi_\mathrm{d}), \ldots, \widehat{V}_{N_\mathrm{d}}(\Pi_\mathrm{d})\right\} \\
&= \prod_{j=\widetilde{J}_i^\mathrm{d}}^{\widetilde{J}_i^\mathrm{d}+k} \left(1 - \frac{\sum_{\ell=1}^{j}(d_{\widehat{V}_\ell(\Pi_\mathrm{d})} - 1) - i}{n-j}\right)
\end{aligned}$$

and, for $k > N_\mathrm{d}$,

$$\mathbf{P}\left\{\widetilde{J}_{i+1}^\mathrm{d} > k \;\middle|\; \widehat{V}_1(\Pi_\mathrm{d}), \ldots, \widehat{V}_{N_\mathrm{d}}(\Pi_\mathrm{d})\right\} = 0.$$

The proof of Lemma 2.3.3 is analogous to that of Lemma 2.3.2 and is therefore omitted. Finally, a bound we will need in Section 2.5, whose proof relies on the bijective construction of T_n, is the following; its proof is postponed to Section 2.8.

Lemma 2.3.4. *Let* $\mathrm{d} = (d_1, \ldots, d_n)$ *be a degree sequence and let* $\mathcal{B} \subset [n]$ *be a set of vertices. Suppose that* $|\mathcal{B}| \leq K$ *and suppose that* $\max_{1 \leq i \leq n} d_i \leq \Delta$*. Let* B_d *be the smallest distance between two vertices in* $\mathcal{B}$ *that are ancestrally related in* $T_\mathrm{d} = B(\Pi_\mathrm{d})$ *(with* $B_\mathrm{d} = \infty$ *if no vertices in* $\mathcal{B}$ *are ancestrally related). Then, for any* $b \geq 0$

$$\mathbf{P}\{B_\mathrm{d} \leq b\} \leq K\left(1 - \left(1 - \frac{K\Delta}{n-1-b\Delta}\right)^b\right).$$

2.4 Random finite-dimensional distributions

In this section we use the bijection B to prove the convergence of the random finite-dimensional distributions of the head of the discrete snake (H_n, R_n). We assume throughout this section that μ is critical and has variance $\sigma^2 \in (0, \infty)$, and that assumption [**A1**] holds.

Recall that T_n is a Bienaymé tree with offspring distribution μ conditioned to have n vertices, and that $\mathbf{T}_n = (\mathrm{T}_n, Y)$ denotes the conditioned (μ, ν)-branching random walk. By Section 2.2.2, $\mathbf{T}_n$ has the same distribution as $(B(\Pi_{D^n}), Y)$, where $D^n = (D_1^n, \ldots, D_n^n)$ is a sequence of IID μ-distributed random variables conditioned to have total sum $\sum_{i=1}^n D_i^n = n-1$ and, conditionally on D^n, $\Pi_{D^n} = ((V_1, C_1), \ldots, (V_{n-1}, C_{n-1})) \in_{\mathcal{U}} \mathcal{P}_{D^n}$.

Fix $k \geq 1$. Let $U_1^n, \ldots, U_k^n$ be a uniformly random k-set of indices chosen from $[n]$. Let $\mathrm{T}_n(U_1^n, \ldots, U_k^n)$ be the subtree of T_n spanned by the root of T_n and the vertices $v_{U_1^n}, \ldots,$ $v_{U_k^n}$, where for $i \in [n]$, we recall that v_i is the i-th vertex in the lexicographical order of T_n. (For fixed k, as $n \to \infty$, a collection of k IID Uniform($[n]$) random variables will be distinct with probability tending to 1, so we can treat $U_1^n, \ldots, U_k^n$ as indistinguishable from independent uniform picks from the vertices.) We immediately observe that $\mathrm{T}_n(U_1^n, \ldots, U_k^n)$ has the same distribution as T_n^k, the subtree of $B(\Pi_{D^n})$ spanned by the root and the vertices $1, \ldots, k$. Since T_n^k is more convenient for our analysis, we will work with it instead. Note that T_n^k is a labeled ordered rooted tree whose leaves are labeled by $1, 2, \ldots, k$. Write ℓ_n^k for the map from T_n^k into $\mathbb{R}$ which gives the spatial locations of the vertices, so that $(\mathrm{T}_n^k, \ell_n^k)$ is the spatial tree (T_n, ℓ) restricted to the subtree spanned by the root and the vertices $1, \ldots, k$.

Let $\mathcal{T}_{2\mathbf{e}}$ denote the Brownian tree encoded by the excursion $2\mathbf{e}$, and let $U_1, \ldots, U_k$ be IID Uniform($[0,1]$) random variables, independent of $\mathbf{e}$. Recall that $\mathcal{T}_{2\mathbf{e}}^k = \mathcal{T}_{2\mathbf{e}}(U_1, \ldots, U_k)$ denotes the subtree of $\mathcal{T}_{2\mathbf{e}}$ spanned by the images of 0 and $U_1, \ldots, U_k$ in $\mathcal{T}_{2\mathbf{e}}$, thought of as an ordered rooted tree with leaves labeled by $1, 2, \ldots, k$ and with real-valued edge lengths. Recall that $\mathcal{T}_{2\mathbf{e}}^k$ has the same distribution as the tree $\mathcal{T}^k$ built by Aldous' line-breaking construction. We now introduce a version of the line-breaking construction which incorporates spatial locations.

Line-breaking construction of the Brownian tree with spatial locations

We construct a sequence $(\mathcal{T}^k)_{k\geq 1}$ of trees along with two functions $\mathbf{h} : [0,\infty) \to [0,\infty)$ and $\mathbf{l} : [0,\infty) \to \mathbb{R}$ recursively. Let $J_1, J_2, \dots$ be the jump times of a Poisson point process on $[0,\infty)$ with intensity tdt at time t, listed in increasing order. Independently, let $(B_t)_{t\geq 0}$ be a standard Brownian motion. Start from the tree $\mathcal{T}^1$ which consists of the line-segment $[0, J_1]$. Define $\mathbf{h}(t) = t$ and $\mathbf{l}(t) = B_t$ for $0 \leq t \leq J_1$. Recursively, for $k \geq 2$, conditionally on J_{k-1}, sample an attachment point $A_{k-1} \sim \text{Uniform}([0, J_{k-1}])$, independent of $(A_j)_{j<k-1}$. Take the completion of the line segment $(J_{k-1}, J_k]$, and let J^*_{k-1} denote the limit point as $x \downarrow J_{k-1}$. Identify the points J^*_{k-1} and A_{k-1}. This has the effect of gluing the line-segment $(J_{k-1}, J_k]$ onto $\mathcal{T}^{k-1}$. We do this with probability $1/2$ to the left side and with probability $1/2$ to the right side. This yields $\mathcal{T}^k$. Define $\mathbf{h}(t) = \mathbf{h}(A_{k-1}) + t - J_{k-1}$ and $\mathbf{l}(t) = \mathbf{l}(A_{k-1}) + B_t - B_{J_{k-1}}$ for $t \in (J_{k-1}, J_k]$ to determine the height and location processes on the new line-segment.

The planar embedding of $\mathcal{T}^k$ is captured by a permutation $\tau^k : [k] \to [k]$ which is such that $\tau^k(1), \dots, \tau^k(k)$ is the order in which we observe the leaves when exploring the tree from left to right. Using the notation $U^k_{(1)}, \dots, U^k_{(k)}$ for the increasing ordering of $U_1, \dots, U_k$ as in Proposition 2.1.10, we then have

$$\begin{aligned}\Big(\mathbf{h}(J_{\tau^k(1)}), \dots, \mathbf{h}(J_{\tau^k(k)}), \mathbf{l}(J_{\tau^k(1)}), \dots, \mathbf{l}(J_{\tau^k(k)})\Big) \\ \overset{\mathrm{d}}{=} \left(2\mathbf{e}_{U^k_{(1)}}, \dots, 2\mathbf{e}_{U^k_{(k)}}, \sqrt{2}\mathbf{r}_{U^k_{(1)}}, \dots, \sqrt{2}\mathbf{r}_{U^k_{(k)}}\right),\end{aligned} \tag{2.4.1}$$

where the equality in distribution of the first k co-ordinates on the two sides is a consequence of Corollary 22 of Aldous [9], and that of the final k co-ordinates is a consequence of the definition of the Brownian snake given at (2.1.1). So the line-breaking construction indeed realises the random finite-dimensional distributions of the head of the Brownian snake.

We show that the scaling limit of $(\mathrm{T}^k_n, \ell^k_n)$ is $(\mathcal{T}^k, \mathbf{l}|_{[0,J_k]})$ in an appropriate sense, which will allow us to prove the convergence of the random finite-dimensional distributions, along with a certain amount of extra information which will be useful to us in Section 2.5 where we prove tightness.

Recall that the tree T^k_n necessarily sits within the first k paths, $P^{(1)}, \dots, P^{(k)}$, in the discrete line-breaking construction. We need to understand the lengths of these paths, and the positions at which the paths are glued onto one another. It is convenient to use the *indices* of the vertices in Π_{D^n} for this purpose rather than the vertex labels themselves. Recall that $J^{D^n}_1, J^{D^n}_2, \dots, J^{D^n}_k$ are the first k indices at which we see either a repeat or an element of $\{1, 2, \dots, k\}$. Let us henceforth write $J^n_i = J^{D^n}_i$ (and also $\widetilde{J}^n_i = \widetilde{J}^{D^n}_i$) for $i \geq 1$. Then the lengths of the paths $P^{(1)}, \dots, P^{(k)}$

are given by $J_1^n, J_2^n - J_1^n, \dots, J_k^n - J_{k-1}^n$. For $1 \le m \le k-1$, the index at which the path $P^{(m+1)}$ attaches onto the subtree constructed from the first m paths is given by the value i such that $V_{J_m^n} = \widehat{V}_i(\Pi_{D^n})$ (i.e. we find the index of the vertex $V_{J_m^n}$ within the vector $(\widehat{V}_1(\Pi_{D^n}), \dots, \widehat{V}_{N_n}(\Pi_{D^n})))$. We write A_m^n for this value i and call this the m-th *attachment point*. See Figure 2.4.

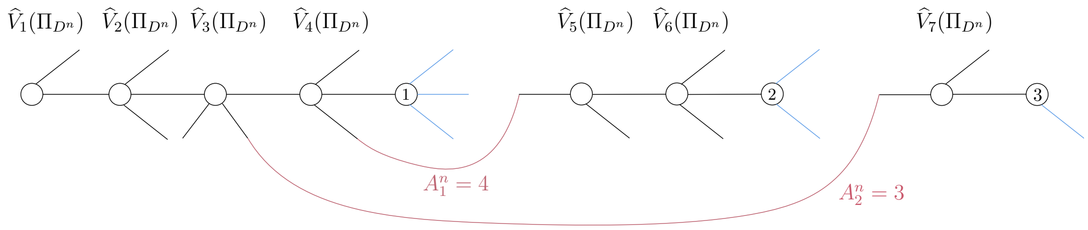

Figure 2.4: Illustration of the first and second attachment points, A_1^n and A_2^n.

Since T_n^k is an ordered tree, we will need to understand where the paths $P^{(2)}, \dots, P^{(k)}$ attach relative to the pre-existing children of their attachment points. If we are looking to attach to a vertex which has only one pre-existing child (i.e. for which there has been no previous repeat) then that vertex must have degree $d \ge 2$, and then whether we attach to the left or to the right of the pre-existing child is simply determined by the relative ordering of the corresponding second coordinates in the sequence Π_{D^n}. If there has been no previous repeat at this vertex then this pair of second coordinates is chosen uniformly at random without replacement from $[d]$ and, in particular, we attach to the left and right sides each with probability $1/2$. This ceases to be true after the first repeat (not least because then there are three or more children whose relative ordering we need to understand), but as we shall show below, we observe a second repeat of any vertex in T_n^k with vanishing probability as $n \to \infty$. Let $F_1^n, \dots, F_k^n$ be random variables taking values in $\{0, 1, 2\}$ such that $F_i^n = 1$ if $P^{(i+1)}$ attaches at a first repeat and to the left-hand side, $F_i^n = 2$ if $P^{(i+1)}$ attaches at a first repeat and to the right-hand side and $F_{i+1}^n = 0$ otherwise, for $1 \le i \le k$.

Finally, recall that vertex $\widehat{V}_i(\Pi_{D^n})$ has degree $D^n_{\widehat{V}_i(\Pi_{D^n})}$ for $i \le N_{D^n} = |\{i \in [n] : D_i^n > 0\}|$. Let $L^n(0) = 0$ and let $L^n(i)$ be the spatial location of the C_i-th child of vertex V_i in line-breaking construction $B(\Pi_{D^n})$, for $1 \le i \le n-1$.

The following proposition shows that, on rescaling, these quantities converge in distribution to their analogues in the line-breaking construction of the Brownian tree with spatial locations.

Proposition 2.4.1. *Fix $k \geq 1$. Then*

$$\frac{\sigma}{\sqrt{n}}(J_1^n, J_2^n, \ldots, J_k^n, A_1^n, \ldots, A_k^n) \xrightarrow{\mathrm{d}} (J_1, J_2, \ldots, J_k, A_1, \ldots, A_k) \tag{2.4.2}$$

as $n \to \infty$. Jointly with this convergence, we have that

$$(F_1^n, F_2^n, \ldots, F_k^n) \xrightarrow{\mathrm{d}} (F_1, F_2, \ldots, F_k), \tag{2.4.3}$$

where $F_1, F_2, \ldots, F_k$ are IID *random variables, independent of everything else, such that $\mathbf{P}\{F_i = 1\} = \mathbf{P}\{F_i = 2\} = 1/2$ and*

$$\left(\frac{L^n(\lfloor tn^{1/2}\rfloor \wedge (J_1^n - 1))}{n^{1/4}}\right)_{t\geq 0} \xrightarrow{\mathrm{d}} \beta(B_{t\wedge(J_1/\sigma)})_{t\geq 0},$$
$$\left(\frac{L^n((J_i^n + \lfloor tn^{1/2}\rfloor) \wedge (J_{i+1}^n - 1))}{n^{1/4}}\right)_{t\geq 0} \xrightarrow{\mathrm{d}} \beta(B_{A_i/\sigma} + B_{((J_i/\sigma)+t)\wedge(J_{i+1}/\sigma)} - B_{(J_i/\sigma)})_{t\geq 0} \tag{2.4.4}$$

for $1 \leq i \leq k-1$, in each case for the uniform norm.

As a corollary, we obtain the convergence of the random finite-dimensional distributions in (2.1.3).

Corollary 2.4.2. *For any $k \geq 1$, as $n \to \infty$*

$$\left(\frac{H_n(nU_{(1)}^k)}{\sqrt{n}}, \ldots, \frac{H_n(nU_{(k)}^k)}{\sqrt{n}}, \frac{R_n(nU_{(1)}^k)}{n^{1/4}}, \ldots, \frac{R_n(nU_{(k)}^k)}{n^{1/4}}\right)$$
$$\xrightarrow{\mathrm{d}} \left(\frac{2}{\sigma}\mathbf{e}_{U_{(1)}^k}, \ldots, \frac{2}{\sigma}\mathbf{e}_{U_{(k)}^k}, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_{U_{(1)}^k}, \ldots, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_{U_{(k)}^k}\right),$$

where $U_{(1)}^k, \ldots, U_{(k)}^k$ are the order statistics of k IID *Uniform$([0,1])$ random variables.*

Proof. Let $(U_1^n, \ldots, U_k^n)$ be a uniformly random k-set chosen from $[n]$, and let $(U_{(1)}^{n,k}, \ldots, U_{(k)}^{n,k})$ be the order statistics of $(U_1^n, \ldots, U_k^n)$. As argued above, we may straightforwardly replace $nU_{(1)}^k, \ldots, nU_{(k)}^k$ by $(U_{(1)}^{n,k} - 1, \ldots, U_{(k)}^{n,k} - 1)$ at no asymptotic cost. Recall that $H_n(i)$ gives the distance from the root of the $(i+1)$-th vertex visited in a depth-first exploration of the tree. The random variables

$$(H_n(U_1^n - 1), \ldots, H_n(U_k^n - 1))$$

have the joint law of the distances from the root to the leaves labeled $1, 2, \ldots, k$ in T_n^k (these may be expressed in terms of sums and differences of elements of $(J_1^n, \ldots, J_k^n, A_1^n, \ldots, A_k^n)$ analogously to the definition of $\mathbf{h}$ in the line-breaking construction of the Brownian tree with spatial locations), and

$$(R_n(U_1^n - 1), \ldots, R_n(U_k^n - 1)) = (L^n(J_1^n - 1), \ldots, L^n(J_k^n - 1)).$$

The effect of ordering the uniforms is simply to apply the same permutation of the entries to each of $(H_n(U_1^n-1),\ldots,H_n(U_k^n-1))$ and $(R_n(U_1^n-1),\ldots,R_n(U_k^n-1))$. This permutation is straightforwardly induced by the choices $(F_1^n,\ldots,F_{k-1}^n)$. By (2.4.3), this permutation then converges in distribution to τ^k. But then the claimed convergence follows from Proposition 2.4.1 using the scaling property of Brownian motion and (2.4.1). □

We begin by studying the vertex degrees at the start of the bijective construction, and show that, on the timescale of $\sqrt{n}$, the degrees that we observe are asymptotically indistinguishable from IID copies of $\bar{\xi}$. We show further that the subtree T_n^k is constructed on a timescale of order $\sqrt{n}$. This allows us to prove (2.4.2) in Proposition 2.4.7. To get the convergence of the spatial locations, we observe that, with the exception of branch points, the displacements along the ancestral lineages in T_n^k are asymptotically indistinguishable from IID copies of $Y_{\bar{\xi},U_{\bar{\xi}}}$. Combining this with the convergence of the tree allows us to obtain the convergence of the spatial locations along the branches of the subtree.

2.4.1 A discrete change of measure

In this subsection, we show that the size-biased random re-ordering of the positive entries of D^n may be viewed as a vector of IID copies of the size-biased offspring random variable $\bar{\xi}$ up to a change of measure. We study the behaviour of the Radon-Nikodym derivative and show that its effect is trivial on the first $O(\sqrt{n})$ entries of the vector. Recall that $N_{D^n}=|\{i\in[n]:D_i^n>0\}|$. To ease the notation, we write $N_n=N_{D^n}$. Let

$$\widehat{D}^n=\left(\widehat{D}_1^n,\ldots,\widehat{D}_{N_n}^n\right)$$

be the size-biased random re-ordering of the positive entries of D^n. We note that

$$\widehat{D}^n\overset{\mathrm{d}}{=}\left(D^n_{\widehat{V}_1(\Pi_{D^n})},\ldots,D^n_{\widehat{V}_{N_n}(\Pi_{D^n})}\right).$$

Later we will often somewhat abuse notation and write $(\widehat{D}_1^n,\ldots,\widehat{D}_{N_n}^n)$ in place of $(D^n_{\widehat{V}_1(\Pi_{D^n})},\ldots D_{\widehat{V}_{N_n}(\Pi_{D^n})})$, for example in the proof of Proposition 2.4.7.

Proposition 2.4.3. *Let $\xi_1,\xi_2,\ldots,\xi_n$ be IID random variables with distribution μ. Further, let $\bar{\xi}_1,\bar{\xi}_2,\ldots$ be IID samples from the size-biased distribution of ξ_1. Then for $1\le m<n$, and any non-negative measurable function $f:\mathbb{Z}^m\to\mathbb{R}_+$,*

$$\mathbf{E}\left[f(\widehat{D}_1^n,\ldots,\widehat{D}_m^n)\mathbf{1}_{[N_n\ge m]}\right]=\mathbf{E}\left[f(\bar{\xi}_1,\ldots,\bar{\xi}_m)\Theta^n(\bar{\xi}_1,\ldots,\bar{\xi}_m)\right],$$

where for $k_1, \dots, k_m \in \mathbb{N}$,

$$\Theta^n(k_1,\dots,k_m) = \frac{\mathbf{P}\left\{\sum_{i=m+1}^n \xi_i = n-1-\sum_{i=1}^m k_i\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \prod_{i=1}^m \left(\frac{n-i+1}{n-1-\sum_{j=1}^{i-1} k_j}\right), \tag{2.4.5}$$

if $k_1 + \dots + k_m \le n-1$, *and* $\Theta^n(k_1,\dots,k_m) = 0$ *otherwise.*

Proposition 2.4.3 is a special case of Proposition 2.8.4 that we state and prove in Section 2.8, and use in full generality to prove Theorems 2.1.4 and 2.1.5 in Section 2.7. We state only the special case here as the more general formulation is much more technical and requires definitions are only relevant in settings where assumption [**A3**] holds.

The next lemma shows that the change of measure Θ^n appearing in Proposition 2.4.3 is asymptotically unimportant provided that $m = \Theta(\sqrt{n})$.

Lemma 2.4.4. *Let μ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$, and let $(\bar{\xi}_i)_{i\ge 1}$ be* IID *samples from the size-biased distribution of μ. Suppose that $m = m(n) = \Theta(\sqrt{n})$. Then as $n \to \infty$*

$$\Theta^n(\bar{\xi}_1,\dots,\bar{\xi}_m) \xrightarrow{\text{P}} 1,$$

and $(\Theta^n(\bar{\xi}_1,\dots,\bar{\xi}_m))_{n\ge 1}$ is a uniformly integrable sequence of random variables.

Proof. By a subsubsequence argument we may assume that $m/\sqrt{n} \to t$ as $n \to \infty$ for some $t > 0$. Let $\xi_1,\dots,\xi_n$ be IID random variables with distribution μ. We deal with the ratio of probabilities in the definition of Θ^n using the local central limit theorem. Specifically, since $\mathbf{E}[\xi_1] = 1$ and $\mathbf{Var}\{\xi_1\} = \sigma^2$, we have that

$$\sup_{k\in\mathbb{Z}} \left| \sqrt{n-m}\cdot \mathbf{P}\left\{\sum_{i=m+1}^n \xi_i = n-1-m+k\right\} - \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{k^2}{2\sigma^2(n-m)}\right)\right| \to 0$$

as $n \to \infty$, so for $k_1,\dots,k_m \in \mathbb{N}$,

$$\begin{aligned}
&\mathbf{P}\left\{\sum_{i=m+1}^n \xi_i = n-1-\sum_{i=1}^m k_i\right\} \\
&\quad= \mathbf{P}\left\{\sum_{i=m+1}^n \xi_i - (n-m) = -1 - m\sigma^2 - \sum_{i=1}^m (k_i - 1 - \sigma^2)\right\} \\
&\quad= \frac{\exp\left(-\frac{1}{2\sigma^2(n-m)}\left(1 + m\sigma^2 + \sum_{i=1}^m (k_i - 1-\sigma^2)\right)^2\right)}{\sqrt{2\pi\sigma^2(n-m)}} + o(n^{-1/2}).
\end{aligned}$$

Similarly, we have that

$$\mathbf{P}\left\{\sum_{i=1}^{n} \xi_i = n-1\right\} = \frac{1}{\sqrt{2\pi\sigma^2 n}} + o(n^{-1/2}).$$

Therefore,

$$\frac{\mathbf{P}\left\{\sum_{i=m+1}^{n} \xi_i = n-1-\sum_{i=1}^{m} k_i\right\}}{\mathbf{P}\left\{\sum_{i=1}^{n} \xi_i = n-1\right\}} \\ = \exp\left(-\left(\frac{1+m\sigma^2+\sum_{i=1}^{m}(k_i-1-\sigma^2)}{\sqrt{2\sigma^2(n-m)}}\right)^2\right) + o(1). \tag{2.4.6}$$

Since the random variables $\bar{\xi}_1, \ldots, \bar{\xi}_n$ are IID with mean σ^2+1, by the functional strong law of large numbers (as stated in Lemma 2.8.1), as $n \to \infty$,

$$\frac{1}{\sqrt{n}} \max_{1 \le i \le \lfloor t\sqrt{n} \rfloor} \left| \sum_{j=1}^{i} (\bar{\xi}_j - 1 - \sigma^2) \right| \xrightarrow{\text{a.s.}} 0. \tag{2.4.7}$$

Since $m = (1+o(1))t\sqrt{n}$ this in particular yields that

$$\exp\left(-\left(\frac{1+m\sigma^2+\sum_{i=1}^{m}(\bar{\xi}_i-(1+\sigma^2))}{\sqrt{2\sigma^2(n-m)}}\right)^2\right) \xrightarrow{\text{p}} \exp\left(-\frac{t^2\sigma^2}{2}\right). \tag{2.4.8}$$

We claim that $n \to \infty$,

$$\prod_{i=1}^{m} \left(\frac{n-i+1}{n-1-\sum_{j=1}^{i-1} \bar{\xi}_j}\right) \xrightarrow{\text{p}} \exp\left(\frac{t^2\sigma^2}{2}\right). \tag{2.4.9}$$

Indeed,

$$\prod_{i=1}^{m} \left(\frac{n-i+1}{n-1-\sum_{j=1}^{i-1} \bar{\xi}_j}\right) = \exp\left(-\sum_{i=1}^{m} \log\left(1 - \frac{\sum_{j=1}^{i-1}(\bar{\xi}_j-1-\sigma^2)+\sigma^2(i-1)}{n-i+1}\right)\right).$$

It follows by Taylor's theorem and (2.4.7) that the last expression is equal to

$$\exp\left(\sum_{i=1}^{m} \frac{\sum_{j=1}^{i-1}(\bar{\xi}_j-1-\sigma^2)+\sigma^2(i-1)}{n-i+1} + o_{\mathbf{P}}(1)\right) \\ = \exp\left(\frac{\sigma^2\lfloor t\sqrt{n}\rfloor(\lfloor t\sqrt{n}\rfloor-1)}{2n} + o_{\mathbf{P}}(1)\right) \xrightarrow{\text{p}} \exp\left(\frac{t^2\sigma^2}{2}\right), \tag{2.4.10}$$

establishing (2.4.9). Combining this with (2.4.6) and (2.4.8) yields that

$$\Theta^n(\bar{\xi}_1, \ldots, \bar{\xi}_m) \xrightarrow{\text{p}} 1.$$

To prove uniform integrability, notice that, by applying Proposition 2.4.3 with $f \equiv 1$,

$$\mathbf{E}\left[\Theta^n(\bar{\xi}_1, \ldots, \bar{\xi}_m)\right] = \mathbf{P}\left\{N_n \ge m\right\}.$$

We claim that this tends to 1 as $n \to \infty$. To this end, note that

$$\#\{i \in [n] \;:\; \xi_i > 0\} \overset{\mathrm{d}}{=} \mathrm{Binomial}(n, 1-\mu_0).$$

So by conditioning on the event $\{\sum_{i=1}^n \xi_i = n-1\}$, which occurs with probability $\Theta(n^{-1/2})$, there are $(1+o_{\mathbf{P}}(1))n(1-\mu_0)$ non-zero entries of $(\xi_1, \ldots, \xi_n)$. Since $m = (1+o(1))t\sqrt{n}$ it follows that as $n \to \infty$,

$$\mathbf{P}\left\{N_n \geq m\right\} = \mathbf{P}\left\{\#\{i \in [n] \;:\; \xi_i > 0\} \geq m \;\middle|\; \sum_{i=1}^n \xi_i = n-1\right\} \to 1.$$

Uniform integrability then follows by the generalised Scheffé lemma, see [63, Theorem 5.12]. □

Lemma 2.4.4 implies that if the offspring distribution has finite variance then, on a timescale of $\sqrt{n}$ in the bijective construction $B(\Pi_{D^n})$ of T_n, the degrees we observe are asymptotically indistiguishable from IID copies of $\bar{\xi}$. To prove Proposition 2.4.1, we use this fact in the form of Proposition 2.4.5 stated below.

Proposition 2.4.5. *Given $\widehat{D}^n$, let $U_1, \ldots, U_n$ be independent random variables such that, for each $i \in [n]$, U_i is uniformly distributed on $[\widehat{D}^n_i]$. Further, let $Y_{\widehat{D}^n_1, U_1}, \ldots Y_{\widehat{D}^n_n, U_n}$ be independent random variables such that, for each $i \in [n]$, $Y_{\widehat{D}^n_i, U_i}$ is a uniform entry of a $\nu_{\widehat{D}^n_i}$-distributed displacement vector. If [**A1**] holds then as $n \to \infty$,*

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor t\sqrt{n}\rfloor} (\widehat{D}^n_i - 1), \frac{1}{n^{1/4}} \sum_{i=1}^{\lfloor t\sqrt{n}\rfloor} Y_{\widehat{D}^n_i, U_i}\right)_{t\geq 0} \overset{\mathrm{d}}{\longrightarrow} (\sigma^2 t, \beta B_t)_{t\geq 0},$$

for the topology of uniform convergence on compact time-intervals, where $(B_t)_{t\geq 0}$ is a standard Brownian motion.

Proof. Fix $T > 0$ and let $F : \mathbf{D}([0,T], \mathbb{R})^2 \to \mathbb{R}$ be a bounded continuous function, where $\mathbf{D}([0,T], \mathbb{R})$ is the space of real-valued functions on $[0,T]$ that are right-continuous with left limits equipped with the Skorokhod topology. Let $\bar{\xi}_1, \bar{\xi}_2, \ldots$ be IID samples from the size biased distribution of ξ. Further, independently for $i \geq 1$, let $\overline{U}_i$ be a Uniform($[\bar{\xi}_i]$) random variable.

By Proposition 2.4.3,

$$\begin{aligned}
&\mathbf{E}\left[F\left(\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor t\sqrt{n}\rfloor} (\widehat{D}^n_i - 1), \frac{1}{n^{1/4}} \sum_{i=1}^{\lfloor t\sqrt{n}\rfloor} Y_{\widehat{D}^n_i, U_i}\right)_{0\leq t\leq T}\right) \mathbf{1}_{[N_n \geq \lfloor T\sqrt{n}\rfloor]}\right] \\
&\quad = \mathbf{E}\left[F\left(\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor t\sqrt{n}\rfloor} (\bar{\xi}_i - 1), \frac{1}{n^{1/4}} \sum_{i=1}^{\lfloor t\sqrt{n}\rfloor} Y_{\bar{\xi}, \overline{U}_i}\right)_{0\leq t\leq T}\right) \Theta^n\left(\bar{\xi}_1, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}\right)\right],
\end{aligned} \tag{2.4.11}$$

where the random variables $(Y_{\bar{\xi}_i,\overline{U}_i})_{i\geq 1}$ are independent, and given $\bar{\xi}_i$, $Y_{\bar{\xi}_i,\overline{U}_i}$ is a uniform entry of a $\nu_{\bar{\xi}_i}$ distributed displacement vector. Since $\mathbf{E}\left[\bar{\xi}_1\right] = \sigma^2 + 1$, by the functional strong law of large numbers (Lemma 2.8.1), as $n \to \infty$,

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor t\sqrt{n}\rfloor}(\bar{\xi}_i - 1)\right)_{t\geq 0} \xrightarrow{\text{p}} (\sigma^2 t)_{t\geq 0}$$

in $\mathbf{D}((0,T),\mathbb{R})$.

Furthermore, the random variables $(Y_{\bar{\xi},\overline{U}_i})_{i\geq 1}$ are IID with mean and variance given by

$$\mathbf{E}\left[Y_{\bar{\xi}_1,\overline{U}_1}\right] = \sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\mathbf{E}\left[Y_{k,j}\right] = 0, \quad \mathbf{Var}\left\{Y_{\bar{\xi}_1,\overline{U}_1}\right\} = \sum_{k=1}^{\infty}\mu_k\sum_{j=1}^{k}\mathbf{E}\left[Y_{k,j}^2\right] = \beta^2.$$

It then follows from Donsker's theorem that as $n \to \infty$

$$\left(\frac{1}{n^{1/4}}\sum_{i=1}^{\lfloor t\sqrt{n}\rfloor}Y_{\bar{\xi}_i,\overline{U}_i}\right)_{t\geq 0} \xrightarrow{\text{d}} (\beta B_t)_{t\geq 0}$$

in $\mathbf{D}([0,T],\mathbb{R})$. Therefore by the continuity of F, as $n \to \infty$

$$\mathbf{E}\left[F\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor t\sqrt{n}\rfloor}(\bar{\xi}_i - 1), \frac{1}{n^{1/4}}\sum_{i=1}^{\lfloor t\sqrt{n}\rfloor}Y_{\bar{\xi}_i,\overline{U}_i}\right)_{0\leq t\leq T}\right)\right] \to \mathbf{E}\left[F\left((\sigma^2 t, \beta B_t)_{0\leq t\leq T}\right)\right].$$

Combining this with Lemma 2.4.4, and the boundedness of F, yields that (2.4.11) converges to

$$\mathbf{E}\left[F\left((\sigma^2 t, \beta B_t)_{0\leq t\leq T}\right)\right],$$

as $n \to \infty$ and the result follows. □

2.4.2 Bijective construction on the timescale $\sqrt{n}$

In this subsection we show that the subtree T_n^k is constructed on a timescale of order $\sqrt{n}$ with high probability. We then prove that the lengths of the paths which are glued together to form T_n^k converge on rescaling, as do the positions at which they attach to one another.

We begin by showing that, with high probability, the vertices $1, \ldots, k$ do not appear in the first $\Theta(\sqrt{n})$ entries of Π_{D^n}.

Lemma 2.4.6. *Fix $T > 0$ and $k \geq 1$, and let*

$$\mathcal{G}_{n,k}(T) = \left\{\left\{\widehat{V}_1(\Pi_{D^n}), \ldots, \widehat{V}_{\lfloor T\sqrt{n}\rfloor}(\Pi_{D^n})\right\} \cap \{1,\ldots,k\} = \emptyset, N_n \geq \lfloor T\sqrt{n}\rfloor\right\}.$$

Then

$$\mathbf{P}\left\{\mathcal{G}_{n,k}(T)\right\} \to 1$$

as $n \to \infty$.

Notice that on the *good event* $\mathcal{G}_{n,k}(T)$, if $J_k^n \le \lfloor T\sqrt{n} \rfloor$ then $J_i^{D^n} = \widetilde{J}_i^{D^n}$ for all $i \in [k]$, and T_n^k is precisely the tree spanned by the root and the paths $P^{(1)}, \dots, P^{(k)}$ in the bijective construction $B(\Pi_{D^n})$ of T_n.

Proof. We have

$$\begin{aligned}\mathbf{P}\left\{\mathcal{G}_{n,k}(T)\right\} &= \mathbf{P}\left\{\{\widehat{V}_1(\Pi_{D^n}), \dots, \widehat{V}_{\lfloor T\sqrt{n}\rfloor}(\Pi_{D^n}))\} \cap \{1, \dots, k\} = \emptyset, N_n \ge \lfloor T\sqrt{n}\rfloor\right\} \\ &\ge \mathbf{E}\left[\left(1 - \frac{D_1^n + \cdots + D_k^n}{n - 1 - T\sqrt{n}\max_{1\le i\le n} D_i^n}\right)^{\lfloor T\sqrt{n}\rfloor}\right] - \mathbf{P}\left\{N_n < \lfloor T\sqrt{n}\rfloor\right\}.\end{aligned}$$

Let $\varepsilon > 0$ and $(\xi_i)_{i\ge1}$ be a sequence of IID random variables with distribution μ. Then,

$$\begin{aligned}\mathbf{P}\left\{\max_{1\le i\le n} D_i^n > \varepsilon\sqrt{n}\right\} &= \frac{\mathbf{P}\left\{\max_{1\le i\le n}\xi_i > \varepsilon\sqrt{n}, \sum_{i=1}^n \xi_i = n-1\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \\ &\le \frac{n\mathbf{P}\left\{\xi_1 > \varepsilon\sqrt{n}, \sum_{i=1}^n \xi_i = n-1\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}}. \qquad (2.4.12)\end{aligned}$$

Since $\mathbf{E}\left[\xi^2\right] < \infty$ we have $n\mathbf{P}\left\{\xi_1 > \varepsilon\sqrt{n}\right\} \to 0$ as $n \to \infty$. Hence,

$$\begin{aligned}&\frac{n\mathbf{P}\left\{\xi_1 > \varepsilon\sqrt{n}, \sum_{i=1}^n \xi_i = n-1\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \\ &\quad\le \frac{n\mathbf{P}\left\{\xi_1 > \varepsilon\sqrt{n}\right\}\max_{\varepsilon\sqrt{n}<m\le n-1}\mathbf{P}\left\{\sum_{i=2}^n \xi_i = n-1-m\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \to 0\end{aligned}$$

as $n \to \infty$. Combining this with (2.4.12) gives that

$$\frac{1}{\sqrt{n}}\max_{1\le i\le n} D_i^n \xrightarrow{\mathrm{P}} 0$$

and so

$$\frac{D_1^n + \cdots + D_k^n}{\sqrt{n}} \xrightarrow{\mathrm{P}} 0$$

as $n \to \infty$. Therefore, by the bounded convergence theorem,

$$\mathbf{E}\left[\left(1 - \frac{D_1^n + \cdots + D_k^n}{n - 1 - T\sqrt{n}\max_{1\le i\le n} D_i}\right)^{\lfloor T\sqrt{n}\rfloor}\right] \to 1$$

as $n \to \infty$. By noting (as at the end of the proof of Lemma 2.4.4) that as $n \to \infty$,

$$\mathbf{P}\left\{N_n < \lfloor T\sqrt{n}\rfloor\right\} \le \frac{\mathbf{P}\left\{\mathrm{Binomial}(n, 1-\mu_0) < \lfloor T\sqrt{n}\rfloor\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \to 0,$$

the result follows. $\square$

Proposition 2.4.7. *Fix $k \geq 1$. Then as $n \to \infty$,*

$$\frac{\sigma}{\sqrt{n}}(J_1^n, J_2^n, \dots, J_k^n, A_1^n, \dots, A_k^n) \xrightarrow{\mathrm{d}} (J_1, J_2, \dots, J_k, A_1, \dots, A_k) \tag{2.4.13}$$

as $n \to \infty$, where $J_1, J_2, \dots, J_k$ are the first k jump-times of an inhomogeneous Poisson process of intensity t with respect to the Lebesgue measure at $t \in \mathbb{R}_+$ and, for $i \in [k]$, conditionally on $J_1, \dots, J_i$, A_i is uniform on $[0, J_i]$, independently of $A_1, \dots, A_{i-1}$.

Proof. Fix $T > 0$. Let $0 \leq t_1 \leq \dots \leq t_k \leq T$ and $s_1 < t_1, \dots, s_k < t_k$. We will prove that

$$\begin{aligned} &\mathbf{P}\left\{J_1^n \leq t_1\sqrt{n}, \dots, J_k^n \leq t_k\sqrt{n}, A_1^n \leq s_1\sqrt{n}, \dots, A_k^n \leq s_k\sqrt{n}\right\} \\ &\quad \to \sigma^{2k}\left(\prod_{j=1}^k s_j\right)\int_0^{t_1}\cdots\int_{r_{k-1}}^{t_k}\exp(-\sigma^2 t_k^2/2)dr_k\dots dr_1 \\ &\quad = \mathbf{P}\left\{J_1 \leq \sigma t_1, \dots J_k \leq \sigma t_k, A_1 \leq \sigma s_1, \dots, A_k \leq \sigma s_k\right\}. \end{aligned} \tag{2.4.14}$$

We will often work conditionally on the random variables $\widehat{D}^n = (\widehat{D}_1^n, \dots, \widehat{D}_{N_n}^n)$. To make the equations easier to read, we write $\mathbf{P}_{\widehat{D}^n}$ for the conditional probability given $\widehat{D}^n$ and $\mathbf{E}_{\widehat{D}^n}$ for the corresponding expectation.

Fix $T' > T$. By Skorokhod's representation theorem, there exists a probability space on which the uniform convergence

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor t\sqrt{n}\rfloor}(\widehat{D}_i^n - 1)\right)_{0\leq t\leq T'} \xrightarrow{\mathrm{d}} (\sigma^2 t)_{0\leq t\leq T'} \tag{2.4.15}$$

from Proposition 2.4.5 occurs in the almost sure sense. We work on this probability space for the rest of the proof. Note, in particular, that if the above convergence occurs almost surely then it is also the case that

$$\mathbf{1}_{[N_n \geq \lfloor T\sqrt{n}\rfloor]} \xrightarrow{\text{a.s.}} 1.$$

We first show that

$$\begin{aligned} &n^{k/2}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor, \widetilde{J}_2^n = \lfloor t_2\sqrt{n}\rfloor, \dots, \widetilde{J}_k^n = \lfloor t_k\sqrt{n}\rfloor\right\}\mathbf{1}_{[N_n \geq \lfloor T\sqrt{n}\rfloor]} \\ &\quad \xrightarrow{\text{a.s.}} \sigma^{2k} t_1 t_2 \dots t_k \exp\left(-\sigma^2 t_k^2/2\right), \end{aligned} \tag{2.4.16}$$

as $n \to \infty$.

By Lemma 2.3.1, whenever the bijective construction $B(\Pi_{D^n})$ of T_n encounters a new vertex, its degree is distributionally equivalent to the next one on the list

$(\widehat{D}_1^n, \ldots, \widehat{D}_{N_n}^n)$. So by Lemma 2.3.2, on the event $\{N_n \geq \lfloor T\sqrt{n}\rfloor\}$, we have

$$\begin{aligned}
&\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor\right\} \\
&= \frac{\sum_{\ell=1}^{\lfloor t_1\sqrt{n}\rfloor-1}(\widehat{D}_\ell^n - 1)}{n - \lfloor t_1\sqrt{n}\rfloor} \prod_{j=1}^{\lfloor t_1\sqrt{n}\rfloor-2}\left(1 - \frac{\sum_{\ell=1}^{j}(\widehat{D}_\ell^n - 1)}{n-1-j}\right) \\
&= \frac{\sum_{\ell=1}^{\lfloor t_1\sqrt{n}\rfloor-1}(\widehat{D}_\ell^n - 1)}{n - \lfloor t_1\sqrt{n}\rfloor} \exp\left(\sum_{j=1}^{\lfloor t_1\sqrt{n}\rfloor-2}\log\left(1 - \frac{\sum_{\ell=1}^{j}(\widehat{D}_\ell^n - 1)}{n-1-j}\right)\right) \\
&= \frac{\sum_{\ell=1}^{\lfloor t_1\sqrt{n}\rfloor-1}(\widehat{D}_\ell^n - 1)}{n - \lfloor t_1\sqrt{n}\rfloor} \exp\left(\sum_{j=1}^{\lfloor t_1\sqrt{n}\rfloor-2}\log\left(1 - \frac{\sum_{\ell=1}^{j}(\widehat{D}_\ell^n - 1 - \sigma^2) + \sigma^2 j}{n-1-j}\right)\right).
\end{aligned}$$

By (2.4.15) and a similar argument to that used in the proof of (2.4.10), we get that as $n \to \infty$,

$$\sqrt{n}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor\right\}\mathbf{1}_{[N_n \geq \lfloor T\sqrt{n}\rfloor]} \xrightarrow{\text{a.s.}} \sigma^2 t_1 \exp\left(-\frac{t_1^2\sigma^2}{2}\right).$$

We now proceed to prove the joint convergence of the first k coordinates in (2.4.16) by induction. Suppose that the claimed convergence holds for $\widetilde{J}_1^n, \ldots, \widetilde{J}_{m-1}^n$. By Lemma 2.3.3, on the event $\{N_n \geq \lfloor T\sqrt{n}\rfloor\}$,

$$\begin{aligned}
&\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_m^n - \widetilde{J}_{m-1}^n = \lfloor t_m\sqrt{n}\rfloor - \lfloor t_{m-1}\sqrt{n}\rfloor \;\middle|\; \widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor, \ldots, \widetilde{J}_{m-1}^n = \lfloor t_{m-1}\sqrt{n}\rfloor\right\} \\
&= \frac{\sum_{\ell=1}^{\lfloor t_m\sqrt{n}\rfloor-m}(\widehat{D}_\ell^n - 1) - m + 1}{n - \lfloor t_m\sqrt{n}\rfloor} \prod_{j=\lfloor t_{m-1}\sqrt{n}\rfloor}^{\lfloor t_m\sqrt{n}\rfloor-2}\left(1 - \frac{\sum_{\ell=1}^{j-m+1}(\widehat{D}_\ell^n - 1) - m + 1}{n-1-j}\right).
\end{aligned}$$

Arguing as above, we obtain

$$\begin{aligned}
&\sqrt{n}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_m^n = \lfloor t_m\sqrt{n}\rfloor \;\middle|\; \widetilde{J}_1^n = \lfloor t_1\sqrt{n}\rfloor, \ldots, \widetilde{J}_{m-1}^n = \lfloor t_{m-1}\sqrt{n}\rfloor\right\}\mathbf{1}_{[N_n \geq \lfloor t_m\sqrt{n}\rfloor]} \\
&\xrightarrow{\text{a.s.}} \sigma^2 t_m \exp\left(-\int_{t_{m-1}}^{t_m}\sigma^2 r dr\right).
\end{aligned}$$

By induction on m, we get this for all $1 \leq m \leq k$. Taking the product of the conditional probabilities, we obtain (2.4.16).

We now wish to add in the random variables $(A_i^n)_{i\in[k]}$. We work conditionally on the event $\mathcal{G}_{n,k}(T)$. Given also $J_1^n = \lfloor t_1\sqrt{n}\rfloor, \ldots, J_m^n = \lfloor t_m\sqrt{n}\rfloor$, $\widehat{D}_1^n, \ldots, \widehat{D}_{N_n}^n$ and $A_1^n, \ldots, A_{m-1}^n$, since $N_n \geq \lfloor T\sqrt{n}\rfloor$, at time J_m^n there are $\widehat{D}_i^n - 1 - \sum_{\ell=1}^{m-1}\mathbf{1}_{[A_\ell^n = i]}$ remaining instances of the vertex $\widehat{V}_i(\Pi_{D^n})$ to appear in the bijective construction. So, the repeated vertex that we see is $\widehat{V}_i(\Pi_{D^n})$, i.e. $A_m^n = i$, with probability

$$\frac{\widehat{D}_i^n - 1 - \sum_{\ell=1}^{m-1}\mathbf{1}_{[A_\ell^n = i]}}{\sum_{j=1}^{\lfloor t_m\sqrt{n}\rfloor-m}(\widehat{D}_i^n - 1) - m},$$

for $1 \le i \le \lfloor t_m\sqrt{n}\rfloor - m$. Hence,

$$\mathbf{P}_{\widehat{D}^n}\left\{A_m^n \le s_m\sqrt{n} \mid \mathcal{G}_{n,k}(T), J_1^n = \lfloor t_1\sqrt{n}\rfloor, \dots, J_m^n = \lfloor t_m\sqrt{n}\rfloor, A_1^n, \dots, A_{m-1}^n\right\}$$
$$= \frac{\sum_{i=1}^{\lfloor s_m\sqrt{n}\rfloor}(\widehat{D}_i^n - 1) - \sum_{\ell=1}^{m-1}\mathbf{1}_{[A_\ell^n \le s_m\sqrt{n}]}}{\sum_{j=1}^{\lfloor t_m\sqrt{n}\rfloor - m}(\widehat{D}_j^n - 1) - m}. \tag{2.4.17}$$

This quantity lies in the interval

$$\left[\frac{\sum_{i=1}^{\lfloor s_m\sqrt{n}\rfloor}(\widehat{D}_i^n - 1) - m + 1}{\sum_{j=1}^{\lfloor t_m\sqrt{n}\rfloor - m}(\widehat{D}_j^n - 1) - m}, \frac{\sum_{i=1}^{\lfloor s_m\sqrt{n}\rfloor}(\widehat{D}_i^n - 1)}{\sum_{j=1}^{\lfloor t_m\sqrt{n}\rfloor - m}(\widehat{D}_j^n - 1) - m}\right]$$

whose end-points do not depend on $A_1^n, A_2^n, \dots, A_{m-1}^n$. Iterating, we thus obtain that

$$\mathbf{P}_{\widehat{D}^n}\left\{A_1^n \le s_1\sqrt{n}, \dots, A_k^n \le s_k\sqrt{m} \;\middle|\; \mathcal{G}_{n,k}(T), J_1^n = \lfloor t_1\sqrt{n}\rfloor, \dots, J_m^n = \lfloor t_m\sqrt{n}\rfloor\right\}$$

lies in a random interval depending only on $\widehat{D}_1^n, \dots, \widehat{D}_{\lfloor t_k\sqrt{n}\rfloor}$, both of whose end-points converge almost surely to $\prod_{m=1}^k (s_m/t_m)$ by (2.4.15). So the same is true by sandwiching for our conditional probability which lies in that interval.

Putting everything together, we then have

$$\begin{aligned}
&\mathbf{P}\left\{J_1^n \le t_1\sqrt{n}, \dots, J_k^n \le t_k\sqrt{n}, A_1^n \le s_1\sqrt{n}, \dots, A_k \le s_k\sqrt{n}\right\} \\
&\quad= \mathbf{P}\left\{J_1^n \le t_1\sqrt{n}, \dots, J_k^n \le t_k\sqrt{n}, A_1^n \le s_1\sqrt{n}, \dots, A_k^n \le s_k\sqrt{n}, \mathcal{G}_{n,k}(T)^c\right\} \\
&\quad+ \mathbf{E}\left[\mathbf{P}_{\widehat{D}^n}\left\{J_1^n \le t_1\sqrt{n}, \dots, J_k^n \le t_k\sqrt{n}, A_1^n \le s_1\sqrt{n}, \dots, A_k^n \le s_k\sqrt{n}, \mathcal{G}_{n,k}(T)\right\}\right].
\end{aligned}$$

The first term on the right-hand side of this equation clearly tends to 0 by Lemma 2.4.6. Since the second is the expectation of a conditional probability, it is sufficient to show that the conditional probability itself tends to $\exp(-\sigma^2 t_k/2)\prod_{m=1}^k s_k$ in distribution. For $1 \le m \le k$ and $n \ge 1$, let us write

$$t_m^n = \frac{\lfloor t_m\sqrt{n}\rfloor + 1}{\sqrt{n}}.$$

Then we have

$$\begin{aligned}
&\mathbf{P}_{\widehat{D}^n}\left\{J_1^n \le t_1\sqrt{n}, \dots, J_k^n \le t_k\sqrt{n}, A_1^n \le s_1\sqrt{n}, \dots, A_k^n \le s_k\sqrt{n}, \mathcal{G}_{n,k}(T)\right\} \\
&= \int_0^{t_1^n}\dots\int_{r_{k-1}}^{t_k^n} \mathbf{P}_{\widehat{D}^n}\left\{A_i^n \le s_i\sqrt{n}\ \forall i \in [k] \middle| \mathcal{G}_{n,k}(T), J_1^n = \lfloor r_1\sqrt{n}\rfloor, \dots, J_k^n = \lfloor r_k\sqrt{n}\rfloor\right\} \\
&\qquad\qquad \times n^{k/2}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor r_1\sqrt{n}\rfloor, \dots, \widetilde{J}_k^n = \lfloor r_k\sqrt{n}\rfloor, \mathcal{G}_{n,k}(T)\right\} dr_k \dots dr_1 \\
&= \int_0^{t_1^n}\dots\int_{r_{k-1}}^{t_k^n} \mathbf{P}_{\widehat{D}^n}\left\{A_i^n \le s_i\sqrt{n}\ \forall i \in [k] \middle| \mathcal{G}_{n,k}(T), J_1^n = \lfloor r_1\sqrt{n}\rfloor, \dots, J_k^n = \lfloor r_k\sqrt{n}\rfloor\right\} \\
&\qquad\qquad \times n^{k/2}\mathbf{P}_{\widehat{D}^n}\left\{\widetilde{J}_1^n = \lfloor r_1\sqrt{n}\rfloor, \dots, \widetilde{J}_k^n = \lfloor r_k\sqrt{n}\rfloor\right\} \mathbf{1}_{[N_n \ge \lfloor T\sqrt{n}\rfloor]} dr_k \dots dr_1 - E_n,
\end{aligned}$$

where E_n is an error term with the property that $0 \le E_n \le \mathbf{P}_{\widehat{D}^n}\left\{\mathcal{G}_{n,k}(T)^c\right\}$ and so tends to 0 in distribution as $n \to \infty$. The first term in the product which forms the integrand tends to $\prod_{m=1}^k (s_m/r_m)$ as $n \to \infty$ and the second term tends to $\sigma^{2k} r_1 \dots r_k \exp(-\sigma^2 r_k^2/2)$, both almost surely. Write $g_n(r_1, \dots, r_k)$ for the integrand above, considered as a function of $r_1, \dots, r_k$. Then we have just shown that

$$g_n(r_1, \dots, r_k) \xrightarrow{\text{a.s.}} \sigma^{2k} \prod_{m=1}^{k} s_m \exp(-\sigma^2 r_k^2/2),$$

It is straightforward to see that this convergence is, in fact, uniform on compacts. Hence,

$$\int_0^{t_1^n} \cdots \int_{r_{k-1}}^{t_k^n} g_n(r_1, \dots, r_k) dr_k \dots dr_1 \\ \xrightarrow{\text{a.s.}} \sigma^{2k} \left(\prod_{m=1}^{k} s_m \right) \int_0^{t_1} \cdots \int_{r_{k-1}}^{t_k} \exp(-\sigma^2 r_k^2/2) dr_k \dots dr_1,$$

which yields (2.4.14). The result follows, since $T > 0$ was arbitrary. □

This completes the proof of (2.4.2) in Proposition 2.4.1.

2.4.3 Displacements at repeats

As shown above, for fixed k and large n, T_n^k is with high probability the subtree of $B(\Pi_{D^n})$ composed of the union of the paths $P^{(1)}, \dots, P^{(k)}$. Moreover, for $i \in [k]$, under the bijective construction, by Proposition 2.4.5, with the exception of the first vertex in each path $P^{(i)}$, the displacements of the vertices in $P^{(i)}$ away from their parents are asymptotically indistinguishable from IID copies of uniform entries of a $\nu_{\bar{\xi}}$ distributed displacement vector. On the other hand, the displacement away from of the first vertex in $P^{(i)}$ cannot be compared to a random variable with the same distribution as a uniform entry of a $\nu_{\bar{\xi}}$ distributed displacement vector. However, in the following lemma we will prove that such displacements are $O_{\mathbf{P}}(1)$ and so negligible on the scale of $n^{1/4}$.

We first introduce some notation. Recall that for $i \in [\ell^*(\Pi_{D^n})]$, vertex $V_{J_i^n}$ is the i-th repeated vertex encountered in the bijective construction $(B(\Pi_{D^n}), Y)$ of $\mathbf{T}_n = (\mathrm{T}_n, Y)$ (and hence a branchpoint). For $i \in [\ell^*(\Pi_{D^n})]$, let Δ_i^n be the displacement of $V_{J_i^n+1}$ away from its parent $V_{J_i^n}$ in $\mathbf{T}_n$.

Lemma 2.4.8. *For any $\ell \ge 0$, $\max\{|\Delta_1^n|, \dots, |\Delta_\ell^n|\}$ is a tight sequence of random variables for $n \ge 1$.*

Proof. We will prove that for all $\varepsilon > 0$ there exists $N > 0$ such that for all $n \geq N$, $\mathbf{P}\{|\Delta_1^n| > N\} < \varepsilon$.

To prove the result for $|\Delta_2^n|, \ldots, |\Delta_\ell^n|$, note that by Proposition 2.4.7, since $(A_i)_{i\in[k]}$ are almost surely distinct, we have

$$\mathbf{P}\left\{(A_i^n)_{i\in[k]} \text{ are distinct}\right\} \to 1$$

as $n \to \infty$. On the event $\{(A_i^n)_{i\in[k]} \text{ are distinct}\}$ the proof for $|\Delta_2^n|, \ldots, |\Delta_\ell^n|$ is analogous to that for $|\Delta_1^n|$ and so we omit it.

Recall from Proposition 2.4.7 that $\sigma n^{-1/2} J_1^n \xrightarrow{\mathrm{d}} J_1$. Recalling also that A_1^n is such that $V_{J_1^n} = \widehat{V}_{A_1^n}(\Pi_{D^n})$, it follows that conditionally on $\widehat{D}^n_{A_1^n} = k$, $\Delta_1^n \stackrel{\mathrm{d}}{=} Y_{k,U_k}$, where $U_k \stackrel{\mathrm{d}}{=} \text{Uniform}([k])$ and Y_{k,U_k} is distributed as a uniform entry of a displacement vector with law ν_k, independent of D^n. Fix $T > 0$ large. We work on the event $\{J_1^n \leq T\sqrt{n}\}$. For $N > 0$ and $K \geq 1$,

$$\begin{aligned}
&\mathbf{P}\left\{|\Delta_1^n| > N,\ J_1^n \leq T\sqrt{n}\right\} \\
&\leq \mathbf{P}\left\{\widehat{D}^n_{A_1^n} > K,\ J_1^n \leq T\sqrt{n}\right\} + \mathbf{P}\left\{|\Delta_1^n| > N,\ \widehat{D}^n_{A_1^n} \leq K,\ J_1^n \leq T\sqrt{n}\right\} \\
&\leq \mathbf{P}\left\{\widehat{D}^n_{A_1^n} > K,\ J_1^n \leq T\sqrt{n}\right\} + \sum_{k=2}^{K} \frac{k(k-1)\mu_k}{\sigma^2} \mathbf{P}\{|Y_{k,U_k}| > N\}\, \mathbf{P}\{J_1 \leq \sigma T\} \\
&+ \sum_{k=2}^{K} \left| \mathbf{P}\left\{|\Delta_1^n| > N, \widehat{D}^n_{A_1^n} = k, J_1^n \leq T\sqrt{n}\right\} - \frac{k(k-1)\mu_k}{\sigma^2} \mathbf{P}\{|Y_{k,U_k}| > N\}\, \mathbf{P}\{J_1 \leq \sigma T\} \right|.
\end{aligned}$$

We have

$$\mathbf{P}\left\{|\Delta_1^n| > N, \widehat{D}^n_{A_1^n} = k, J_1^n \leq T\sqrt{n}\right\} = \mathbf{P}\{|Y_{k,U_k}| > N\}\, \mathbf{P}\left\{\widehat{D}^n_{A_1^n} = k, J_1^n \leq T\sqrt{n}\right\}$$

and so

$$\begin{aligned}
&\mathbf{P}\left\{|\Delta_1^n| > N,\ J_1^n \leq T\sqrt{n}\right\} \\
&\leq \mathbf{P}\left\{\widehat{D}^n_{A_1^n} > K,\ J_1^n \leq T\sqrt{n}\right\} + \sum_{k=2}^{K} \frac{k(k-1)\mu_k}{\sigma^2} \mathbf{P}\{|Y_{k,U_k}| > N\} \\
&\quad + \sum_{k=2}^{K} \left| \mathbf{P}\left\{\widehat{D}^n_{A_1^n} = k,\ J_1^n \leq T\sqrt{n}\right\} - \frac{k(k-1)\mu_k}{\sigma^2} \mathbf{P}\{J_1 \leq \sigma T\} \right|. \qquad (2.4.18)
\end{aligned}$$

Since $k\mu_k \leq 1$ for all k, it follows that (2.4.18) is at most

$$\begin{aligned}
&\mathbf{P}\left\{\widehat{D}^n_{A_1^n} > K,\ J_1^n \leq T\sqrt{n}\right\} + \frac{K-1}{\sigma^2} \mathbf{P}\left\{|Y_{\bar{\xi},U_{\bar{\xi}}}| > N\right\} \\
&\quad + K \max_{2\leq k\leq K} \left| \mathbf{P}\left\{\widehat{D}^n_{A_1^n} = k,\ J_1^n \leq T\sqrt{n}\right\} - \frac{k(k-1)\mu_k}{\sigma^2} \mathbf{P}\{J_1 \leq \sigma T\} \right|. \qquad (2.4.19)
\end{aligned}$$

Fix $\varepsilon > 0$. Since $|Y_{\bar{\xi},U_{\bar{\xi}}}|$ is a random variable with support in $[0,\infty)$, we may take $M = M(K) > 0$ large enough so that

$$\frac{K-1}{\sigma^2}\mathbf{P}\left\{|Y_{\bar{\xi},U_{\bar{\xi}}}| > N\right\} < \frac{\varepsilon}{4}.$$

It remains to prove that for sufficiently large $n \geq 1$ and $K \geq 1$ the sum of the first and third terms in (2.4.19) is at most $3\varepsilon/4$. To this end, observe that for $i \geq 1$,

$$\mathbf{P}_{\widehat{D}^n}\left\{A_1^n = i,\ J_1^n \leq T\sqrt{n}\ \middle|\ J_1^n\right\} = \frac{\widehat{D}_i^n - 1}{\sum_{j=1}^{J_1^n-1}(\widehat{D}_j^n - 1)}\mathbf{1}_{[1\leq i\leq J_1^n]}\mathbf{1}_{[J_1^n\leq T\sqrt{n}]}.$$

Therefore, for any $k \geq 2$,

$$\mathbf{P}_{\widehat{D}^n}\left\{\widehat{D}_{A_1^n}^n = k,\ J_1^n \leq T\sqrt{n}\ \middle|\ J_1^n\right\} = (k-1)\frac{\left|\left\{1 \leq i \leq J_1^n\ :\ \widehat{D}_i^n = k\right\}\right|}{\sum_{j=1}^{J_1^n-1}(\widehat{D}_j^n - 1)}\mathbf{1}_{[J_1^n\leq T\sqrt{n}]}.$$

It follows that

$$\begin{aligned}
&\mathbf{P}\left\{\widehat{D}_{A_1^n}^n = k, J_1^n \leq T\sqrt{n}\right\} \\
&= (k-1)\mathbf{E}\left[\frac{\left|\left\{1 \leq i < J_1^n\ :\ \widehat{D}_i^n = k\right\}\right|}{\sum_{j=1}^{J_1^n-1}(\widehat{D}_j^n - 1)}\mathbf{1}_{[J_1^n\leq T\sqrt{n}]}\right] \\
&= (k-1)\mathbf{E}\left[\frac{\left|\left\{1 \leq i < J_1^n\ :\ \bar{\xi}_i = k\right\}\right|}{\sum_{j=1}^{J_1^n-1}(\bar{\xi}_j - 1)}\mathbf{1}_{[J_1^n\leq T\sqrt{n}]}\Theta^n\left(\bar{\xi}_1,\ldots,\bar{\xi}_{\lfloor T\sqrt{n}\rfloor}\right)\right],
\end{aligned}$$

where the final equality holds by Proposition 2.4.3.

By Proposition 2.4.7, $J_1^n = \Theta_{\mathbf{P}}(\sqrt{n})$, and so by a functional law of large numbers (see Lemma 2.8.1 in Section 2.8),

$$\frac{\left|\left\{1 \leq i < J_1^n\ :\ \bar{\xi}_i = k\right\}\right|}{\sum_{j=1}^{J_1^n-1}(\bar{\xi}_j - 1)} \xrightarrow{\mathrm{P}} \frac{k\mu_k}{\sigma^2}.$$

Combining this with Lemma 2.4.4 we obtain that as $n \to \infty$

$$\mathbf{P}\left\{\widehat{D}_{A_1^n}^n = k,\ J_1^n \leq T\sqrt{n}\right\} \to \frac{k(k-1)\mu_k}{\sigma^2}\mathbf{P}\left\{J_1 \leq \sigma T\right\}. \tag{2.4.20}$$

Since $\sum_{k=1}^{\infty} k(k-1)\mu_k/\sigma^2 = 1$, we can take $K \geq 1$ and $T > 0$ large enough so that

$$\mathbf{P}\left\{J_1 \leq \sigma T\right\}\sum_{k=2}^{K}\frac{k(k-1)\mu_k}{\sigma^2} > 1 - \frac{\varepsilon}{4}.$$

Further, by (2.4.20) we can take $n \geq 1$ large enough such that

$$\max_{2\leq k\leq K}\left|\mathbf{P}\left\{\widehat{D}_{A_1^n}^n = k,\ J_1^n \leq T\sqrt{n}\right\} - \frac{k(k-1)\mu_k}{\sigma^2}\mathbf{P}\left\{J_1 \leq \sigma T\right\}\right| < \frac{\varepsilon}{4K},$$

For such n and T, we have $\mathbf{P}\left\{\widehat{D}_{A_1^n}^n > K,\ J_1^n \leq T\sqrt{n}\right\} < \varepsilon/2$. The result follows. □

2.4.4 Convergence to the continuous line-breaking construction

We are now ready to complete the proof of Proposition 2.4.1.

Proof of Proposition 2.4.1. In view of Proposition 2.4.7, it remains to prove (2.4.3) and (2.4.4).

For (2.4.3), we recall from the discussion at the start of Section 2.4 (where $(F_1^n, \dots, F_k^n)$ were defined) that at attachment points which are first repeats, the attachment is to the left with probability 1/2 and to the right with probability 1/2. By Proposition 2.4.7, the first k attachment points are distinct and are, therefore, all first repeats with probability tending to 1 as $n \to \infty$. The statement (2.4.3) follows.

For (2.4.4), we must consider the spatial locations of the vertices along the first k paths in the bijective construction. We work on the event that the paths $P^{(1)}, \dots, P^{(k)}$ terminate in vertices $1, 2, \dots, k$ respectively, which we have already shown holds with high probability as $n \to \infty$. For the first path, we have

$$L^n(\lfloor tn^{1/2} \rfloor \wedge (J_1^n - 1)) = \sum_{j=1}^{\lfloor tn^{1/2} \rfloor \wedge (J_1^n - 1)} Y_{\widehat{D}_j^n, U_j}$$

and, for $1 \le i \le k-1$,

$$L^n((J_i^n + \lfloor tn^{1/2} \rfloor) \wedge (J_{i+1}^n - 1)) = L^n(A_i^n + i - 2) + \Delta_i^n + \sum_{j=J_i^n+1}^{(J_i^n + \lfloor tn^{1/2} \rfloor) \wedge (J_{i+1}^n - 1)} Y_{\widehat{D}_j^n, U_j}.$$

The desired convergence then follows from Proposition 2.4.5, Proposition 2.4.7 and Lemma 2.4.8. □

2.5 Tightness

We assume throughout the section that μ is critical and has finite variance $\sigma^2 \in (0, \infty)$, and that [**A1**] and [**A2**] hold.

Let $k \ge 1$. Recall that T_n^k is the subtree of T_n spanned by the root and the vertices $v_{U_1^n}, \dots v_{U_k^n} \in \mathrm{T}_n$, where $(U_1^n, \dots, U_k^n)$ is a uniformly random k-set sampled from $[n]$ and, for $i \in [n]$, v_i is the i-th vertex in the lexicographical order of T_n. In what follows we write $(U_{(1)}^{n,k}, \dots, U_{(k)}^{n,k})$ for the increasing rearrangement of $(U_1^n, \dots, U_k^n)$. Further, recall from Section 2.2 that $\mathbf{T}_n = (\mathrm{T}_n, Y)$ is the (μ, ν)-branching random walk conditioned to have size n.

Proposition 2.5.1. *For all $\gamma > 0$,*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |H_n(s) - H_n(t)| > \gamma n^{1/2} \right\} = 0 \qquad (2.5.1)$$

*and, additionally, if [**A1**] and [**A2**] hold, then*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |R_n(s) - R_n(t)| > \gamma n^{1/4} \right\} = 0. \qquad (2.5.2)$$

Under our assumptions on μ have that

$$\left(\frac{H_n(nt)}{\sqrt{n}}\right)_{0\le t\le 1} \xrightarrow{\text{d}} \left(\frac{2}{\sigma}\mathbf{e}_t\right)_{0\le t\le 1}$$

in $\mathbf{C}([0,1],\mathbb{R})$ and so (2.5.1) holds. It follows that we only need to prove (2.5.2).

Let us immediately observe that the vertices of the tree T_n either belong to T^k_n or belong to a subtree hanging off T^k_n. In Proposition 2.4.1, we showed the convergence of the spatial locations along the subtree T^k_n to those given by a Brownian motion indexed by $\mathcal{T}^k$. This has the consequence that for values $s,t \in [U^{n,k}_{(i)} - 1, U^{n,k}_{(i+1)} - 1]$ such that both corresponding vertices lie in T^k_n, we have that $|R_n(s) - R_n(t)|$ is bounded above by the maximum modulus $\Upsilon^{n,k}_i$ of an increment of the location process along the path from $U^{n,k}_{(i)}$ to $U^{n,k}_{(i+1)}$ in T^k_n. Moreover, this upper bound converges in distribution on rescaling to the analogous quantity in the limit tree, which has the same distribution as the maximum modulus Υ^k_i of an increment of β times a Brownian motion run for time D^k_i, where D^k_i is the distance between the ith and $(i+1)$st leaves of $(2/\sigma)\mathcal{T}^k$ in planar order. We thus have that

$$\max_{0\le i\le k} \Upsilon^k_i \stackrel{\text{d}}{=} \beta\sqrt{\frac{2}{\sigma}} \max_{0\le i\le k} \sup_{s,t\in[U^k_{(i)}, U^k_{(i+1)}]} |\mathbf{r}_s - \mathbf{r}_t|.$$

But

$$\max_{0\le i\le k}\left(U^k_{(i+1)} - U^k_{(i)}\right) \xrightarrow{\text{a.s.}} 0$$

as $k \to \infty$ and so, since $\mathbf{r}$ is uniformly continuous, we may deduce that for any $\gamma > 0$,

$$\lim_{k\to\infty}\lim_{n\to\infty} \mathbf{P}\left\{\max_{0\le i\le k} \Upsilon^{n,k}_i > \gamma n^{1/4}\right\} = \lim_{k\to\infty} \mathbf{P}\left\{\max_{0\le i\le k} \Upsilon^k_i > \gamma\right\} = 0. \qquad (2.5.3)$$

For values $s,t \in [U^{n,k}_{(i)}, U^{n,k}_{(i+1)}]$ for some $0 \le i \le k$ such that at least one of the corresponding vertices does not lie in T^k_n, we may bound $|R_n(s) - R_n(t)|$ by $\Upsilon^{n,k}_i$ plus twice the maximum modulus of the difference in spatial location between the parent in T^k_n of the root of a pendant subtree and some other vertex inside the tree. We have already dealt with $\Upsilon_i^{n,k}$, and so it remains to deal with the pendant subtrees. Before we can do so, we need to do some truncation of the displacements.

Fix $\gamma > 0$ and $\delta \in (0, 1/4)$. We will consider three "restrictions" of the branching random walk $\mathbf{T}_n = (\mathrm{T}_n, Y)$, which we denote by $\mathbf{T}_{n,\delta} = (\mathrm{T}_n, Y_{n,\delta})$, $\mathbf{T}_{n,\delta}^\gamma = (\mathrm{T}_n, Y_{n,\delta}^\gamma)$, and $\mathbf{T}_n^\gamma = (\mathrm{T}_n, Y_n^\gamma)$. These branching random walks capture the "typical", "mid-range", and "large" spatial displacements in $\mathbf{T}_n$.

1. **(typical displacements):** $Y_{n,\delta} = (Y_{n,\delta}^{(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ is such that for $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$,
$$Y_{n,\delta}^{(v)} = Y^{(v)} \mathbf{1}_{[\|Y^{(v)}\|_\infty \le n^{1/4-\delta}]}.$$
2. **(mid-range displacements):** $Y_{n,\delta}^\gamma = (Y_{n,\delta}^{\gamma,(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ is such that for all $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$,
$$Y_{n,\delta}^{\gamma,(v)} = Y^{(v)} \mathbf{1}_{[n^{1/4-\delta} < \|Y^{(v)}\|_\infty \le \gamma n^{1/4}]}.$$
3. **(large displacements):** $Y_n^\gamma = (Y_n^{\gamma,(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ is such that for $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$,
$$Y_n^{\gamma,(v)} = Y^{(v)} \mathbf{1}_{[\|Y^{(v)}\|_\infty > \gamma n^{1/4}]}.$$

For $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$, the vectors $Y_{n,\delta}^{(v)}, Y_{n,\delta}^{\gamma,(v)}, Y_n^{\gamma,(v)}$ are all of length $c(v, \mathrm{T}_n)$, however in what follows we will not refer to their individual entries.

Let $R_{n,\delta}$, $R_{n,\delta}^\gamma$, and R_n^γ denote the functions encoding the spatial locations of the branching random walks $\mathbf{T}_{n,\delta}$, $\mathbf{T}_{n,\delta}^\gamma$, and $\mathbf{T}_n^\gamma$, respectively. Then, for all n large enough so that $n^{1/4-\delta} \le \gamma n^{1/4}$,
$$R_n = R_{n,\delta} + R_{n,\delta}^\gamma + R_n^\gamma.$$

By the triangle inequality, for all $\gamma > 0$, we then have
$$\begin{aligned} &\max_{0 \le i \le k} \sup_{s,t \in [U_{(i)}^{n,k}-1, U_{(i+1)}^{n,k}-1]} |R_n(s) - R_n(t)| \\ &\le \max_{0 \le i \le k} \sup_{s,t \in [U_{(i)}^{n,k}-1, U_{(i+1)}^{n,k}-1]} |R_{n,\delta}(s) - R_{n,\delta}(t)| + 2\|R_{n,\delta}^\gamma\|_\infty + 2\|R_n^\gamma\|_\infty \end{aligned} \tag{2.5.4}$$

We deal with each of these three terms separately.

2.5.1 Large and mid-range displacements

Under assumption [**A2**], we show that the probability that there is a displacement in $\mathbf{T}_n$ with modulus exceeding $\gamma n^{1/4}$ goes to zero, so that the contribution of the large displacements is negligible.

Proposition 2.5.2. *For all $\gamma > 0$, as $n \to \infty$,*

$$\mathbf{P}\left\{\|R_n^\gamma\|_\infty > \gamma n^{1/4}\right\} = o(1).$$

Proof. Let

$$M_n^\gamma := \left|\left\{v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n \ : \ \|Y^{(v)}\|_\infty > \gamma n^{1/4}\right\}\right|.$$

It suffices to prove that $\mathbf{P}\{M_n^\gamma > 0\} \to 0$ as $n \to \infty$. To this end, let $\xi_1, \ldots, \xi_n$ be IID random variables with distribution μ. By assumption [**A2**],

$$\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > \gamma n^{1/4}\right\} = o(n^{-1}).$$

Fixing $\varepsilon > 0$, this implies that for n large enough,

$$\widetilde{M}_n^\gamma := \left|\left\{i \in [n] \ : \ \|Y_{\xi_i}\|_\infty > \gamma n^{1/4}\right\}\right| \preceq_{st} \mathrm{Bin}\left(n, \frac{\varepsilon}{n}\right), \tag{2.5.5}$$

where $\preceq_{st}$ denotes stochastic domination. It follows from a Chernoff bound that there exists $c > 0$ such that for n sufficiently large,

$$\mathbf{P}\left\{\widetilde{M}_n^\gamma \geq n^\varepsilon\right\} \leq \exp\left(-cn^\varepsilon\right).$$

Since $\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\} = \Theta(n^{-1/2})$, we obtain

$$\begin{aligned}
\mathbf{P}\{M_n^\gamma \geq n^\varepsilon\} &= \mathbf{P}\left\{\widetilde{M}_n^\gamma \geq n^\varepsilon \mid \sum_{i=1}^n \xi_i = n-1\right\} \\
&\leq \frac{\mathbf{P}\left\{\widetilde{M}_n^\gamma \geq n^\varepsilon\right\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}} \\
&= O\left(n^{1/2}\exp(-cn^\varepsilon)\right).
\end{aligned}$$

Let $\widetilde{S}_n^\gamma := \sum_{i=1}^n \xi_i \mathbf{1}_{[\|Y_{\xi_i}\|_\infty > \gamma n^{1/4}]}$ and let $S_n^\gamma := \sum_{v \in v(\mathrm{T}_n)} c(v, \mathrm{T}_n)\mathbf{1}_{[\|Y^{(v)}\|_\infty > \gamma n^{1/4}]}$. Since $\mathbf{E}[\xi^3] < \infty$, by [59, Corollary 19.11], both $\max_{1 \leq i \leq n} \xi_i$ and $\max_{v \in v(\mathrm{T}_n)} c(v, \mathrm{T}_n)$ are $O_{\mathbf{P}}(n^{1/3})$, and so

$$\begin{aligned}
\mathbf{P}\left\{M_n^\gamma \geq n^\varepsilon \text{ or } S_n^\gamma \geq n^{1/3+\varepsilon}\right\} &\leq o(1) + \mathbf{P}\left\{S_n^\gamma \geq n^{1/3+\varepsilon} \cap \max_{v \in v(\mathrm{T}_n)} c(v, \mathrm{T}_n) \leq n^{1/3}\right\} \\
&\leq o(1) + \mathbf{P}\left\{\sum_{v \in v(\mathrm{T}_n)} \mathbf{1}_{[\|Y^{(v)}\|_\infty \geq \gamma n^{1/4}]} > n^\varepsilon\right\} \\
&\leq o(1) + \frac{\mathbf{P}\left\{\mathrm{Bin}\left(n, \frac{\varepsilon}{n}\right) > n^\varepsilon\right\}}{\mathbf{P}\{\sum_{i=1}^n \xi_i = n-1\}} \\
&= o(1),
\end{aligned} \tag{2.5.6}$$

where the final inequality holds by (2.5.5). Further, for $\xi_1^n, \xi_2^n, \ldots$ independent random variables such that for each $i \geq 1$, ξ_i^n is distributed as ξ_i conditional on $\|Y_{\xi_i}\|_\infty < \gamma n^{1/4}$, we have that

$$\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1 \;\middle|\; \widetilde{S}_n^\gamma, \widetilde{M}_n^\gamma\right\} = \mathbf{P}\left\{\widetilde{S}_n^\gamma + \sum_{i=1}^{n-\widetilde{M}_n^\gamma} \xi_i^n = n-1 \;\middle|\; \widetilde{S}_n^\gamma, \widetilde{M}_n^\gamma\right\}.$$

Therefore,

$$\begin{aligned}
&\mathbf{P}\left\{M_n^\gamma > 0\right\} \\
&\quad= \mathbf{P}\left\{0 < M_n^\gamma < n^\varepsilon,\ S_n^\gamma < n^{1/3+\varepsilon}\right\} + o(1) \\
&\quad= \mathbf{P}\left\{0 < \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon} \;\middle|\; \sum_{i=1}^n \xi_i = n-1\right\} + o(1) \\
&\quad= \frac{\mathbf{P}\left\{0 < \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}, \sum_{i=1}^n \xi_i = n-1\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} + o(1) \\
&\quad= \mathbf{E}\left[\frac{\mathbf{P}\left\{\widetilde{S}_n^\gamma + \sum_{i=1}^{n-\widetilde{M}_n^\gamma} \xi_i^n = n-1 \;\middle|\; \widetilde{S}_n^\gamma, \widetilde{M}_n^\gamma\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \mathbf{1}_{[0<\widetilde{M}_n^\gamma<n^\varepsilon,\ \widetilde{S}_n^\gamma<n^{1/3+\varepsilon}]}\right] + o(1).
\end{aligned}$$

By a quantitative local limit theorem (see Lemma 2.8.3 in Section 2.8), we obtain that as $n \to \infty$

$$\frac{\mathbf{P}\left\{\sum_i^{n-m} \xi_i^n = n-1-s\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \to 1,$$

uniformly over all $0 < m < n^\varepsilon$ and $0 < s < n^{1/3+\varepsilon}$. It follows that

$$\mathbf{P}\left\{M_n^\gamma > 0\right\} = \mathbf{P}\left\{0 < \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}\right\} + o(1) \leq \mathbf{P}\left\{\widetilde{M}_n^\gamma > 0\right\} + o(1).$$

The result follows since for n sufficiently large $\widetilde{M}_n^\gamma \preceq_{st} \text{Bin}(n, \varepsilon/n)$, and $\varepsilon > 0$ is arbitrary. □

Similarly to the large displacements, the mid-range displacements are also negligible on the order of $n^{-1/4}$. However, the argument required to prove this is more refined.

Proposition 2.5.3. *Fix $\gamma > 0$. For $\delta > 0$ sufficiently small, as $n \to \infty$,*

$$\mathbf{P}\left\{\|R_{n,\delta}^\gamma\|_\infty > \gamma n^{1/4}\right\} = o(1).$$

To prove this proposition, we will require some further results pertaining to the positions of non-typical displacements in the branching random walk $\mathbf{T}_n$. More specifically, we will need to study the law of the *number* and *positions* of the vertices $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ such that $\|Y^{(v)}\|_\infty > n^{1/4-\delta}$, for fixed, small $\delta > 0$. The next lemma pertains to the number of such vertices.

Lemma 2.5.4. *For $\delta > 0$ sufficiently small,*

$$\left|\left\{v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n \text{ such that } \|Y^{(v)}\|_\infty > n^{1/4-\delta}\right\}\right| = o_{\mathbf{P}}(n^{1/12}).$$

Proof. Let $\xi_1, \ldots, \xi_n$ be IID with distribution μ. By [**A2**] there exists $C > 0$ such that $\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > n^{1/4-\delta}\right\} \leq Cn^{-1+4\delta}$. It follows that

$$A_n := \left|\left\{i \in [n] \ : \ \|Y_{\xi_i}\|_\infty > n^{1/4-\delta}\right\}\right| \preceq_{st} \mathrm{Bin}\left(n, Cn^{-1+4\delta}\right).$$

By a Chernoff bound, this implies that for $\delta \in (0, 1/48)$, and $n \geq 1$ sufficiently large, for any $\varepsilon > 0$,

$$\begin{aligned}\mathbf{P}\left\{A_n > \varepsilon n^{1/12}\right\} &\leq \mathbf{P}\left\{\mathrm{Bin}\left(n, Cn^{-1+4\delta}\right) > \varepsilon n^{1/12}\right\} \\ &= \mathbf{P}\left\{\mathrm{Bin}\left(n, Cn^{-1+4\delta}\right) > Cn^{4\delta}\left(1 + \left(\frac{\varepsilon}{C}n^{1/12-4\delta} - 1\right)\right)\right\} \\ &= O\left(\exp(-n^{4\delta})\right),\end{aligned}$$

so

$$\begin{aligned}\mathbf{P}\left\{A_n > \varepsilon n^{1/12} \;\middle|\; \sum_{i=1}^n \xi_i = n-1\right\} &= O\left(n^{1/2}\exp(-n^{4\delta})\right) \\ &= o(1),\end{aligned}$$

and the result follows. □

We say that two vertices $u, v \in \mathcal{U}$ are *ancestrally related*, if either $u \prec v$ or $v \prec u$. The following lemma establishes that with high probability there are no ancestrally related vertices $u, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ such that $\|Y^{(u)}\|_\infty \wedge \|Y^{(v)}\|_\infty > n^{1/4-\delta}$.

Proposition 2.5.5. *for $\delta > 0$ sufficiently small, as $n \to \infty$,*

$$\mathbf{P}\left\{\exists u, v \in \mathbf{T}_n, u \prec v, \text{ such that } \|Y^{(u)}\|_\infty \wedge \|Y^{(v)}\|_\infty > n^{1/4-\delta}\right\} = o(1).$$

The proof of this proposition relies on an application of the technical lemma, Lemma 2.3.4, which we prove in Section 2.8.

Proof. We generate T_n using the bijective construction $B(\Pi_{D^n})$ described in Section 2.2.2. Sample the displacement vectors $(Y_{D_i^n})_{1\leq i\leq n}$ with $Y_{D_i^n} = (Y_{D_i^n,1}, \ldots, Y_{D_i^n,D_i^n})$, and let

$$\mathcal{B} = \left\{i \in [n] \ : \ \|Y_{D_i^n}\|_\infty > n^{1/4-\delta}\right\}.$$

$$\begin{aligned}&\mathbf{P}\left\{\exists u, v \in \mathbf{T}_n, u \prec v, \text{ such that } \|Y^{(u)}\|_\infty \wedge \|Y^{(v)}\|_\infty > n^{1/4-\delta}\right\} \\ &\quad\leq \mathbf{P}\left\{\max_{0\leq i\leq n} H_n(i) > t\sqrt{n}\right\} + \mathbf{P}\left\{|\mathcal{B}| > sn^{1/12}\right\} + \mathbf{P}\left\{\max_{1\leq i\leq n} D_i^n > Tn^{1/3}\right\} \\ &\qquad + \mathbf{P}\left\{\left\{\max_{1\leq i\leq n} D_i^n \leq Tn^{1/3}, |\mathcal{B}| \leq sn^{1/12}\right\} \cap \left\{\exists i, j \in \mathcal{B} : i \prec j, d_n(i,j) \leq t\sqrt{n}\right\}\right\}\end{aligned} \tag{2.5.7}$$

where, for vertices $i, j \in v(\mathrm{T}_n)$, $d_n(i,j)$ denotes the length of the shortest path between i and j in $B(\Pi_{D^n}) \stackrel{\mathrm{d}}{=} \mathrm{T}_n$. Take t and T large enough so that

$$\mathbf{P}\left\{\max_{0\le i\le n} H_n(i) > t\sqrt{n}\right\} < \varepsilon/4,$$

and $\mathbf{P}\left\{\max_{1\le i\le n} D_i^n > Tn^{1/3}\right\} < \varepsilon/4$. (The latter inequality is possible by [59, Corollary 19.11] since $\mathbf{E}\left[\xi^3\right] < \infty$.) By Lemma 2.5.4, we may take n large enough so that $\mathbf{P}\left\{|\mathcal{B}| > sn^{1/12}\right\} < \frac{\varepsilon}{4}$. Therefore, for t, T and n sufficiently large (2.5.7) is at most

$$\frac{3\varepsilon}{4} + \mathbf{P}\left\{\left\{\max_{1\le i\le n} D_i^n \le Tn^{1/3}, |\mathcal{B}| \le sn^{1/12}\right\} \cap \left\{\exists i, j \in \mathcal{B} : i \prec j, d_n(i,j) \le t\sqrt{n}\right\}\right\}$$

Then by Lemma 2.3.4 with $\mathrm{d} = D^n$, $K \le sn^{1/12}$, $\Delta \le Tn^{1/3}$, and $b = t\sqrt{n}$, for n sufficiently large, this is at most

$$\frac{3\varepsilon}{4} + sn^{1/12}\left(1 - \left(1 - \frac{sTn^{-7/12}}{1 - n^{-1} - tTn^{-1/6}}\right)^{t\sqrt{n}}\right).$$

The result follows by taking $s > 0$ small enough and n large enough so that

$$sn^{1/12}\left(1 - \left(1 - \frac{sTn^{-7/12}}{1 - n^{-1} - tTn^{-1/6}}\right)^{t\sqrt{n}}\right) < \frac{\varepsilon}{4},$$

which is possible since

$$sn^{1/12}\left(1 - \left(1 - \frac{sTn^{-7/12}}{1 - n^{-1} - tTn^{-1/6}}\right)^{t\sqrt{n}}\right) < s^2Tt\frac{1}{1 - n^{-1} - tTn^{-1/6}},$$

for n large enough because $(1-x)^r > 1 - rx$ for $x < 1$ and $r > 1$. □

Lemma 2.5.6. *Let $v^*(\mathrm{T}_n) \subseteq v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ be the set of vertices $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ such that $\|Y^{(v)}\|_\infty \le n^{1/4-\delta}$ and there exists an ancestor $u \prec v$ with $\|Y^{(u)}\|_\infty > n^{1/4-\delta}$. For $\delta > 0$ sufficiently small, $v^*(\mathrm{T}_n) = o_{\mathbf{P}}(n)$.*

Proof. The result holds if and only if the probability that a uniformly random vertex in $v \in v(\mathrm{T}_n)$ is ancestrally related to a vertex $u \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ with $\|Y^{(u)}\|_\infty > n^{1/4-\delta}$ is $o_{\mathbf{P}}(1)$. By exchangeability, this holds if and only if the probability that vertex 1 is ancestrally related to a vertex $u \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ with $\|Y^{(u)}\|_\infty > n^{1/4-\delta}$ is $o_{\mathbf{P}}(1)$. To prove this we may adapt the proof of Proposition 2.5.5 by including vertex 1 in the set $\mathcal{B}$. Then by Lemma 2.5.4, $|\mathcal{B}| = o_{\mathbf{P}}(n^{1/12})$ still holds and so the proof carries over verbatim. □

As an immediate consequence of Lemma 2.5.6, with probability $1 - o(1)$ none of the increments of the branching random walk $\mathbf{T}_{n,\delta}^{\gamma}$ are ancestrally related with high probability, and Proposition 2.5.3 follows.

2.5.2 Typical displacements

In this subsection we will prove the following proposition.

Proposition 2.5.7. *For all $\gamma > 0$,*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |R_{n,\delta}(s) - R_{n,\delta}(t)| > \gamma n^{1/4} \right\} = 0.$$

Notice that $R_{n,\delta}$ is equal in distribution to the function encoding the spatial locations of the branching random walk with underlying tree T_n and displacements $Y^{n,\delta} = (Y^{n,\delta,(v)}, v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n)$ such that if $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ has k children, then $Y^{n,\delta,(v)}$ has the same distribution as

$$Y^{n,\delta}_k = (Y^{n,\delta}_{k,1}, \dots Y^{n,\delta}_{k,k}) := \begin{cases} (Y_{k,1}, \dots, Y_{k,k}) & \text{if } \max_{1\le j\le k} |Y_{k,j}| \le n^{1/4-\delta} \\ (0, \dots, 0) & \text{else.} \end{cases}$$

This branching random walk is *not* globally centered, and in particular has "global" drift

$$\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right].$$

Thus for all $t \in [0, n]$ we have that

$$R_{n,\delta}(t) \stackrel{\mathrm{d}}{=} \breve{R}_{n,\delta}(t) + \mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right] \cdot H_n(t).$$

where $\breve{R}_{n,\delta} : [0, n] \to \mathbb{R}$ is the function encoding the spatial locations of the globally centered branching random walk $(\mathrm{T}_n, \breve{Y}^{n,\delta})$, where conditionally on T_n, $\breve{Y}^{n,\delta} = (\breve{Y}^{n,\delta,(v)}, v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n)$ is a vector of independent random variables, such that if $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ has k children then $\breve{Y}^{n,\delta,(v)}$ has the same distribution as

$$\breve{Y}^{n,\delta}_k := Y^{n,\delta}_k - \mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right] = Y^{n,\delta}_k - \mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}} \mathbf{1}_{[\|Y_{\bar{\xi}}\|_\infty \le n^{1/4-\delta}]}\right].$$

Moreover, by the triangle inequality, for all $\gamma > 0$,

$$\begin{aligned} &\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |R_{n,\delta}(s) - R_{n,\delta}(t)| > \gamma n^{1/4} \right\} \\ &\le \lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1, U^{n,k}_{(i+1)}-1]} |\breve{R}_{n,\delta}(s) - \breve{R}_{n,\delta}(t)| > \frac{\gamma}{2} n^{1/4} \right\} \\ &\quad + \limsup_{n\to\infty} \mathbf{P}\left\{ \left|\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right]\right| \cdot \|H_n\|_\infty > \frac{\gamma}{4} n^{1/4} \right\}. \end{aligned}$$

Lemma 2.5.8. *It holds that*

$$\left|\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right]\right| = O(n^{-5/12+5\delta/3}).$$

and furthermore, as $n \to \infty$,

$$\mathbf{Var}\left\{Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right\} \to \beta^2.$$

This result is a special case of Lemma 2.8.11, which is stated and proved in Section 2.8. Since $\|H_n\|_\infty = O_{\mathbf{P}}(\sqrt{n})$, Lemma 2.5.8 implies that for δ sufficiently small,

$$\limsup_{n\to\infty} \mathbf{P}\left\{\left|\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right]\right| \cdot \|H_n\|_\infty > \gamma n^{1/4}\right\} = 0.$$

It follows that to prove Proposition 2.5.7, it suffices to prove that for all $\gamma > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{\max_{0\le i\le k} \sup_{s,t\in[U^{n,k}_{(i)}-1,U^{n,k}_{(i+1)}-1]} |\breve{R}_{n,\delta}(s) - \breve{R}_{n,\delta}(t)| > \gamma n^{1/4}\right\} = 0.$$

As discussed above, we need to deal with the maximum modulus of the difference in spatial location (for the branching random walk $\mathrm{T}_{n,\delta}$) between the parent of the root of a pendant subtree and a vertex of that subtree. There are

$$c(\mathrm{T}^k_n) := \sum_{v\in V(\mathrm{T}^k_n)} (c(v, T_n) - 1) + 1$$

edges in T_n with one endpoint in T^k_n and another in $\mathrm{T}_n \setminus \mathrm{T}^k_n$. Conditionally on T^k_n, if we remove all such edges we obtain a Bienaymé(μ) forest conditioned to have $n - |V(\mathrm{T}^{(k)}_n)|$ vertices and $c(\mathrm{T}^k_n)$ trees. We denote this forest by $\mathrm{F}^k_n = (\mathrm{T}^k_{n,j})_{j\ge1}$, where the trees are listed in decreasing order of size, and $|\mathrm{T}^k_{n,j}| = 0$ for $j > c(\mathrm{T}^k_n)$. Write $\|\breve{R}_{n,\delta}(\mathrm{T}^k_{n,j})\|_\infty$ for maximum modulus of the difference in spatial location between the root and any other vertex of $\mathrm{T}^k_{n,j}$.

The trees $(\mathrm{T}^k_{n,j})_{j\ge1}$ are independent Bienaymé trees, conditioned on their sizes. Therefore, conditionally on F^k_n, we have $\|\breve{R}_{n,\delta}(\mathrm{T}^k_{n,j})\|_\infty \overset{\mathrm{d}}{=} \|\breve{R}_{|\mathrm{T}^k_{n,j}|,\delta}\|_\infty$. Moreover, displacements on the tree $\mathrm{T}^k_{n,j}$ (from the branching random walk $\mathbf{T}_{n,\delta}$) depend on those in other parts of T_n only through the displacement $\breve{Z}^{n,\delta}_j$ of the root of $\mathrm{T}^k_{n,j}$ away from its parent in T^k_n; see Figure 2.5.

It follows that

$$\begin{aligned}\max_{0\le i\le k} &\sup_{s,t\in[U^{n,k}_{(i)}-1,U^{n,k}_{(i+1)}-1]} |\breve{R}_{n,\delta}(s) - \breve{R}_{n,\delta}(t)| \\ &\le \max_{0\le i\le k} \Upsilon^{n,k}_i + 2 \max_{1\le j\le c(\mathrm{F}^k_n)} \left(\left\|\breve{R}_{n,\delta}(\mathrm{T}^k_{n,j})\right\|_\infty + |\breve{Z}^{n,\delta}_j|\right)\end{aligned}$$

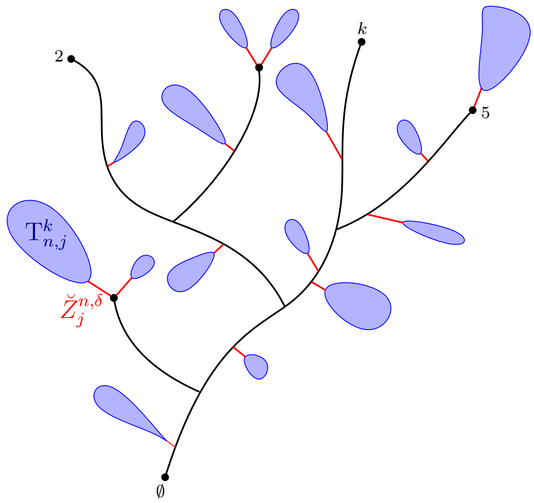

Figure 2.5: In black, the tree T^k_n. In blue, the forest $\mathrm{F}^k_n = (\mathrm{T}^k_{n,j})_{j\geq 1}$. The root of tree $\mathrm{T}^k_{n,j}$ is displaced $\breve{Z}^{n,\delta}_j$ away from its parent in T^k_n.

Consequently, using (2.5.3), in order to prove Proposition 2.5.7, it is sufficient to prove that for $\gamma > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{1\leq j\leq c(\mathrm{F}^k_n)} \left(\left\| \breve{R}_{n,\delta}(\mathrm{T}^k_{n,j}) \right\|_\infty + |\breve{Z}^{n,\delta}_j| \right) \geq \gamma n^{1/4} \right\} = 0$$

The proof requires two key ingredients: (1) a scaling limit for the sizes of the trees in F^k_n; (2) quantitative control on the tail of $\|\breve{R}_{n,\delta}\|_\infty$. We begin by establishing (1).

Proposition 2.5.9. *As $n \to \infty$,*

$$\frac{c(\mathrm{T}^k_n)}{\sigma\sqrt{n}} \xrightarrow{\mathrm{d}} J_k, \tag{2.5.8}$$

where J_k is $\mathrm{Gamma}(k, 1/2)$ *distributed. Jointly with this convergence, we have*

$$\frac{\sigma}{n - |V(\mathrm{T}^k_n)|}(|\mathrm{T}^k_{n,j}|, j \geq 1) \xrightarrow{\mathrm{d}} (|\gamma^k_j|, j \geq 1), \tag{2.5.9}$$

where, conditionally on J_k, $(|\gamma^k_j|, j \geq 1)$ lists the sizes of the excursions above the past minimum of a Brownian motion stopped on first hitting $-J_k$, listed in decreasing order.

Proof. By Skorokhod's representation theorem we may work in a probability space where the convergence in Proposition 2.4.7 holds almost surely so that in particular $\sigma n^{-1/2} J^n_k \xrightarrow{\mathrm{a.s.}} J_k$.

Let $T > 0$ and recall the event

$$\mathcal{G}_{n,k}(T) = \left\{ \left\{ \widehat{V}_1(\Pi_{D^n}), \ldots, \widehat{V}_{\lfloor T\sqrt{n}\rfloor}(\Pi_{D^n}) \right\} \cap \{1, \ldots, k\} = \emptyset, N_n \geq \lfloor T\sqrt{n} \rfloor \right\}$$

from Lemma 2.4.6. On $\mathcal{G}_{n,k}(T) \cap \{J_k^n \le \lfloor T\sqrt{n} \rfloor\}$, the tree T_n^k is precisely the subtree of T_n spanned by the root and the vertices $1, \ldots, k$. Therefore, on $\mathcal{G}_{n,k}(T) \cap \{J_k^n \le \lfloor T\sqrt{n} \rfloor\}$,

$$\frac{V(\mathrm{T}_n^k)}{\sqrt{n}} = \frac{J_k^n}{\sqrt{n}}.$$

Since $T > 0$ is arbitrary and $\sigma n^{-1/2} J_k^n \xrightarrow{\text{a.s.}} J_k$ we obtain that $n - |V(\mathrm{T}_n^k)| = n - o_{\mathbf{P}}(n)$. Hence, we are essentially considering a forest of Bienaymé trees conditioned to have n vertices. We now need to show that the number of trees in such a forest is $\sim \sigma\sqrt{n} J_k$. We note that on the event $\mathcal{G}_{n,k}(T) \cap \{J_k^n \le T\sqrt{n}\}$, there are $\sum_{i=1}^{J_k^n - k} (\widehat{D}_i^n - 1) + \sum_{i=1}^k D_{J_i^n}^n$ subtrees of T_n whose root has a parent in T_n^k, and $(k-1)$ branch points in T_n^k. Therefore for $s \ge 0$

$$\begin{aligned} &\mathbf{P}\left\{ \frac{c(\mathrm{T}_n^k)}{\sigma\sqrt{n}} \ge s,\ \mathcal{G}_{n,k}(T),\ J_k^n \le T\sqrt{n} \right\} \\ &\quad = \mathbf{P}\left\{ \frac{1}{\sigma\sqrt{n}} \left(\sum_{i=1}^{J_k^n - k} (\widehat{D}_i^n - 1) + \sum_{i=1}^{k} D_{J_i^n}^n - (k-1) \right) \ge s,\ \mathcal{G}_{n,k}(T),\ J_k^n \le T\sqrt{n} \right\}. \end{aligned} \tag{2.5.10}$$

Since, $\sigma n^{-1/2} J_k^n \xrightarrow{\text{a.s.}} J_k$, by Proposition 2.4.5

$$\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{J_k^n - k} (\widehat{D}_i - 1) \xrightarrow{\text{d}} J_k.$$

Combining this with Lemma 2.4.6 and Proposition 2.4.7 we obtain that (2.5.10) converges to

$$\mathbf{P}\left\{ J_k > s, J_k \le \sigma T \right\}.$$

Then, (2.5.8) follows as $T > 0$ is arbitrary. The scaling limit in (2.5.9) now follows from [78, Proposition 1.4] and [23, Lemma 11]. □

The control on $\|\breve{R}_{n,\delta}\|_\infty$ needed to prove Proposition 2.5.7 is given by the next proposition.

Proposition 2.5.10. *There exists $A > 0$ such that for all $\gamma > 0$, $\delta \in (0, 1/4)$, and $n \ge 1$,*

$$\mathbf{P}\left\{ \|\breve{R}_{n,\delta}\|_\infty > \gamma n^{1/4} \right\} \le \frac{A}{\gamma^8}.$$

The proof of this proposition is long and somewhat technical, so we postpone it until Section 2.6.

Proof of Proposition 2.5.7 assuming Proposition 2.5.10. By Skorohod's representation theorem we may assume that we are working on a probability space where the convergence in Proposition 2.4.7 is almost sure. In particular, $\sigma n^{-1/2} J_k^n \xrightarrow{\text{a.s.}} J_k$ as $n \to \infty$.

As argued above, it remains to show that, for $\gamma > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{1\le j\le c(\mathrm{F}_n^k)} \left(\left\| \breve{R}_{n,\delta}(\mathrm{T}_{n,j}^k) \right\|_\infty + |\breve{Z}_j^{n,\delta}| \right) \ge \gamma n^{1/4} \right\} = 0.$$

Since $(\mathrm{T}_{n,j}^k)_{1\le j\le c(\mathrm{F}_n^k)}$ are independent Bienaymé(μ) trees conditionally on their sizes, we obtain

$$\begin{aligned}
&\mathbf{P}\left\{ \max_{1\le j\le c(\mathrm{F}_n^k)} \left(\left\| \breve{R}_{n,\delta}(\mathrm{T}_{n,j}^k) \right\|_\infty + |\breve{Z}_j^{n,\delta}| \right) \ge \gamma n^{1/4} \right\} \\
&= \mathbf{E}\left[\mathbf{P}\left\{ \max_{1\le j\le c(\mathrm{F}_n^k)} \left(\left\| \breve{R}_{|\mathrm{T}_{n,j}^k|,\delta} \right\|_\infty + |\breve{Z}_j^{n,\delta}| \right) \ge \gamma n^{1/4} \mid \mathrm{F}_n^k, (\breve{Z}_j^{n,\delta})_{j\ge1} \right\} \right] \\
&\le \mathbf{E}\left[\sum_{j=1}^{c(\mathrm{F}_n^k)} \mathbf{P}\left\{ \left\| \breve{R}_{|\mathrm{T}_{n,j}^k|,\delta} \right\|_\infty \ge n^{1/4}(\gamma - |\breve{Z}_j^{n,\delta}|/n^{1/4}) \mid \mathrm{F}_n^k, (\breve{Z}_j^{n,\delta})_{j\ge1} \right\} \right] \\
&\le \mathbf{E}\left[\sum_{j=1}^{\infty} \mathbf{P}\left\{ \| \breve{R}_{|\mathrm{T}_{n,j}^k|} \|_\infty \ge \frac{\gamma n^{1/4}}{2} \,\middle|\, \mathrm{F}_n^k \right\} \right],
\end{aligned}$$

for all n sufficiently large, since $|\breve{Z}_j^{n,\delta}| \le n^{1/4-\delta}$ for all $1 \le j \le c(\mathrm{T}_k^n)$ and all $n \ge 1$. Applying Proposition 2.5.10 to each of the conditional probabilities in the above sum, we obtain that the right-hand side is at most

$$\frac{2^8 A}{\gamma^8} \mathbf{E}\left[\sum_{j=1}^{\infty} \left(\frac{|\mathrm{T}_{n,j}^k|}{n} \right)^2 \right] = \frac{2^8 A}{\gamma^8} \mathbf{E}\left[\left(\frac{n - |V(\mathrm{T}_n^k)|}{n} \right)^2 \cdot \frac{|\widehat{\mathrm{T}^k}|}{n - |V(\mathrm{T}_n^k)|} \right], \tag{2.5.11}$$

where $|\widehat{\mathrm{T}^k}|$ is a size-biased pick from $(|\mathrm{T}_{n,j}^k|)_{j\ge1}$. Clearly,

$$\frac{n - |V(\mathrm{T}_n^k)|}{n} \le 1.$$

By Proposition 2.5.9, as $n \to \infty$, $|\widehat{\mathrm{T}^k}|/(n - |V(\mathrm{T}_n^k)|) \xrightarrow{\mathrm{d}} \sigma^{-1}|\widehat{\gamma^k}|$ where $|\widehat{\gamma^k}|$ is a size-biased pick from $(|\gamma_j^k|, j \ge 1)$. By [95, Section 8.1], conditionally on J_k,

$$|\widehat{\gamma^k}| \overset{\mathrm{d}}{=} \frac{B^2}{J_k^2 + B^2},$$

where B is a $N(0,1)$ random variable independent of J_k. Combining this with (2.5.11), we obtain that

$$\limsup_{n\to\infty} \mathbf{P}\left\{ \max_{1\le j\le c(\mathrm{F}_n^k)} \left(\left\| \breve{R}_{n,\delta}(\mathrm{T}_{n,j}^k) \right\|_\infty + |\breve{Z}_j^{n,\delta}| \right) \ge \gamma n^{1/4} \right\} \le \frac{2^8 A}{\sigma \gamma^8} \mathbf{E}\left[\frac{B^2}{J_k^2 + B^2} \right].$$

As $k \to \infty$, $J_k \xrightarrow{\mathrm{p}} \infty$. Therefore, by bounded convergence,

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbf{P}\left\{ \max_{1\le j\le c(\mathrm{F}_n^k)} \left(\left\| \breve{R}_{n,\delta}(\mathrm{T}_{n,j}^k) \right\|_\infty + |\breve{Z}_j^{n,\delta}| \right) \ge \gamma n^{1/4} \right\} = 0,$$

and the result follows. □

Assuming Proposition 2.5.10, Proposition 2.5.1 now follows from (2.5.4) by taking $\delta \in (0, 1/4)$ sufficiently small so that Proposition 2.5.3 holds and combining that with Propositions 2.5.2 and 2.5.7.

2.6 The maximum spatial location: proof of Proposition 2.5.10

We assume throughout this section that μ is critical and has finite variance $\sigma^2 \in (0, \infty)$, and that [**A1**] and [**A2**] hold.

For $n \ge 1$, let $\Lambda^{(n)} := (\Lambda_1^{(n)}, \Lambda_2^{(n)}, \ldots, \Lambda_{\widehat{D}_1^n}^{(n)})$ be the sizes of the subtrees of the root of T_n, so that $\Lambda_i^{(n)}$ is the size of the subtree rooted at the i-th child of the root. We will make extensive use of the fact that, conditionally on $\widehat{D}_1^n$, these are exchangeable random variables (i.e. their distribution is invariant under permutations of the labels). To prove Proposition 2.5.10 we will make extensive use of the following consequence of Lemma 26 of Haas and Miermont [50] which, roughly speaking, tells us that typically only one subtree of a child of the root is macroscopic and, moreover, the probability of a non-trivial macroscopic split at the root is on the order of $n^{-1/2}$.

Lemma 2.6.1 (Lemma 26 of Haas and Miermont [50])**.** *It holds that*

$$\mathbf{E}\left[1 - \sum_{i=1}^{\widehat{D}_1^n} \left(\frac{\Lambda_i^{(n)}}{n} \right)^2 \right] = \Theta(n^{-1/2}). \tag{2.6.1}$$

In the proof of Proposition 2.5.10, we encounter terms directly related to the global centering and global finite variance conditions, respectively. The latter is more challenging to control, and is the reason for the third moment condition on the offspring distribution. These terms, and the control we will require on them are stated in the following technical lemma. Recall the definition of $\widehat{D}^m$, the size-biased ordering of $D^m = (D_1^m, \ldots, D_m^m)$, IID samples random distribution μ conditioned to sum to $m - 1$.

The proof of Proposition 2.5.10 is inductive, and requires that we control the maximum of $\breve{R}_{n,\delta}^k$ when restricted to subtrees of T_n^k. We henceforth use $m \ge 1$ to denote the number of vertices in the underlying tree, T_m, and $n \ge 1$ to denote the truncation threshold $n^{1/4-\delta}$ on the displacements. More specifically, in this section, we will consider branching random walks on T_m with displacements $\breve{Y}_k^{n,\delta}$, $k \geq 1$.

Lemma 2.6.2. *Let $n \geq 1$ and $m \leq n$. There exists $B > 0$ such that*

$$\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m}\right)^2 (\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i})^2\right] \leq B. \tag{2.6.2}$$

*If in addition (μ, ν) satisfies [**A1**] and [**A2**] then there exists $B' > 0$ such that*

$$\left|\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m} \left(\frac{\Lambda_i^{(m)}}{m}\right)^2 \breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right]\right| \leq \frac{B' n^{1/4-\delta}}{\sqrt{m}} + \frac{B'}{m^{1/4}} \tag{2.6.3}$$

Condition [**A1**] pertains to the mean and variance of the displacement of a uniform child of a vertex with a size-biased number of offspring, $Y_{\bar{\xi},U_{\bar{\xi}}}$. The displacement from the root of T_m to a uniform child is distributed as $Y_{\widehat{D}_1^m,U_{\widehat{D}_1^m}}$ and we have $\widehat{D}_1^m \xrightarrow{\mathrm{d}} \bar{\xi}$ as $m \to \infty$. However, in order to use the global centering and global finite variance conditions in the proof of Lemma 2.6.2, we need something stronger, namely an explicit rate of decay for the total variation distance between the laws of $\bar{\xi}$ and $\widehat{D}_1^m$. This is provided by the next lemma.

Lemma 2.6.3. *As $m \to \infty$,*

$$d_{\mathrm{TV}}(\widehat{D}_1^m, \bar{\xi}) = \frac{1}{2}\sum_{k=1}^{\infty} \left|\mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\}\right| = o(m^{-1/2}).$$

Proof. Let $k \geq 1$, and let $(S_m)_{m\geq 1}$ be a random walk with IID μ-distributed increments. Recall from (2.1.14) that

$$\mathbf{P}\left\{\widehat{D}_1^m = k\right\} = \left(\frac{m}{m-1}\right) \frac{\mathbf{P}\left\{S_{m-1} = m-1-k\right\}}{\mathbf{P}\left\{S_m = m-1\right\}} \mathbf{P}\left\{\bar{\xi} = k\right\}.$$

Since $\mathbf{E}[\xi] = 1$ and $\mathbf{E}[\xi^3] < \infty$, by Theorem 2.8.2,

$$\begin{aligned}&\sqrt{2\pi(m-1)}\sigma\mathbf{P}\left\{S_{m-1} = m-1-k\right\} \\ &\quad = e^{-k^2/(2\sigma^2(m-1))}\left(1 + \frac{1}{\sqrt{m-1}}\frac{\gamma_3}{6\sigma^3}\left(\frac{k^3}{\sigma^3(m-1)^{3/2}} - \frac{3k}{\sigma\sqrt{m-1}}\right)\right) + o(m^{-1/2}).\end{aligned}$$

If $k = O(m^{1/4})$,

$$\left|\frac{k^3}{\sigma^3(m-1)^{3/2}} - \frac{3k}{\sigma\sqrt{m-1}}\right| = O(m^{-1/4}),$$

and

$$e^{-k^2/(2\sigma^2(m-1))} = 1 - \frac{k^2}{2\sigma^2(m-1)} + O(m^{-1}).$$

Hence for $k = O(m^{1/4})$,

$$\sqrt{2\pi(m-1)}\sigma\mathbf{P}\left\{S_{m-1} = m-1-k\right\} = 1 - \frac{k^2}{2\sigma^2(m-1)} + o(m^{-1/2}).$$

It follows that for $k = O(m^{1/4})$,

$$\begin{aligned}\left(\frac{m}{m-1}\right)\frac{\mathbf{P}\left\{S_{m-1} = m-1-k\right\}}{\mathbf{P}\left\{S_m = m-1\right\}} &= \frac{1 - \frac{k^2}{2\sigma^2(m-1)} + o(m^{-1/2})}{1 + o(m^{-1/2})} \\ &= 1 - \frac{k^2}{2\sigma^2(m-1)} + o(m^{-1/2}),\end{aligned}$$

and, consequently,

$$\mathbf{P}\left\{\widehat{D}_1^m = k\right\} = \left(1 - \frac{k^2}{2\sigma^2(m-1)} + o(m^{-1/2})\right)\mathbf{P}\left\{\bar{\xi} = k\right\}.$$

Therefore,

$$\begin{aligned}&\sum_{k=1}^{\infty}\left|\mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\}\right| \\ &= \sum_{k=1}^{\lfloor m^{1/4}\rfloor}\left|\mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\}\right| + \sum_{k=\lfloor m^{1/4}\rfloor+1}^{\infty}\left|\mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\}\right| \\ &= \sum_{k=1}^{\lfloor m^{1/4}\rfloor}\left(\frac{k^2}{2\sigma^2(m-1)} + o(m^{-1/2})\right)\mathbf{P}\left\{\bar{\xi} = k\right\} + \sum_{k=\lfloor m^{1/4}\rfloor+1}^{\infty}\left|\mathbf{P}\left\{\widehat{D}_1^m = k\right\} - \mathbf{P}\left\{\bar{\xi} = k\right\}\right| \\ &\leq \frac{\mathbf{E}\left[\xi^3\right]}{2\sigma^2(m-1)} + o(m^{-1/2}) + \sum_{k=\lfloor m^{1/4}\rfloor+1}^{\infty}\left(\mathbf{P}\left\{\widehat{D}_1^m = k\right\} + \mathbf{P}\left\{\bar{\xi} = k\right\}\right) \\ &\leq o(m^{-1/2}) + (c+1)\cdot\mathbf{P}\left\{\bar{\xi} > m^{1/4}\right\},\end{aligned}$$

where the final inequality follows since $\mathbf{P}\left\{\widehat{D}_1^m = k\right\} \leq c\mathbf{P}\left\{\bar{\xi} = k\right\}$ for all $k \in [m]$. Since $\mathbf{E}\left[\xi^3\right] < \infty$, $\bar{\xi}$ has a finite second moment. Therefore, $\mathbf{P}\left\{\bar{\xi} > k\right\} = o(k^{-2})$ as $k \to \infty$ and so $\mathbf{P}\left\{\bar{\xi} > m^{1/4}\right\} = o(m^{-1/2})$. The result follows. □

The terms (2.6.2) and (2.6.3) relate to the variance and mean (respectively) of the displacement of a uniform child of the root in branching random walk $(\mathrm{T}_m, \breve{Y}^{n,\delta})$. Since this branching random walk is globally centered, it is reasonable to expect that the mean will be small and that the second moment will be bounded. A key technical lemma follows.

Lemma 2.6.4. *There exists a constant $C > 0$ such that for $m \leq n$,*

$$\left|\mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right]\right| \leq \frac{Cn^{1/4-\delta}}{\sqrt{m}}.$$

Proof. Let $(\widehat{D}_1^m, \bar{\xi})$ be a coupling of the degree of the root of T_m and the size-biased distribution of μ. We consider the events $\{\bar{\xi} = \widehat{D}_1^m\}$ and $\{\bar{\xi} \neq \widehat{D}_1^m\}$ separately:

$$\left|\mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m} \breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right]\right| \leq \left|\mathbf{E}\left[\frac{1}{\bar{\xi}}\sum_{i=1}^{\bar{\xi}} \breve{Y}^{n,\delta}_{\bar{\xi},i}\mathbf{1}_{[\bar{\xi}=\widehat{D}_1^m]}\right]\right| + \mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m} |\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}|\mathbf{1}_{[\bar{\xi}\neq\widehat{D}_1^m]}\right]$$

$$= \left|\mathbf{E}\left[\frac{1}{\bar{\xi}}\sum_{i=1}^{\bar{\xi}} \breve{Y}^{n,\delta}_{\bar{\xi},i}\mathbf{1}_{[\bar{\xi}\neq\widehat{D}_1^m]}\right]\right| + \mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m} |\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}|\mathbf{1}_{[\bar{\xi}\neq\widehat{D}_1^m]}\right] \quad (2.6.4)$$

where the equality holds since

$$\mathbf{E}\left[\frac{1}{\bar{\xi}}\sum_{i=1}^{\bar{\xi}} \breve{Y}^{n,\delta}_{\bar{\xi},i}\right] = \mathbf{E}\left[\breve{Y}^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right] = 0.$$

Since $|\breve{Y}^{n,\delta}_{k,j}| \leq 2n^{1/4-\delta}$ for all $k \geq 1$ and $j \in [k]$ it follows that (2.6.4) is at most $4n^{1/4-\delta}\mathbf{P}\left\{\bar{\xi} \neq \widehat{D}_1^m\right\}$. The result follows from Lemma 2.6.3 by taking an optimal coupling of $(\widehat{D}_1^m, \bar{\xi})$. □

We now proceed to prove Lemma 2.6.2.

Proof of Lemma 2.6.2. We first prove (2.6.2). Note that by exchangeability of $(\Delta_1^{(m)}, \ldots, \Delta^{(m)}_{\widehat{D}^1_m})$ and linearity of conditional expectation

$$\begin{aligned}
\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2 (\breve{Y}^{n,\delta}_{\widehat{D}^1_m,i})^2\right] &= \mathbf{E}\left[\left(\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}(\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i})^2\right)\left(\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)\right] \\
&\leq \mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}(\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i})^2\right] \\
&\leq c\mathbf{E}\left[\frac{1}{\bar{\xi}}\sum_{i=1}^{\bar{\xi}}(\breve{Y}^{n,\delta}_{\bar{\xi},i})^2\right] \\
&= c\mathbf{E}\left[(\breve{Y}^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}})^2\right], \qquad (2.6.5)
\end{aligned}$$

where the second inequality follows since $\mathbf{P}\left\{\widehat{D}_1^m = k\right\} \leq c\mathbf{P}\left\{\bar{\xi} = k\right\}$. By Lemma 2.5.8, (2.6.5) tends to β^2 as $n \to \infty$ and hence (2.6.2) holds.

We now proceed to proving (2.6.3). By linearity and the triangle inequality we have

$$\left|\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2 \breve{Y}^{n,\delta}_{\widehat{D}^1_m,i}\right]\right|$$

$$\leq \left|\mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right]\right| + \left|\mathbf{E}\left[\left(\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m}\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right)\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)\right]\right|. \quad (2.6.6)$$

By Lemma 2.6.4, (2.6.6) is at most

$$\frac{Cn^{1/4-\delta}}{\sqrt{m}} + \left|\mathbf{E}\left[\left(\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m} \breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right)\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)\right]\right|.$$

Applying the Cauchy–Schwarz inequality to the second term yields an upper bound of

$$\begin{aligned}
&\frac{Cn^{1/4-\delta}}{\sqrt{m}} + \mathbf{E}\left[\left(\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m} |\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}|\right)^2\right]^{1/2}\mathbf{E}\left[\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)^2\right]^{1/2} \\
&\le \frac{Cn^{1/4-\delta}}{\sqrt{m}} + \mathbf{E}\left[\frac{1}{\widehat{D}_1^m}\sum_{i=1}^{\widehat{D}_1^m} (\breve{Y}^{n,\delta}_{\widehat{D}_1^m,i})^2\right]^{1/2}\mathbf{E}\left[\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)^2\right]^{1/2}, \qquad (2.6.7)
\end{aligned}$$

where we have again used the Cauchy–Schwarz inequality on the sum inside the expectation to get the second inequality. Since $\mathbf{P}\left\{\widehat{D}_1^m = k\right\} \le c\mathbf{P}\left\{\bar{\xi} = k\right\}$, by the same methods used in (2.6.2), there exists $c' > 0$ such that (2.6.7) is at most

$$\frac{Cn^{1/4-\delta}}{\sqrt{m}} + c'\mathbf{E}\left[\left(1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right)^2\right]^{1/2}.$$

Lastly, since $x^2 \le x$ for all $x \in [0,1]$, we obtain the bound

$$\left|\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2 \breve{Y}^{n,\delta}_{\widehat{D}_1^m,i}\right]\right| \le \frac{Cn^{1/4-\delta}}{\sqrt{m}} + c'\mathbf{E}\left[1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right]^{1/2},$$

and the result follows by Lemma 2.6.1. □

We now present the proof of Proposition 2.5.10.

Proof of Proposition 2.5.10. For $n \ge 1$ and $m \le n$, let $\breve{R}_{m,n,\delta}$ be the spatial process of a branching random walk $\breve{\mathbf{T}}_{m,n} = (\mathrm{T}_m, \breve{Y}^{n,\delta})$ where the displacement vector of a vertex $v \in v(\mathrm{T}_m) \setminus \partial\mathrm{T}_m$ with k children is distributed as

$$\breve{Y}^{n,\delta}_k = Y^{n,\delta}_k - \mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right].$$

Furthermore, let

$$\breve{R}^{+}_{m,n,\delta} := \max\left\{0, \max_{0\le i\le m} \breve{R}_{m,n,\delta}\right\},$$

and

$$\breve{R}^{-}_{m,n,\delta} := -\min\left\{0, \min_{0\le i\le m} \breve{R}_{m,n,\delta}\right\}.$$

It suffices to prove that there exists $A > 0$ such that for all $m \geq 0$, all $n \geq m$ and all $\gamma > 0$,

$$\mathbf{P}\left\{\breve{R}^{+}_{m,n,\delta} > \gamma n^{1/4}\right\} \leq \frac{A}{\gamma^8} \quad \text{and} \quad \mathbf{P}\left\{\breve{R}^{-}_{m,n,\delta} > \gamma n^{1/4}\right\} \leq \frac{A}{\gamma^8},$$

since Proposition 2.5.10 then follows by taking $n = m$. We only prove the tail bound for $\breve{R}^{+}_{m,n,\delta}$ as the bound for $\breve{R}^{-}_{m,n,\delta}$ then follows by symmetry.

Notice that $\breve{R}^{+}_{1,n,\delta} = 0$ for all $n \geq 0$, and so the claim holds trivially if $m = 1$. Moreover, at the cost of taking $A > 0$ larger, it is sufficient to prove the result for $\gamma > 0$ sufficiently large. We will proceed by induction on $m \geq 2$, and hence assume that for $1 \leq k \leq m-1$ and $\gamma > 0$,

$$\mathbf{P}\left\{\breve{R}^{+}_{k,n,\delta} > \gamma n^{1/4}\right\} \leq \frac{A}{\gamma^8},$$

for all $n \geq k$.

Observe that conditionally on $\widehat{D}^m_1$ and $\Lambda^{(m)}$,

$$\breve{R}^{+}_{m,n,\delta} \overset{\mathrm{d}}{=} \max\left\{0, \max_{1 \leq i \leq \widehat{D}^m_1}\left\{\breve{R}^{+}_{\Lambda^{(m)}_i,n,\delta} + \breve{Y}^{n,\delta}_{\widehat{D}^m_1,i}\right\}\right\}.$$

For the rest of the proof, we write $\breve{Y}^{n,\delta}_i$ in place of $\breve{Y}^{n,\delta}_{\widehat{D}^m_1,i}$ to ease the notation. Take $u_0 \in (0,1)$ such that for all $0 < u < u_0$, $(1-u)^{-8} \leq 1 + 8u + 72u^2$. Then, taking $\gamma > 2/u_0$ (recall this is possible at the cost of taking $A > 0$ larger), it follows that

$$\begin{aligned}
&\mathbf{P}\left\{\breve{R}^{+}_{m,n,\delta} \leq \gamma n^{1/4}\right\} \\
&= \mathbf{E}\left[\mathbf{P}\left\{\max_{1 \leq i \leq \widehat{D}^m_1}\left\{\breve{R}^{+}_{\Lambda^{(m)},n,\delta} + \breve{Y}^{n,\delta}_i\right\} \leq \gamma n^{1/4} \middle| \widehat{D}^m_1, \Lambda^{(m)}\right\}\right] \\
&= \mathbf{E}\left[\prod_{i=1}^{\widehat{D}^m_1}\mathbf{P}\left\{\breve{R}^{+}_{\Lambda^{(m)}_i,n,\delta} \leq \gamma n^{1/4} - \breve{Y}^{n,\delta}_i \middle| \widehat{D}^m_1, \Lambda^{(m)}, \breve{Y}^{n,\delta}_i\right\}\right],
\end{aligned}$$

where in the second equality we have used the tower law and the branching property. We will bound the right-hand side of the above equality by applying induction on each term in the product. More specifically, taking $n = k = \Lambda^{(m)}_i$ and for the i-th term of the product, by the induction hypothesis we obtain

$$\begin{aligned}
&\mathbf{E}\left[\prod_{i=1}^{\widehat{D}^m_1}\mathbf{P}\left\{\breve{R}_{\Lambda^{(m)}_i,n,\delta} \leq \gamma n^{1/4} - \breve{Y}^{n,\delta}_i \middle| \widehat{D}^m_1, \Lambda^{(m)}, \breve{Y}^{n,\delta}_i\right\}\right] \\
&= \mathbf{E}\left[\prod_{i=1}^{\widehat{D}^m_1}\mathbf{P}\left\{\breve{R}_{\Lambda^{(m)}_i,n,\delta} \leq \left(\frac{\gamma n^{1/4} - \breve{Y}^{n,\delta}_i}{(\Lambda^{(m)}_i)^{1/4}}\right)(\Lambda^{(m)}_i)^{1/4} \middle| \widehat{D}^m_1, \Lambda^{(m)}, \breve{Y}^{n,\delta}_i\right\}\right] \\
&\geq \mathbf{E}\left[\prod_{i=1}^{\widehat{D}^m_1}\left(1 - \frac{A(\Lambda^{(m)}_i)^2}{(\gamma n^{1/4} - \breve{Y}^{n,\delta}_i)^8}\right)_+\right]
\end{aligned}$$

Furthermore, since $\prod_{i=1}^{k}(1-x_i)_+ \geq 1-\sum_{i=1}^{k} x_i$ for any non-negative sequence $(x_i)_{i\geq 1}$, we may lower bound the above as

$$\mathbf{E}\left[\prod_{i=1}^{\widehat{D}_1^m}\left(1-\frac{A(\Lambda_i^{(m)})^2}{(\gamma n^{1/4}-\breve{Y}_i^{n,\delta})^8}\right)_+\right] \geq 1-\frac{A}{\gamma^8}\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{n}\right)^2\left(1-\frac{\breve{Y}_i^{n,\delta}}{\gamma n^{1/4}}\right)^{-8}\right]$$
$$\geq 1-\frac{A}{\gamma^8}\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\left(1-\frac{\breve{Y}_i^{n,\delta}}{\gamma n^{1/4}}\right)^{-8}\right],$$

where the final inequality holds since $m \leq n$. Moreover, since $\gamma > 2/u_0$, we have that $|\breve{Y}_i^{n,\delta}|/(\gamma n^{1/4}) < u_0$ for any n and so $(1-\breve{Y}_i^{n,\delta}/(\gamma n^{1/4}))^{-8} \leq 1+8\breve{Y}_i^{n,\delta}/(\gamma n^{1/4}) + 72(\breve{Y}_i^{n,\delta})^2/(\gamma^2\sqrt{n})$. Hence,

$$\mathbf{P}\left\{\breve{R}_{m,n,\delta}^{+} \leq \gamma n^{1/4}\right\} \geq 1-\frac{A}{\gamma^8}+\frac{A}{\gamma^8}\mathbf{E}\left[1-\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\right] \tag{2.6.8}$$
$$-\frac{8A}{\gamma^9 n^{1/4}}\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2\breve{Y}_i^{n,\delta}\right] \tag{2.6.9}$$
$$-\frac{72A}{\gamma^{10}\sqrt{m}}\mathbf{E}\left[\sum_{i=1}^{\widehat{D}_1^m}\left(\frac{\Lambda_i^{(m)}}{m}\right)^2(\breve{Y}_i^{n,\delta})^2\right], \tag{2.6.10}$$

where we may take the denominator of the final term of the above expression to be $\gamma^{10}\sqrt{m}$ rather than $\gamma^{10}\sqrt{n}$ as $m \leq n$ and the expectation in this term is non-negative. Applying (2.6.1), (2.6.2), and (2.6.3) to bound the expectations in (2.6.8), (2.6.10), and (2.6.9), respectively, we obtain that there exist constants $B, B', B'' > 0$ such that

$$\mathbf{P}\left\{\breve{R}_{m,n,\delta}^{+} \leq \gamma n^{1/4}\right\} \geq 1-\frac{A}{\gamma^8}+\frac{AB''}{\gamma^8\sqrt{m}}-\frac{8A}{\gamma^9 n^{1/4}}\left(\frac{B' n^{1/4-\delta}}{\sqrt{m}}+\frac{B'}{m^{1/4}}\right)-\frac{72AB}{\gamma^{10}\sqrt{m}}$$
$$\geq 1-\frac{A}{\gamma^8}+\frac{A}{\gamma^8\sqrt{m}}\left(B''-\frac{8B'}{\gamma n^{\delta}}-\frac{8B'}{\gamma}-\frac{72B}{\gamma^2}\right).$$

For $\gamma > 0$ large enough, the final term in parentheses is positive so the whole expression is at least $1-A/\gamma^8$, and the result follows. □

2.7 The hairy tour

In this section we prove Theorems 2.1.4 and 2.1.5. In particular, we show that under assumptions [**A1**] and [**A3**] for a given measure π with $\eta \in [0,2)$, we have that $(n^{-1/2}H_n, n^{-1/(4-\eta)}R_n)$ converges in distribution to a generalisation of the hairy tour introduced by Janson and Marckert [60] if $\eta = 0$, and to a process whose second coordinate is a pure jump process if $\eta \in (0,2)$. Recall that by [**A3**], π is a Borel measure on $\mathbb{R}^2 \setminus \{(0,0)\}$ such that for for any $\varepsilon > 0$, both $\pi(\mathbb{R}_+ \times (\varepsilon, \infty)) < \infty$ and $\pi((\varepsilon,\infty) \times \mathbb{R}_+) < \infty$, and that for all Borel sets $A \subset \mathbb{R}_+^2 \setminus \{(0,0)\}$ for which $\pi(\partial A) = 0$,

$$r^{4-\eta}\mathbf{P}\left\{\frac{1}{r}\left(\max_{1\le i\le \xi} Y^+_{\xi,i}, \max_{1\le i\le \xi} Y^-_{\xi,i}\right) \in A)\right\} \to \pi(A)$$

as $r \to \infty$, where $Y^+_{k,j} = Y_{k,j} \vee 0$ and $Y^-_{k,j} = (-Y_{k,j}) \vee 0$. The measure π will be the intensity measure for a Poisson point process which drives the second coordinate of the limit.

Recall that $\mathbf{T}_n = (\mathrm{T}_n, Y)$ is such that given T_n, $Y = (Y^{(v)}, v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ is a collection of independent random vectors, where if $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ has k children then $Y^{(v)}$ has distribution ν_k. Observe that, for fixed $\eta \in [0,2)$, by assumption [**A3**], if the measure π has non-zero mass then

$$\max_{v\in v(\mathrm{T}_n)} \|Y^{(v)}\|_\infty = \Theta_{\mathbf{P}}(n^{1/(4-\eta)}).$$

Fix $\gamma > 0$, $\delta \in (0, 1/(4-\eta))$, and suppose that $n \ge 1$ is sufficiently large so that $n^{1/(4-\eta)-\delta} \le \gamma n^{1/(4-\eta)}$. As in the proof of tightness for Theorem 2.1.1, and more specifically as in Section 2.5, in order to prove Theorems 2.1.4 and 2.1.5, we will need to consider three "restrictions" of the branching random walk $\mathbf{T}_n$. These restrictions are a generalisation of those used in Section 2.5 from the case $\eta = 0$ to that of general $\eta \in [0,2)$; the modified definitions are given below.

We denote the restrictions of $\mathbf{T}_n$ by $\mathbf{T}_{n,\delta} = (\mathrm{T}_n, Y_{n,\delta})$, $\mathbf{T}^\gamma_{n,\delta} = (\mathrm{T}_n, Y^\gamma_{n,\delta})$, and $\mathbf{T}^\gamma_n = (\mathrm{T}_n, Y^\gamma_n)$. Again, these branching random walks will respectively capture the "typical", "mid-range", and "large" displacements in $\mathbf{T}_n$, as follows:

1. **(typical displacements):** For all $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$,

$$Y^{(v)}_{n,\delta} = Y^{(v)}\mathbf{1}_{[\|Y^{(v)}\|_\infty \le n^{1/(4-\eta)-\delta}]};$$

2. **(mid-range displacements):** For all $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$,

$$Y^{\gamma,(v)}_{n,\delta} = Y^{(v)}\mathbf{1}_{[n^{1/(4-\eta)-\delta} < \|Y^{(v)}\|_\infty \le \gamma n^{1/(4-\eta)}]};$$

3. **(large displacements):** For all $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$,

$$Y^{\gamma,(v)}_{n} = Y^{(v)}\mathbf{1}_{[\|Y^{(v)}\|_\infty > \gamma n^{1/(4-\eta)}]}.$$

We note that, informally, taking $\gamma \downarrow 0$ in $\mathbf{T}^\gamma_n$ captures all displacements of the largest order. We define $R_{n,\delta}, R^\gamma_{n,\delta}$, and R^γ_n to be the functions encoding the spatial locations of the vertices of $\mathbf{T}_{n,\delta}, \mathbf{T}^\gamma_{n,\delta}$, and $\mathbf{T}^\gamma_n$, respectively.

Before studying the convergence of the functions $R_{n,\delta}$, $R^{\gamma}_{n,\delta}$, and R^{γ}_{n}, we will prove convergence upon rescaling of the *values* of the large displacements. For $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$, let

$$Y^{(v,+)} := 0 \vee \max_{j\in[c(v,\mathrm{T}_n)]} Y^{(v)}_j,$$
$$Y^{(v,-)} := 0 \vee \max_{j\in[c(v,\mathrm{T}_n)]} (-Y^{(v)}_j),$$

be the largest positive and negative terms (respectively) in the displacement vector $Y^{(v)}$ from v to its children and, for $v \in \partial\mathrm{T}_n$, set $Y^{(v,+)} = Y^{(v,-)} = 0$. For a finite multiset $S \subset \mathbb{R}^2$, by "the decreasing ordering of S" we mean the vector $(s_1, \ldots, s_m)$ which lists the elements of S in decreasing order of their largest coordinate, breaking ties in decreasing order of their smallest coordinate. Let $L^{\eta,\gamma}_n$ be the decreasing ordering of the multiset

$$\left\{(Y^{(v,+)}, Y^{(v,-)})\mathbf{1}_{[\|Y^{(v)}\|_\infty > \gamma n^{1/(4-\eta)}]},\ v \in v(\mathrm{T}_n)\right\}, \tag{2.7.1}$$

concatenated with an infinite sequence with all entries $(0,0)$.

Lemma 2.7.1. *Fix $\gamma > 0$ and suppose that [**A1**] holds and [**A3**] holds for a given measure π with $\eta \in [0,2)$. Then as $n \to \infty$,*

$$\frac{L^{\eta,\gamma}_n}{n^{1/(4-\eta)}} \xrightarrow{\mathrm{d}} L^{\eta,\gamma}$$

in ℓ_∞, where $L^{\eta,\gamma}$ is the decreasing ordering of the points of a Poisson process on $\mathbb{R}^2_{\geq 0}$ with intensity $\pi(dx,dy)\mathbf{1}_{[(x\vee y)>\gamma]}$ concatenated with an infinite sequence with all entries $(0,0)$.

Proof. Let $(\xi_i, i \geq 1)$ be IID samples from the offspring distribution μ. Further, for $i \geq 1$, sample Y_{ξ_i} independently and let

$$Y^{+}_{\xi_i} := 0 \vee \max_{j\in[\xi_i]} Y_{\xi_i,j},$$
$$Y^{-}_{\xi_i} := 0 \vee \max_{j\in[\xi_i]} (-Y_{\xi_i,j}).$$

By definition, the multiset $\{(Y^{(v,+)}, Y^{(v,-)}), v \in v(\mathrm{T}_n)\}$ is distributed as

$$\{(Y^{+}_{\xi_i}, Y^{-}_{\xi_i}), i \in [n]\}$$

conditioned on the event that $\sum_{i=1}^n \xi_i = n-1$.

For $n \geq 1$ let $\widetilde{L}^{\eta,\gamma}_n$ be the decreasing ordering of

$$\left\{(Y^{+}_{\xi_i}, Y^{-}_{\xi_i})\mathbf{1}_{[\|Y_{\xi_i}\|_\infty > \gamma n^{1/(4-\eta)}]},\ i \in [n]\right\},$$

concatenated with an infinite sequence with all entries $(0,0)$. We will first show that as $n \to \infty$,

$$n^{-1/(4-\eta)} \widetilde{L}_n^{\eta,\gamma} \xrightarrow{\mathrm{d}} L^{\eta,\gamma} \tag{2.7.2}$$

in ℓ_∞. To this end, note that by [**A3**], for any $x, y \geq 0$ such that $x \vee y > \gamma$ and such that $\pi((\{x\} \times [y,\infty)) \cup ([x,\infty) \times \{y\})) = 0$,

$$n\mathbf{P}\left\{Y_{\xi_i}^+ > xn^{1/(4-\eta)},\ Y_{\xi_i}^- > yn^{1/(4-\eta)}\right\} \to \pi((x,\infty) \times (y,\infty)),$$

as $n \to \infty$ and moreover, $\pi((x,\infty)\times(y,\infty)) < \infty$. Therefore,

$$\begin{aligned}
&\left|\left\{i \in [n]\ :\ Y_{\xi_i}^+ > xn^{1/(4-\eta)},\ Y_{\xi_i}^- > yn^{1/(4-\eta)}\right\}\right| \\
&\qquad \stackrel{\mathrm{d}}{=} \text{Binomial}\left(n\mathbf{P}\left\{Y_{\xi_i}^+ > xn^{1/(4-\eta)},\ Y_{\xi_i}^- > yn^{1/(4-\eta)}\right\}\right) \\
&\qquad \xrightarrow{\mathrm{d}} \text{Poisson}(\pi((x,\infty)\times(y,\infty))),
\end{aligned} \tag{2.7.3}$$

and (2.7.2) follows from the fact that a Poisson process on $\mathbb{R}^2$ is determined by its distribution on half-infinite rectangles and the continuity of the function $x, y \mapsto x \vee y, x \wedge y$ that we use to order the multisets.

We now show that the convergence in (2.7.2) still holds when we condition on $\sum_{i=1}^n \xi_i = n-1$. We note that the remainder of this proof is similar to the end of the proof of Proposition 2.5.2.

Let $\widetilde{M}_n^\gamma$ be the number of elements in $\widetilde{L}_n^{\eta,\gamma}$ which are not equal to $(0,0)$. Note that by (2.7.3), the sequence $(\widetilde{M}_n^\gamma)_{n\geq 1}$ is tight. Further, let $\widetilde{S}_n^\gamma := \sum_{i\in[n]} \xi_i \mathbf{1}_{[\|Y_{\xi_i}\|_\infty > \gamma n^{1/(4-\eta)}]}$. Since $\xi_1, \ldots \xi_n$ are IID, the law of $\sum_{i=1}^n \xi_i$ depends on $\widetilde{L}_n^{\eta,\gamma}$ solely through $\widetilde{M}_n^\gamma$ and $\widetilde{S}_n^\gamma$. To be precise, let $\xi_1^n, \xi_2^n, \ldots$ be independent random variables such that for each $i \geq 1$, ξ_i^n is distributed as ξ_i conditional on $\|Y_{\xi_i}\|_\infty < \gamma n^{1/(4-\eta)}$. Then,

$$\mathbf{P}\left\{\sum_{i=1}^n \xi_i = k \;\middle|\; \widetilde{L}_n^{\eta,\gamma}\right\} = \mathbf{P}\left\{\widetilde{S}_n^\gamma + \sum_{i=1}^{n-\widetilde{M}_n^\gamma} \xi_i^n = k \;\middle|\; \widetilde{S}_n^\gamma,\ \widetilde{M}_n^\gamma\right\}. \tag{2.7.4}$$

Let $F : \ell_\infty \to \mathbb{R}$ be a bounded measurable function. Then, by analogous arguments to those used to prove (2.5.6),

$$\begin{aligned}
\mathbf{E}\left[F(L_n^{\eta,\gamma})\right] &= \mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\mathbf{1}_{[\widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}]} \;\middle|\; \sum_{i=1}^n \xi_i = n-1\right] + o(1) \\
&= \frac{\mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\mathbf{1}_{[\sum_{i=1}^n \xi_i = n-1,\ \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}]}\right]}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} + o(1) \\
&= \frac{\mathbf{E}\left[\mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\mathbf{1}_{[\sum_{i=1}^n \xi_i = n-1,\ \widetilde{M}_n^\gamma < n^\varepsilon,\ \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}]} \;\middle|\; \widetilde{L}_n^{\eta,\gamma}\right]\right]}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} + o(1) \\
&= \mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\right] \frac{\mathbf{P}\left\{\sum_i^{n-\widetilde{M}_n^\gamma} \xi_i^n = n-1-\widetilde{S}_n^\gamma \mid \widetilde{M}_n^\gamma < n^\varepsilon, \widetilde{S}_n^\gamma < n^{1/3+\varepsilon}\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} + o(1),
\end{aligned}$$

where the last equality holds by (2.7.4). By a quantitative local limit theorem (see Lemma 2.8.3 in Section 2.8), we obtain that as $n \to \infty$

$$\frac{\mathbf{P}\left\{\sum_i^{n-m} \xi_i^n = n-1-s\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \to 1,$$

uniformly over all $m < n^{\varepsilon}$ and $s < n^{1/3+\varepsilon}$. It follows that

$$\mathbf{E}\left[F(L_n^{\eta,\gamma})\right] = \mathbf{E}\left[F(\widetilde{L}_n^{\eta,\gamma})\right] + o(1).$$

The result then follows by (2.7.2). □

In the remainder of the section, we continue to use $L^{n,\gamma}$ to refer to a random vector with the distribution given in Lemma 2.7.1.

To prove Theorems 2.1.4 and 2.1.5, we use similar methods to those used to prove Theorem 2.1.1. First, we will prove convergence of the branching random walk restricted to the subtree spanned by k uniform vertices, by showing that the convergence from Proposition 2.4.1 holds jointly with that in Lemma 2.7.1, and that the limits are independent. This, in particular, implies the convergence of the random finite-dimensional distributions in Theorems 2.1.4 and 2.1.5. The independence is the key issue here, and in order to obtain it, we require adaptations of Proposition 2.4.3 and Lemma 2.4.4 to the setting of n-dependent offspring distributions. The required technical results may be found in Section 2.8.

Following this, using similar techniques to those used in Sections 2.5 and 2.6 to prove tightness for the discrete snake in Theorem 2.1.1, and applying the aforementioned joint convergence, we will show that a discrete snake comprised solely of the "typical" displacements converges to the head of the BSBE on rescaling by $n^{-1/4}$ if $\eta = 0$, and to 0 on rescaling by $n^{-1/(4-\eta)}$ if $\eta \in (0, 2)$. Furthermore, this discrete snake is asymptotically independent of the large displacements. In Section 2.7.2 we show that for $\eta \in [0, 2)$ the mid-range displacements make only a vanishing contribution to the head of the discrete snake on the scale of $n^{1/(4-\eta)}$. Next, by a small variant of Lemma 2.5.6, we deduce that the large displacements appear near the leaves. We apply this result to prove Lemma 2.7.11, which states that the discrete snake associated with the branching random walk $\mathbf{T}'_n$ obtained by pruning sub-branching random walks rooted at vertices with large displacements in $\mathbf{T}_n$ converges upon rescaling by $n^{-1/(4-\eta)}$ to the same limit as that of the "typical displacement" discrete snake (with the limit depending on whether $\eta = 0$ or $\eta \in (0, 2)$). Theorems 2.1.4 and 2.1.5 then follow by showing that the branching random walk obtained by regrafting these pruned sub-branching random walks to uniform leaves of $\mathbf{T}'_n$ has the same law as $\mathbf{T}_n$.

The following proposition establishes the convergence of the branching random walk restricted to the subtree spanned by k uniform vertices, as well as the its asymptotic independence from the large displacements.

Proposition 2.7.2. *Fix $\gamma > 0$ and suppose that [**A1**] holds and [**A3**] holds for a given measure π with $\eta \in [0, 2)$. Fix $k \geq 1$. Then*

$$\frac{\sigma}{\sqrt{n}}(J_1^n, J_2^n, \ldots, J_k^n, A_1^n, \ldots, A_k^n) \xrightarrow{\mathrm{d}} (J_1, J_2, \ldots, J_k, A_1, \ldots, A_k)$$

as $n \to \infty$. Jointly with this convergence, we have that

$$(F_1^n, F_2^n, \ldots, F_k^n) \xrightarrow{\mathrm{d}} (F_1, F_2, \ldots, F_k),$$

where $F_1, F_2, \ldots, F_k$ are IID random variables, independent of everything else, such that $\mathbf{P}\{F_i = 1\} = \mathbf{P}\{F_i = 2\} = 1/2$ and

$$\left(\frac{L^n(\lfloor tn^{1/2} \rfloor \wedge (J_1^n - 1))}{n^{1/4}}\right)_{t \geq 0} \xrightarrow{\mathrm{d}} \beta(B_{t \wedge (J_1/\sigma)})_{t \geq 0},$$

$$\left(\frac{L^n((J_i^n + \lfloor tn^{1/2} \rfloor) \wedge (J_{i+1}^n - 1))}{n^{1/4}}\right)_{t \geq 0} \xrightarrow{\mathrm{d}} \beta(B_{A_i/\sigma} + B_{((J_i/\sigma)+t) \wedge (J_{i+1}/\sigma)} - B_{(J_i/\sigma)})_{t \geq 0}$$

for $1 \leq i \leq k-1$, in each case for the uniform norm. Moreover, jointly with this convergence,

$$\frac{L_n^{\eta,\gamma}}{n^{1/(4-\eta)}} \xrightarrow{\mathrm{d}} L^{\eta,\gamma},$$

in ℓ_∞, where $L^{\eta,\gamma}$ is independent of all the other limiting random variables.

Proof. Fix $k \geq 1$ and $\gamma > 0$ and write

$$\begin{aligned} V_n = (&J_1^n, J_2^n, \ldots, J_k^n, A_1^n, \ldots, A_k^n, F_1^n, F_2^n, \ldots, F_k^n, \\ &(L^n(\lfloor tn^{1/2} \rfloor \wedge (J_1^n - 1)))_{t \geq 0}), (L^n((J_1^n + \lfloor tn^{1/2} \rfloor) \wedge (J_2^n - 1)))_{t \geq 0}), \ldots, \\ &(L^n((J_{k-1}^n + \lfloor tn^{1/2} \rfloor) \wedge (J_k^n - 1)))_{t \geq 0}) \end{aligned}$$

for the vector containing all variables that, in Proposition 2.4.1, have already been shown to converge jointly under rescaling when we equip the first $3k$ entries with the Euclidean topology on $\mathbb{R}$, the last k entries with the topology of uniform convergence, and the whole vector with the product topology. Then, let g be an $\mathbb{R}$-valued bounded continuous function (for this topology), and $h : \ell_\infty \to \mathbb{R}$ be another bounded continuous function. By Proposition 2.4.1 and Lemma 2.7.1 it suffices to prove that

$$\left| \mathbf{E}\left[g(V_n)h(L_n^{\eta,\gamma})\right] - \mathbf{E}\left[g(V_n)\right]\mathbf{E}\left[h(L_n^{\eta,\gamma})\right] \right| \to 0 \tag{2.7.5}$$

as $n \to \infty$.

Let $(M_n^{\eta,\gamma}, S_n^{\eta,\gamma})$ have the joint distribution of the number of vertices with a large displacement, $\sum_{v\in \mathrm{T}_n} \mathbf{1}_{[\|Y^{(v)}\|>\gamma n^{1/(4-\eta)}]}$, and the total number of children of such vertices, $\sum_{v\in \mathrm{T}_n} c(v,\mathrm{T}_n)\mathbf{1}_{[\|Y^{(v)}\|>\gamma n^{1/(4-\eta)}]}$. Fix $\varepsilon \in (0,1/6)$ and define the good event

$$\mathcal{G}_1 = \{M_n^{\eta,\gamma} \le n^{\varepsilon}, S_n^{\eta,\gamma} \le n^{1/3+\varepsilon}\}.$$

By analogous arguments to those used to prove (2.5.6), $\mathcal{G}_1$ occurs with high probability. Now recall that $\sigma n^{-1/2} J_k^n \xrightarrow{\mathrm{d}} J_k$ as $n \to \infty$. Fix $T > 0$ and let $\mathcal{G}_2$ be the (good) event that $J_k^n \le T\sqrt{n}$. (We observe that by choosing T large we may make $\mathbf{P}\{\mathcal{G}_2\}$ as close to 1 as we like, uniformly in n sufficiently large.) Then,

$$\mathbf{E}\left[g(V_n)h(L_n^{\eta,\gamma})\right] = \mathbf{E}\left[g(V_n)h(L_n^{\eta,\gamma})\mathbf{1}_{[\mathcal{G}_1\cap\mathcal{G}_2]}\right] + o(1),$$

where $o(1)$ is to be understood as an error that tends to 0 as $n \to \infty$ and then $T \to \infty$.

Let $\mathcal{F}_n^{\eta,\gamma}$ denote the σ-algebra generated by the degrees and displacement vectors of the vertices v with $\|Y^{(v)}\| > \gamma n^{1/(4-\eta)}$. We see that $\mathcal{G}_1$ and $L_n^{\eta,\gamma}$ are measurable with respect to $\mathcal{F}_n^{\eta,\gamma}$, and so

$$\mathbf{E}\left[g(V_n)h(L_n^{\eta,\gamma})\mathbf{1}_{[\mathcal{G}_1\cap\mathcal{G}_2]}\right] = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \mid \mathcal{F}_n^{\eta,\gamma}\right] h(L_n^{\eta,\gamma})\mathbf{1}_{[\mathcal{G}_1]}\right].$$

Therefore, since g and h are bounded, to prove (2.7.5) it suffices to show that as $n \to \infty$ and $T \to \infty$,

$$\left|\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \mid \mathcal{F}_n^{\eta,\gamma}\right]\mathbf{1}_{[\mathcal{G}_1]} - \mathbf{E}\left[g(V_n)\right]\right| \xrightarrow{\mathrm{p}} 0. \tag{2.7.6}$$

To prove (2.7.6), we will use the measure change between a size-biased random array and a vector of IID size-biased random variables which may be found in Proposition 2.8.4 below. To this end, let ξ^n denote a random variable with distribution μ, conditioned not to yield a large displacement vector (i.e. conditioned on $\max_{1\le i\le \xi^n} |Y_{\xi^n,i}| \le \gamma n^{1/(4-\eta)}$), and let μ^n denote the distribution of ξ^n. Using similar notation to that in Proposition 2.8.4, write r_n for the value of $M_n^{\eta,\gamma}$, s_n for the value of $S_n^{\eta,\gamma}$ and $d_1,\ldots,d_{r_n}$ for the degrees of the vertices v with $\|Y^{(v)}\| > \gamma n^{1/(4-\eta)}$. Then, let $\xi_{r_n+1}^n,\ldots,\xi_n^n$ be IID samples from μ^n and write $\vec{Z} = (Z_1,\ldots,Z_n) = (d_1,\ldots,d_{r_n},\xi_{r_n+1}^n,\ldots,\xi_n^n)$. Further, conditionally given $\vec{Z}$, let $\Sigma = \Sigma_{\vec{Z}}$ be the random permutation in (2.8.1), so that $(Z_{\Sigma(1)},\ldots,Z_{\Sigma(n)})$ is a size-biased random re-ordering of $\vec{Z}$. Also define $\tau_{r_n}(\Sigma) = \min\{j \in [n] : \Sigma(j) \in [r_n]\}$. Finally, write $N = N_{n,r_n} = |\{i \in \{r_n+1,\ldots,n\} : \xi_i^n > 0\}|$.

Note that conditionally on $\mathcal{F}_n^{\eta,\gamma}$, the remaining vertex degrees are distributed as $\xi_{r_n+1}^n, \dots, \xi_n^n$ conditioned on $\xi_{r_n+1}^n + \cdots + \xi_n^n = n-1-s_n$. Therefore,

$$\begin{aligned}
&\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \mid \mathcal{F}_n^{\eta,\gamma}\right] \\
&\quad = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \,\middle|\, \xi_{r_n+1}^n, \dots, \xi_n^n, \sum_{i=r_n+1}^n \xi_i^n = n-1-s_n, \mathcal{F}_n^{\eta,\gamma}\right] \,\middle|\, \mathcal{F}_n^{\eta,\gamma}\right].
\end{aligned}$$

By (2.8.18), $\tau_{r_n}(\Sigma) > T\sqrt{n}$ with high probability. Furthermore, by a Chernoff bound, $N \geq T\sqrt{n}$ with high probability. It follows that

$$\begin{aligned}
&\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \,\middle|\, \xi_{r_n+1}^n, \dots, \xi_n^n, \sum_{i=r_n+1}^n \xi_i^n = n-1-s_n, \mathcal{F}_n^{\eta,\gamma}\right] \,\middle|\, \mathcal{F}_n^{\eta,\gamma}\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]} \middle| \xi_{r_n+1}^n, \dots, \xi_n^n, \sum_{i=r_n+1}^n \xi_i^n = n-1-s_n, \mathcal{F}_n^{\eta,\gamma}\right] \middle| \mathcal{F}_n^{\eta,\gamma}\right] \\
&\quad + o_{\mathbf{P}}(1).
\end{aligned}$$

Now, observe that on the event $\mathcal{G}_2$, T_n^k contains at most $T\sqrt{n}$ vertices, and further on the event $\tau_{r_n}(\Sigma) > T\sqrt{n}$, none of these vertices have a displacement exceeding $\gamma n^{1/(4-\eta)}$. This implies that $g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]}$ only depends on $\xi_{r_n+1}^n, \dots, \xi_n^n$ and $\mathcal{F}_n^{\eta,\gamma}$ through $Z_{\Sigma(1)}, \dots, Z_{\Sigma(\lfloor T\sqrt{n}\rfloor)}$ and $\Sigma(1), \dots, \Sigma(\lfloor T\sqrt{n}\rfloor)$. Therefore,

$$\begin{aligned}
&\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \middle| \mathcal{F}_n^{\eta,\gamma}\right] \\
&\quad = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]} \middle| (Z_{\Sigma(i)})_{i\in[\lfloor T\sqrt{N}\rfloor]}, (\Sigma(i))_{i\in[\lfloor T\sqrt{N}\rfloor]}\right] \middle| \mathcal{F}_n^{\eta,\gamma}\right] \\
&\qquad + o_{\mathbf{P}}(1). \\
&\quad = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \middle| (Z_{\Sigma(i)})_{i\in[\lfloor T\sqrt{N}\rfloor]}, (\Sigma(i))_{i\in[\lfloor T\sqrt{N}\rfloor]}\right] \mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]} \middle| \mathcal{F}_n^{\eta,\gamma}\right] \\
&\qquad + o_{\mathbf{P}}(1)
\end{aligned}$$

where the last equality is implied by the fact that the events $N \geq T\sqrt{n}$ and $\tau_{r_n}(\Sigma) > T\sqrt{n}$ are measurable with respect to $Z_{\Sigma(1)}, \dots, Z_{\Sigma(\lfloor T\sqrt{n}\rfloor)}, \Sigma(1), \dots, \Sigma(\lfloor T\sqrt{n}\rfloor)$. However, observe that $g(V_n)\mathbf{1}_{[\mathcal{G}_2]}$ is independent of $\Sigma(1), \dots, \Sigma(\lfloor T\sqrt{n}\rfloor)$ given $Z_{\Sigma(1)}, \dots, Z_{\Sigma(\lfloor T\sqrt{n}\rfloor)}$, and so

$$\begin{aligned}
&\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \middle| (Z_{\Sigma(i)})_{i\in[\lfloor T\sqrt{N}\rfloor]}, (\Sigma(i))_{i\in[\lfloor T\sqrt{N}\rfloor]}\right] \mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]} \middle| \mathcal{F}_n^{\eta,\gamma}\right] \\
&\quad = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \middle| (Z_{\Sigma(i)})_{i\in[\lfloor T\sqrt{N}\rfloor]}\right] \mathbf{1}_{[N\geq T\sqrt{n},\tau_{r_n}(\Sigma)>T\sqrt{n}]} \middle| \mathcal{F}_n^{\eta,\gamma}\right]. \qquad (2.7.7)
\end{aligned}$$

We now apply the measure change from Proposition 2.8.4 to obtain that, for $\bar{\xi}_1^n, \bar{\xi}_2^n, \ldots$ IID samples from the size-biased law of μ^n, (2.7.7) is equal to

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n\right]\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n) \mid \mathcal{F}_n^{\eta,\gamma}\right], \tag{2.7.8}$$

where the inner conditional expectation of $g(V_n)\mathbf{1}_{[\mathcal{G}_2]}$ is now thought of as a measurable functional of the IID random variables $\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n$ in place of $Z_{\Sigma(1)}, \ldots, Z_{\Sigma(\lfloor T\sqrt{n}\rfloor)}$. This implies that

$$\begin{aligned}&\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]} \mid \mathcal{F}_n^{\eta,\gamma}\right]\mathbf{1}_{[\mathcal{G}_1]}\\ &\quad= \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n\right]\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n)\middle|\mathcal{F}_n^{\eta,\gamma}\right]\mathbf{1}_{[\mathcal{G}_1]} + o_{\mathbf{P}}(1).\end{aligned}$$

By applying Lemma 2.8.5 on $\mathcal{G}_1$ (which occurs with high probability),

$$\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n) \xrightarrow{\mathrm{p}} 1$$

as $n \to \infty$ and $(\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n))_{n\geq 0}$ is uniformly integrable, so (2.7.8) is equal to

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n\right]\middle|\mathcal{F}_n^{\eta,\gamma}\right] + o_{\mathbf{P}}(1).$$

Since $\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n\right]$ does not depend on $\mathcal{F}_n^{\eta,\gamma}$, it follows that

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n\right]\middle|\mathcal{F}_n^{\eta,\gamma}\right] = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n\right]\right].$$

By Corollary 2.8.10, the total variation distance between $(\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n)$ and IID size-biased samples from μ, henceforth denoted by $(\bar{\xi}_1, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor})$, tends to 0 as $n \to \infty$. Therefore, since g is bounded,

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}_1^n, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}^n\right]\right] = \mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\bar{\xi}_1, \ldots, \bar{\xi}_{\lfloor T\sqrt{n}\rfloor}\right]\right] + o(1).$$

Finally, by Lemma 2.4.4 and Proposition 2.4.3, this is in turn equal to

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\widehat{D}_1^n, \ldots, \widehat{D}_{\lfloor T\sqrt{n}\rfloor}^n\right]\mathbf{1}_{[N_n \geq T\sqrt{n}\rfloor]}\right] + o(1).$$

where we recall that $N_n = |\{i \in [n] \ : \ D_i^n > 0\}|$. Again, since the probability of $\mathcal{G}_2$ and $N_n \geq \lfloor T\sqrt{n}\rfloor$ occurring tends to 1 as $n \to \infty$ and subsequently $T \to \infty$, we see that

$$\mathbf{E}\left[\mathbf{E}\left[g(V_n)\mathbf{1}_{[\mathcal{G}_2]}\middle|\widehat{D}_1^n, \ldots, \widehat{D}_{\lfloor T\sqrt{n}\rfloor}^n\right]\mathbf{1}_{[N_n \geq T\sqrt{n}\rfloor]}\right] = \mathbf{E}\left[g(V_n)\right] + o(1),$$

which proves (2.7.6). The result follows. □

2.7.1 Typical displacements

Fix $\eta \in [0,2)$ and $\delta \in (0, 1/(10-4\eta)) \subset (0, 1/(4-\eta))$. In this section we will study the function encoding the spatial locations of the branching random walk $\mathbf{T}_{n,\delta} = (\mathrm{T}_n, Y_{n,\delta})$, namely $R_{n,\delta} : [0,n] \to \mathbb{R}$.

Proposition 2.7.3. *Fix $\gamma > 0$ and suppose that [$\mathbf{A1}$] holds and [$\mathbf{A3}$] holds for a given measure π with $\eta \in [0,2)$. Let $\delta \in (0, 1/(10-4\eta))$. If $\eta = 0$, then*

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_{n,\delta}(nt)}{n^{1/4}} \right)_{0 \le t \le 1}, \frac{L_n^{0,\gamma}}{n^{1/4}} \right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma} \mathbf{e}_t, \beta \sqrt{\frac{2}{\sigma}} \mathbf{r}_t \right)_{0 \le t \le 1}, L^{0,\gamma} \right),$$

as $n \to \infty$, in $\mathbf{C}([0,1], \mathbb{R}^2) \times \ell_\infty$. Furthermore, $L^{0,\gamma}$ is independent of $((\mathbf{e}_t, \mathbf{r}_t))_{0 \le t \le 1}$.

If $\eta \in (0,2)$, then

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_{n,\delta}(nt)}{n^{1/(4-\eta)}} \right)_{0 \le t \le 1}, \frac{L_n^{\eta,\gamma}}{n^{1/(4-\eta)}} \right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma} \mathbf{e}_t, 0 \right)_{0 \le t \le 1}, L^{\eta,\gamma} \right),$$

in $\mathbf{C}([0,1], \mathbb{R}^2) \times \ell_\infty$, where $L^{\eta,\gamma}$ is independent of $(\mathbf{e}_t)_{0 \le t \le 1}$.

Proof. The convergence of the random finite-dimensional distributions follows from Proposition 2.7.2 exactly as Corollary 2.4.2 follows from Proposition 2.4.1, but now with the additional independence from $L^{\eta,\gamma}$.

We will obtain tightness (now on the scale of $n^{1/(4-\eta)}$) via arguments very similar to those in Section 2.5, where we replace the truncations with those defined in Section 2.7. In particular, the key point is that we must show the analogue of Proposition 2.5.7, which states that

$$\lim_{k \to \infty} \limsup_{n \to \infty} \mathbf{P} \left\{ \max_{0 \le i \le k} \sup_{s,t \in [U_{(i)}^{n,k} - 1, U_{(i+1)}^{n,k} - 1]} |R_{n,\delta}(s) - R_{n,\delta}(t)| > \gamma n^{1/(4-\eta)} \right\} = 0.$$

Fix $\delta \in (0, 1/(10-4\eta))$. For all $n \ge 1$ and $k \ge 1$ let $Y_k^{n,\delta} \in \mathbb{R}^k$ be such that

$$Y_k^{n,\delta} = (Y_{k,1}^{n,\delta}, \dots, Y_{k,k}^{n,\delta}) := \begin{cases} (Y_{k,1}, \dots, Y_{k,k}) & \text{if } \max_{1 \le j \le k} |Y_{k,j}| \le n^{1/(4-\eta)-\delta}, \\ (0, \dots, 0) & \text{else.} \end{cases}$$

As discussed in Section 2.6, the displacements of the branching random walk $\mathbf{T}_{n,\delta}$ are not necessarily globally centered and so may not satisfy [$\mathbf{A1}$]. Thus to prove the result, we will need to instead consider the re-centered branching random walk $(\mathrm{T}_n, Y_{n,\delta}^*)$ where conditionally on T_n, the entries of $Y_{n,\delta}^* = (Y_{n,\delta}^{*,(v)},\ v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n)$ are independent random vectors, such that if $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ has k children then $Y_{n,\delta}^{*,(v)}$ has the same distribution as

$$Y_k^{n,\delta} - \mathbf{E}\left[Y_{\bar{\xi}, U_{\bar{\xi}}}^{n,\delta} \right].$$

The function $R^*_{n,\delta} : [0,n] \to \mathbb{R}$ encoding the spatial locations of $(\mathrm{T}_n, Y^*_{n,\delta})$ is such that for all $t \in [0,n]$,

$$R^*_{n,\delta}(t) \overset{\mathrm{d}}{=} R_{n,\delta}(t) - \mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right] \cdot H_n(t). \tag{2.7.9}$$

By Lemma 2.8.11,

$$\left|\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right]\right| = O\left((n^{1/(4-\eta)-\delta})^{1-2(4-\eta)/3}\right).$$

Since $(n^{-1/2}H_n(nt))_{0\le t\le 1} \xrightarrow{\mathrm{d}} \frac{2}{\sigma}(\mathbf{e}_t)_{0\le t\le 1}$ as $n \to \infty$ in $\mathbf{C}([0,1],\mathbb{R})$, it then follows that

$$\frac{\|H_n\|_\infty}{n^{1/(4-\eta)}}\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right] \xrightarrow{\mathrm{p}} 0 \tag{2.7.10}$$

as long as $\delta > 0$ satisfies

$$\left(\frac{1}{4-\eta} - \delta\right)\left(1 - \frac{2(4-\eta)}{3}\right) < \frac{1}{4-\eta} - \frac{1}{2}.$$

Rearranging, this is equivalent to requiring that $\delta < (10-4\eta)^{-1}$. For these values of δ, we then have

$$\sup_{t\in[0,1]} |R^*_{n,\delta}(t) - R_{n,\delta}(t)| \xrightarrow{\mathrm{d}} 0,$$

and so there is no asymptotic cost in doing this re-centering. Arguing again exactly as in Section 2.5, it is sufficient to prove the analogue of Lemma 2.5.10, which states that there exists $A > 0$ such that for any $\gamma > 0$, $\delta \in (0, 1/(4-\eta))$ and $n \ge 1$ we have

$$\mathbf{P}\left\{\|R^*_{n,\delta}\|_\infty > \gamma n^{1/(4-\eta)}\right\} \le \frac{A}{\gamma^8}.$$

It is straightforward to verify that the proof of Lemma 2.5.10 given in Section 2.6 generalises immediately to this setting, on replacing $n^{1/4}$ by $n^{1/(4-\eta)}$. □

2.7.2 Mid-range and large displacements

We will adapt the proof of Proposition 2.5.3 to the case where [**A3**] holds instead of [**A2**]. The proof of Proposition 2.5.3 uses Lemma 2.8.12 to show that, with high probability, there are no vertices with a mid-range or large displacement that are ancestrally related. To apply that lemma, it is sufficient to bound both the maximum degree in the tree and the number of vertices with a mid-range or large displacement, with high probability. The required bound on the maximal degree follows from the assumption that $\mathbf{E}[\xi^3] < \infty$. Therefore, for the adaptation, we need to obtain the same control on the number of mid-range displacements under [**A3**] as we obtained under [**A2**] in Lemma 2.5.4.

Lemma 2.7.4. *Suppose that [**A3**] holds for a given measure π and $\eta \in [0,2)$. For $\delta > 0$ sufficiently small,*

$$\left|\left\{v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n \text{ such that } \|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta}\right\}\right| = o_{\mathbf{P}}(n^{1/12}).$$

Proof. Let $\xi_1, \ldots, \xi_n$ be IID with distribution μ. Let $x \in (0,1)$ be such that $\pi(\{x\} \times \mathbb{R}_+) = \pi(\mathbb{R}_+ \times \{x\}) = 0$. Then, by [**A3**],

$$\begin{aligned} n^{1-(4-\eta)\delta}\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > n^{1/(4-\eta)-\delta}\right\} &\leq n^{1-(4-\eta)\delta}\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > xn^{1/(4-\eta)-\delta}\right\} \\ &\to \pi\Big(((x,\infty)\times\mathbb{R}_+) \cup (\mathbb{R}_+ \times (x,\infty))\Big) < \infty. \end{aligned}$$

This in particular implies that there exists $C > 0$ such that

$$\mathbf{P}\left\{\|Y_{\xi_1}\|_\infty > n^{1/(4-\eta)-\delta}\right\} \leq Cn^{-1+(4-\eta)\delta}$$

for all $n \geq 1$. It follows that

$$A_n := \left|\left\{i \in [n] \ : \ \|Y_{\xi_i}\|_\infty > n^{1/(4-\eta)-\delta}\right\}\right| \preceq_{st} \mathrm{Bin}\left(n, Cn^{-1+(4-\eta)\delta}\right).$$

By a Chernoff bound, this implies that for $\delta \in (0, (12(4-\eta))^{-1})$, and $n \geq 1$ sufficiently large, for any $\varepsilon > 0$,

$$\begin{aligned} \mathbf{P}\left\{A_n > \varepsilon n^{1/12}\right\} &\leq \mathbf{P}\left\{\mathrm{Bin}\left(n, Cn^{-1+(4-\eta)\delta}\right) > \varepsilon n^{1/12}\right\} \\ &= \mathbf{P}\left\{\mathrm{Bin}\left(n, Cn^{-1+(4-\eta)\delta}\right) > Cn^{(4-\eta)\delta}\left(1 + \left(\frac{\varepsilon}{C}n^{1/12-(4-\eta)\delta} - 1\right)\right)\right\} \\ &= O\left(\exp(-n^{(4-\eta)\delta})\right), \end{aligned}$$

so

$$\begin{aligned} \mathbf{P}\left\{A_n > \varepsilon n^{1/12} \;\middle|\; \sum_{i=1}^n \xi_i = n-1\right\} &= O\left(n^{1/2}\exp(-n^{(4-\eta)\delta})\right) \\ &= o(1), \end{aligned}$$

and the result follows. ∎

Thereafter, we obtain the following result on the mid-range displacements under [**A3**] with a proof that is analogous to that of Proposition 2.5.3; we omit the details.

Proposition 2.7.5. *Fix $\gamma > 0$ and suppose that [**A3**] holds for a given measure π and $\eta \in [0,2)$. For $\delta > 0$ sufficiently small, as $n \to \infty$,*

$$\mathbf{P}\left\{\|R^{\gamma}_{n,\delta}\|_\infty > \gamma n^{1/(4-\eta)}\right\} = o(1).$$

In the remainder of this section we will study the function encoding the spatial locations of the "large-displacement" branching random walk $\mathbf{T}_n^\gamma = (\mathrm{T}_n, Y_n^\gamma)$, namely $R_n^\gamma : [0,n] \to \mathbb{R}$.

Let Ξ be a Poisson process on $[0,1] \times \mathbb{R}_+^2 \setminus \{(0,0)\}$ with intensity $dt \otimes \pi(dx,dy)$, and let Ξ^γ be the restriction of Ξ to $[0,1] \times (\mathbb{R}_+^2 \setminus ([0,\gamma] \times [0,\gamma]))$. Also, recall the definition of the function U from just before Theorem 2.1.4.

Proposition 2.7.6. *Fix $\gamma > 0$ and suppose that [**A1**] holds and that [**A3**] holds for a given measure π and $\eta \in [0,2)$. Let $\delta \in \left(0, \frac{1}{64} \wedge \frac{1}{10-4\eta}\right)$.*

If $\eta = 0$ then as $n \to \infty$,

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_{n,\delta}(nt)}{n^{1/4}}\right)_{0\le t\le 1}, U\left(\frac{R_n^\gamma}{n^{1/4}}, \emptyset\right)\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t, \beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right)_{0\le t\le 1}, U(0,\Xi^\gamma)\right)$$

with convergence in the first coordinate in $\mathbf{C}([0,1],\mathbb{R}^2)$, and convergence in the second coordinate with respect to the Hausdorff topology on non-empty compact subsets. Furthermore, $U(0,\Xi^\gamma)$ is independent of $(\mathbf{e}_t, \mathbf{r}_t,\ 0 \le t \le 1)$.

If $\eta \in (0,2)$ then as $n \to \infty$,

$$\left(\left(\frac{H_n(nt)}{\sqrt{n}}, \frac{R_{n,\delta}(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1}, U\left(\frac{R_n^\gamma}{n^{1/(4-\eta)}}, \emptyset\right)\right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma}\mathbf{e}_t, 0\right)_{0\le t\le 1}, U(0,\Xi^\gamma)\right)$$

with convergence in the first coordinate in in $\mathbf{C}([0,1],\mathbb{R}^2)$, and convergence in the second coordinate with respect to the Hausdorff topology on non-empty compact subsets. Furthermore, $U(0,\Xi^\gamma)$ is independent of $(\mathbf{e}_t, 0 \le t \le 1)$.

We first prove Theorems 2.1.4 and 2.1.5 assuming Proposition 2.7.6.

Proof of Theorems 2.1.4 and 2.1.5 assuming Proposition 2.7.6. For γ and δ as in Proposition 2.7.6,

$$\left(\frac{R_n(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1} = \left(\frac{R_{n,\delta}(nt)}{n^{1/(4-\eta)}} + \frac{R_n^\gamma(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1} + \left(\frac{R_{n,\delta}^\gamma(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1}.$$

By Proposition 2.7.6, as $n \to \infty$, $U(n^{-1/(4-\eta)}R_n^\gamma(n\,\cdot), \emptyset) \xrightarrow{\mathrm{d}} U(0,\Xi^\gamma)$ with respect to the Hausdorff topology on non-empty compact subsets, jointly with convergence

$$\left(\frac{R_{n,\delta}(nt)}{n^{1/(4-\eta)}}\right)_{0\le t\le 1} \xrightarrow{\mathrm{d}} \begin{cases} \beta\sqrt{\frac{2}{\sigma}}\mathbf{r} & \text{if } \eta = 0 \\ 0 & \text{if } \eta \in (0,2) \end{cases}$$

in $\mathbf{C}([0,1],\mathbb{R}^2)$ where, for $\eta = 0$, $U(0,\Xi^\gamma)$ and $(\mathbf{r}_t)_{0\le t\le 1}$ are independent. Therefore,

$$U\left(\frac{R_{n,\delta}(n\cdot)}{n^{1/(4-\eta)}} + \frac{R_n^\gamma(n\cdot)}{n^{1/(4-\eta)}}, \emptyset\right) \xrightarrow{\mathrm{d}} \begin{cases} U\left(\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}, \Xi^\gamma\right) & \text{if } \eta = 0, \\ U(0,\Xi^\gamma) & \text{if } \eta \in (0,2). \end{cases}$$

Note that $U(0,\Xi)$ is a compact set by our assumptions on π, and that $U(0,\Xi^\gamma) \xrightarrow{\text{a.s.}} U(0,\Xi)$ in the Hausdorff sense as $\gamma \downarrow 0$. We have

$$d_{\mathrm{H}}\left(U\left(\frac{R_n(n\cdot)}{n^{1/(4-\eta)}},\emptyset\right), U\left(\frac{R_{n,\delta}(n\cdot)}{n^{1/(4-\eta)}} + \frac{R_n^\gamma(n\cdot)}{n^{1/(4-\eta)}},\emptyset\right)\right) \le n^{-1/(4-\eta)}\|R_{n,\delta}^\gamma\|_\infty$$

and, by Proposition 2.7.5,

$$\lim_{\gamma\to 0}\limsup_{n\to\infty} \mathbf{P}\left\{\|R_{n,\delta}^\gamma\|_\infty > \gamma n^{1/(4-\eta)}\right\} = 0.$$

We may now apply the principle of accompanying laws [18, Theorem 3.2] in order to obtain that

$$U\left(\frac{R_n(n\cdot)}{n^{1/(4-\eta)}},\emptyset\right) \xrightarrow{\text{d}} \begin{cases} U\left(\beta\sqrt{\frac{2}{\sigma}}\mathbf{r},\Xi\right) & \text{if } \eta = 0, \\ U(0,\Xi) & \text{if } \eta \in (0,2), \end{cases}$$

which yields Theorems 2.1.4 and 2.1.5. □

The remainder of this section is devoted to the proof of Proposition 2.7.6. We will need a notion of pruning and grafting of branching random walks. We refer to Figure 2.6 as a visual aid in understanding the following three definitions.

Definition 2.7.7 (Pruning branching random walks)**.** *Let* $\mathrm{T} = (T, Y)$ *be a branching random walk with displacements* $Y = (Y^{(v)}, v \in v(T) \setminus \partial T)$*. Let* $v \in v(T)$ *and* $T^{(v)}$ *be the subtree of* T *rooted at* v*. The sub-branching random walk of* T *rooted at* v *is the branching random walk* $\mathrm{T}^{(v)} = (T^{(v)}, Y')$ *with displacements* $Y' = (Y'^{(u)}, u \in v(T^{(v)}) \setminus \partial T^{(v)})$*. Also,* $\mathrm{T}^{\uparrow v}$ *is the branching random walk obtained by removing all descendants of* v *from* T*. More generally, for* $\mathrm{v} = (v_1, \ldots, v_k)$ *a sequence of distinct vertices in* $v(T)$ *such that no two vertices in* v *are ancestrally related, we set* $\mathrm{T}^{\mathrm{v}} = (\mathrm{T}^{(v)}, v \in \mathrm{v})$*, and define* $\mathrm{T}^{\uparrow \mathrm{v}}$ *inductively as* $\mathrm{T}^{\uparrow\mathrm{v}} = (\mathrm{T}^{\uparrow(v_1,\ldots,v_{k-1})})^{\uparrow v_k}$.

Definition 2.7.8 (Grafting branching random walks)**.** *For branching random walks* $\mathrm{T} = (T, Y)$ *and* $\mathrm{T}' = (T', Y')$*, and for a leaf* $l \in \partial T$*, let* $\mathrm{T} \oplus_l \mathrm{T}' = (T \oplus_l T, Y \oplus_l Y')$ *be the branching random walk defined by setting* $T \oplus_l T' = T \cup lT'$ *and, for* $v \in v(T \oplus_l T') \setminus \partial(T \oplus_l T')$ *setting*

$$(Y \oplus_l Y')^{(v)} = \begin{cases} Y^{(v)} & \text{if } v \in v(T) \setminus \partial T \\ Y'^{(u)} & \text{if } v = lu \text{ for some } u \in v(T') \setminus \partial T' \end{cases}$$

More generally, for branching random walks $\mathrm{T}, \mathrm{T}^1, \ldots, \mathrm{T}^k$ *and distinct leaves* $l_1, \ldots, l_k \in \partial T$*, define* $\mathrm{T} \oplus_{l_1,\ldots,l_k} (\mathrm{T}^1, \ldots, \mathrm{T}^k)$ *recursively over* k *as*

$$\mathrm{T} \oplus_{l_1,\ldots,l_k} (\mathrm{T}^1, \ldots, \mathrm{T}^k) = (\mathrm{T} \oplus_{l_1,\ldots,l_{k-1}} (\mathrm{T}^1, \ldots, \mathrm{T}^{k-1})) \oplus_{l_k} \mathrm{T}^k.$$

The previous definitions imply that for a branching random walk $\mathrm{T} = (T, Y)$ and $v \in v(T)$,

$$\mathrm{T}^{\uparrow v} \oplus_v \mathrm{T}^{(v)} = \mathrm{T},$$

and more generally, for a sequence of distinct vertices $\mathrm{v} = (v_1, \ldots, v_k)$ of T such that no two vertices in v are ancestrally related in T, that $\mathrm{T}^{\uparrow \mathrm{v}} \oplus_{\mathrm{v}} \mathrm{T}^{\mathrm{v}} = \mathrm{T}$.

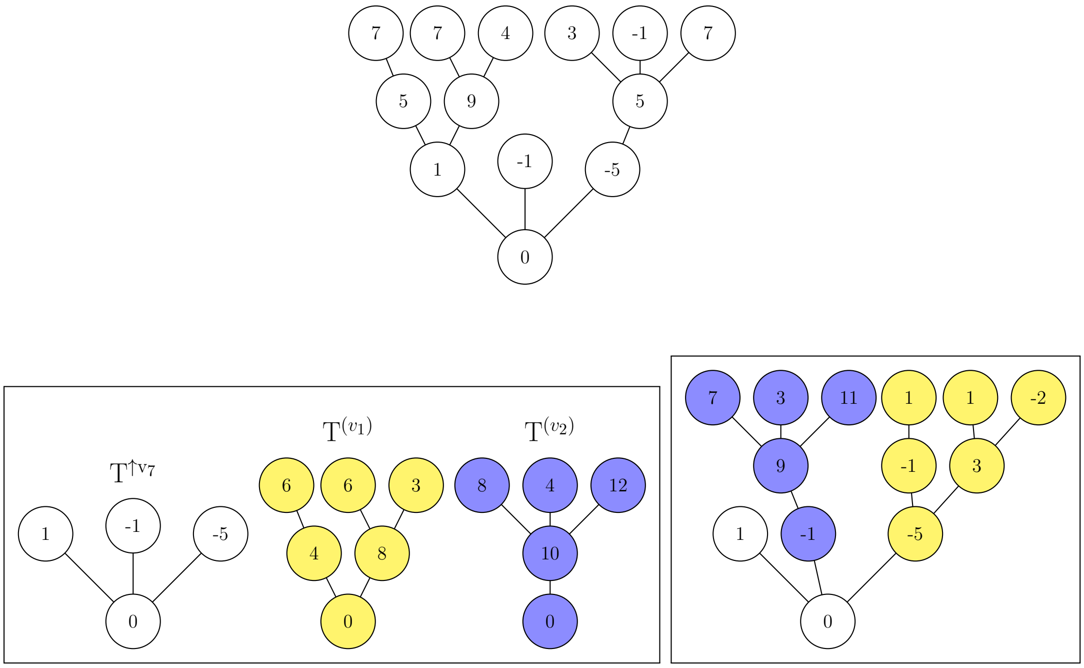

Figure 2.6: On top, a spatial tree. We denote the associated branching walk by T. On the bottom left, we depict $f_7(\mathrm{T}) = (\mathrm{T}^{\uparrow \mathrm{v}_7}, \{\mathrm{T}^{(v_1)}, \mathrm{T}^{(v_2)}\})$, which is obtained from T by pruning the sub-branching walks of T that have a displacement with absolute value exceeding 7 in their first generation. On the bottom right is a spatial tree obtained by grafting the branching walks $(f_7(\mathrm{T}))_2$ to leaves of $(f_7(\mathrm{T}))_1$. $\mathrm{T}^{(v_1)}$ and $\mathrm{T}^{(v_2)}$ to leaves of $\mathrm{T}^{\uparrow \mathrm{v}_7}$.

We next use the above definitions to define a map that prunes the sub-branching random walks of branching random walks that are rooted at ancestrally minimal vertices v with $\|Y^{(v)}\|_\infty \geq \tau$. See Figure 2.6 for an illustration of the coming definition.

Definition 2.7.9. *For a branching random walk* $\mathrm{T} = (T, Y)$ *and for* $\tau > 0$*, let* $\mathrm{v}_\tau = (v_1, \ldots, v_m)$ *be the set of vertices* $v \in v(T)$ *such that* $\|Y^{(v)}\|_\infty > \tau$ *and for all ancestors* $u \preceq v$, $\|Y^{(u)}\| \leq \tau$*, listed in depth-first order. Define a map* f_τ *by*

$$\mathrm{T} \xmapsto{f_\tau} (\mathrm{T}^{\uparrow \mathrm{v}_\tau}, \{\mathrm{T}^{(v_1)}, \ldots, \mathrm{T}^{(v_m)}\}),$$

where the second coordinate is a multiset with elements $\mathrm{T}^{(v_1)}, \ldots, \mathrm{T}^{(v_m)}$ *which are the branching random walks rooted at the vertices* $v_1, \ldots, v_m$.

For $\tau \geq 0$, let

$$v^\tau(\mathbf{T}_n) := \Big\{ v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n \ : \ \|Y^{(v)}\|_\infty > \tau, \text{ and } \|Y^{(u)}\|_\infty \leq \tau \ \forall u \prec v \Big\} .$$

We will apply f_τ to $\mathbf{T}_n$, and then study the law of $\mathbf{T}_n$ conditional on $f_\tau(\mathbf{T}_n)$. Observe that given $f_\tau(\mathbf{T}_n)$, $\mathbf{T}_n$ is determined by $\mathrm{v}^\tau(\mathbf{T}_n)$. We will show that conditional on $f_\tau(\mathbf{T}_n)$, $\mathrm{v}^\tau(\mathbf{T}_n)$ is distributed as a uniformly random subset of leaves in $(f_\tau(\mathbf{T}_n))_1$. We make this formal in the next lemma.

Lemma 2.7.10. *Let $\tau > 0$ and write $f_\tau(\mathbf{T}_n) = (\mathbf{T}'_n, \{\mathbf{T}^1_n, \ldots, \mathbf{T}^m_n\})$. Fix $m \geq 1$ and let $\Sigma \in_{\mathcal{U}} \mathcal{S}_m$, where $\mathcal{S}_m$ is the symmetric group of order m. Further, let $(\mathcal{L}_1, \ldots, \mathcal{L}_m)$ be a uniformly random vector of leaves in $\mathbf{T}'_n$ listed in depth-first order. Then, given $f_\tau(\mathbf{T}_n)$, $\mathbf{T}_n$ is equal in distribution to*

$$\mathbf{T}'_n \oplus_{\mathcal{L}_1,\ldots,\mathcal{L}_m} (\mathbf{T}^{\Sigma(1)}_n, \ldots, \mathbf{T}^{\Sigma(m)}_n).$$

Proof. Let $(\mathrm{t}', \{\mathrm{t}^1, \ldots, \mathrm{t}^m\})$ be in the support of $f_\tau(\mathbf{T}_n)$. We will first show that

$$f_\tau^{-1}(\mathrm{t}', \{\mathrm{t}^1, \ldots, \mathrm{t}^m\}) = \Big\{ \mathrm{t}' \oplus_{l_1,\ldots,l_m} (\mathrm{t}^{\pi(1)}, \ldots, \mathrm{t}^{\pi(m)}) : l_1, \ldots, l_m \in \partial \mathrm{t}'; \pi \in \mathcal{S}_m \Big\} , \tag{2.7.11}$$

where in the right-hand side, $(l_1, \ldots, l_m)$ are listed in depth-first order. Following this, we will show that the law of $\mathbf{T}_n$ conditional on its degrees and displacement vectors assigns equal mass to all elements of the right-hand set in (2.7.11) and that each element of the right-hand set corresponds to the same number of sets of leaves $(l_1, \ldots, l_m)$ listed in depth-first order and permutations π.

For the inclusion of the left-hand set in the right-hand set, observe that if for some spatial tree t it holds that $f_\tau(\mathrm{t}) = (\mathrm{t}', \{\mathrm{t}^1, \ldots, \mathrm{t}^m\})$ then, for $(l_1, \ldots, l_m)$ the minimal vertices in t that have a displacement vector with sup-norm lower bounded by τ, listed in depth-first order, are leaves in t'. Thus there is some $\pi \in \mathcal{S}_m$ such that for all $i = 1, \ldots, m$, $\mathrm{t}^{\pi(i)} = \mathrm{t}^{(l_i)}$. This implies that $\mathrm{t} = \mathrm{t}' \oplus_{l_1,\ldots,l_m} (\mathrm{t}^{\pi(1)}, \ldots, \mathrm{t}^{\pi(m)})$.

For the other inclusion, it is straightforward to see that for leaves $(l_1, \ldots, l_m)$ in t', listed in depth-first order, and $\pi \in \mathcal{S}_m$ it holds that $f_\tau(\mathrm{t}' \oplus_{l_1,\ldots,l_m} (\mathrm{t}^{\pi(1)}, \ldots, \mathrm{t}^{\pi(m)}) = (\mathrm{t}', \{\mathrm{t}^1, \ldots, \mathrm{t}^m\})$.

We now show that the law of $\mathbf{T}_n$ conditional on its degrees and displacement vectors assigns equal mass to all elements of the right-hand set. This follows from the observation that, conditional on its degrees and displacement vectors, $\mathbf{T}_n$ is uniform on all branching random walks with those degrees and displacement vectors. Each element in

$$\Big\{ \mathrm{t}' \oplus_{l_1,\ldots,l_m} (\mathrm{t}^{\pi(1)}, \ldots, \mathrm{t}^{\pi(m)}) : (l_1, \ldots, l_m) \text{ leaves in } \mathrm{t}'; \pi \in \mathcal{S}_m \Big\} \tag{2.7.12}$$

with $l_1, \ldots, l_m$ listed in depth-first order has the same degrees and displacement vectors.

Finally, we show that each element of (2.7.12) corresponds to the same number of sets of leaves $(l_1, \ldots, l_m)$ and permutations $\pi \in \mathcal{S}_m$. To this end, note that every vertex in a spatial tree t with a displacement vector whose sup-norm is at least τ has a non-zero number of children, so for each t in the set (2.7.12), we can recognise $(l_1, \ldots, l_m)$ as the vertices v that are leaves in t' and not leaves in t; thus, the choice of $(l_1, \ldots, l_m)$ is unique. Moreover, if the multiset $\{\mathrm{t}^1, \ldots, \mathrm{t}^m\}$ contains j different spatial trees with multiplicities $m_1, \ldots, m_j$ respectively, then t corresponds to $m!/(m_1! \ldots m_j!)$ different permutations π. This number does not depend on t, and the statement follows. $\square$

For $n \geq 1$, let $\tau_n = n^{1/(4-\eta)-\delta}$. Further, let $\mathbf{T}'_n = (\mathrm{T}'_n, Y')$ denote the first coordinate of $f_{\tau_n}(\mathbf{T}_n)$, and $\mathbf{F}^{\mathrm{pr}}_n = (\mathbf{T}^{(v)}_n)_{v \in v^{\tau_n}(\mathrm{T}_n)}$ denote the second coordinate of $f_{\tau_n}(\mathbf{T}_n)$, where we assume that the trees in $\mathbf{F}^{\mathrm{pr}}_n$ are ordered according to the depth-first order of their roots in T_n. We require one further lemma to prove Proposition 2.7.6.

Lemma 2.7.11. *Fix $\gamma > 0$. Suppose that [**A1**] holds and that [**A3**] holds for a given measure π and $\eta \in [0, 2)$. For $n \geq 1$, let H'_n be the height function of T'_n and R'_n be the function encoding the spatial locations of $\mathbf{T}'_n$. Extend their domains to $[0, n]$ by setting $H'_n(t) = R'_n(t) = 0$ for all $t > |\mathrm{T}'_n|$. If $\eta = 0$, then as $n \to \infty$,*

$$\left(\left(\frac{H'_n(nt)}{\sqrt{n}}, \frac{R'_n(nt)}{n^{1/4}} \right)_{0 \leq t \leq 1}, \frac{L^{0,\gamma}_n}{n^{1/4}}, \right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma} \mathbf{e}_t, \beta \sqrt{\frac{2}{\sigma}} \mathbf{r}_t \right)_{0 \leq t \leq 1}, L^{0,\gamma} \right), \tag{2.7.13}$$

and if $\eta \in (0, 2)$, then

$$\left(\left(\frac{H'_n(nt)}{\sqrt{n}}, \frac{R'_n(nt)}{n^{1/(4-\eta)}} \right)_{0 \leq t \leq 1}, \frac{L^{\eta,\gamma}_n}{n^{1/(4-\eta)}}, \right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma} \mathbf{e}_t, 0 \right)_{0 \leq t \leq 1}, L^{\eta,\gamma} \right), \tag{2.7.14}$$

with convergence in the first coordinate in $\mathbf{C}([0, 1], \mathbb{R}^2)$ endowed with the topology of uniform convergence, and the convergence in the second coordinate in ℓ_∞.

Proof. We prove (2.7.13). The proof of (2.7.14) then follows by identical arguments. By Proposition 2.7.3 it suffices to prove that as $n \to \infty$,

$$\sup_{1 \leq j \leq n} \left\{ n^{-1/2} |H_n(j) - H'_n(j)| \vee n^{-1/4} |R_{n,\delta}(j) - R'_n(j)| \right\} \xrightarrow{\mathrm{p}} 0. \tag{2.7.15}$$

We also prove (2.7.15) using Proposition 2.7.3. Fix $\varepsilon > 0$. The sample paths of both $\mathbf{e}$ and $\mathbf{r}$ are almost surely continuous so since $[0, 1]$ is compact, they are in fact almost surely uniformly continuous. This implies that there exists $\rho > 0$ so that

$$\mathbf{P}\left\{\sup_{0\le s<t\le 1,|s-t|<\rho}\left|\frac{2}{\sigma}\mathbf{e}_s-\frac{2}{\sigma}\mathbf{e}_t\right|\vee\left|\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_s-\beta\sqrt{\frac{2}{\sigma}}\mathbf{r}_t\right|>\varepsilon/2\right\}<\varepsilon/2.$$

Then, the convergence in Proposition 2.7.3 implies that for n sufficiently large, the probability that the event

$$B_n := \left\{\sup_{0\le k<\ell\le n,|k-\ell|<\rho n}\left\{\frac{|H_n(k)-H_n(\ell)|}{n^{1/2}}\vee\frac{|R_{n,\delta}(k)-R_{n,\delta}(\ell)|}{n^{1/4}}\right\}\ge\varepsilon\right\}$$

occurs is less than ε.

Next, let

$$v^*(\mathrm{T}_n) := \left\{v\in v(\mathrm{T}_n)\setminus\partial\mathrm{T}_n : \|Y^{(v)}\|_\infty\le n^{1/4-\delta}, \exists\ u\prec v \text{ with } \|Y^{(u)}\|_\infty > n^{1/4-\delta}\right\}.$$

By identical methods as those used to prove Lemma 2.5.6, it can be seen that $v^*(\mathrm{T}_n) = o_{\mathbf{P}}(n)$ and so for n sufficiently large, $\mathbf{P}\{|\mathrm{T}'_n|\le n-\rho n\}\le\varepsilon$.

Now suppose that neither of the (unlikely, bad) events $\{|\mathrm{T}'_n|\le n-\rho n\}$ or B_n hold. Observe that H'_n and R'_n can respectively be obtained from H_n and $R_{n,\delta}$ by "skipping" all the vertices in $v^*(\mathrm{T}_n)$. To be precise, for $1\le k\le|\mathrm{T}'_n|$, let $P_n(k)$ be the position of the k-th vertex that is not in $v^*(\mathrm{T}_n)$ in the depth-first order of T_n. Then,

$$(H'_n(k),R'_n(k)) = \begin{cases}(H_n(P_n(k)),R_{n,\delta}(P_n(k))) & \text{for } k=1,\ldots,|\mathrm{T}'_n|\\ (0,0) & \text{for } k>|\mathrm{T}'_n|.\end{cases}.$$

By our assumption that $n-|\mathrm{T}'_n|<\rho n$, we have $|P_n(k)-k|<\rho n$ for all k; by our assumption that

$$\sup_{0\le k<\ell\le n,|k-\ell|<\rho n}\left\{\frac{|H_n(k)-H_n(\ell)|}{n^{1/2}}\vee\frac{|R_{n,\delta}(k)-R_{n,\delta}(\ell)|}{n^{1/4}}\right\}<\varepsilon,$$

we then also have

$$\sup_{0\le k\le n}\left\{\frac{|H_n(k)-H'_n(k)|}{\sqrt{n}}\vee\frac{|R_{n,\delta}(k)-R'_n(k)|}{n^{1/4}}\right\}<\varepsilon. \tag{2.7.16}$$

Since $\varepsilon>0$ was arbitrary, the result follows. □

With Lemma 2.7.11 in hand, we proceed to proving Proposition 2.7.6. In the proof, the pair $(\mathbf{T}'_n,\mathbf{F}^{\mathrm{pr}}_n)$ is as in Lemma 2.7.11. Observe that by Lemma 2.7.10, given $f_{\tau_n}(\mathbf{T}_n)$, we can obtain an object with the same law as $\mathbf{T}_n$ by grafting the branching random walks in $\mathbf{F}^{\mathrm{pr}}_n$ at uniformly random leaves of the first coordinate of $f_{\tau_n}(\mathbf{T}_n)$.

Proof of Proposition 2.7.6. Let $n \geq 1$ be large enough so that $n^{1/(4-\eta)-\delta} < \gamma n^{1/(4-\eta)}$. Then if $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ is such that $\|Y^{(v)}\|_\infty > \gamma n^{1/(4-\eta)}$, it also holds that $\|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta}$. The proof of Proposition 2.5.5 can be adapted so that under [**A3**] for a given measure π and $\eta \in [0, 2)$, for $\delta > 0$ sufficiently small, as $n \to \infty$

$$\mathbf{P}\left\{\exists u, v \in \mathbf{T}_n, u \prec v, \text{ such that } \|Y^{(u)}\|_\infty \wedge \|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta}\right\} = o(1).$$

If follows that at the cost of throwing away an event of asymptotically vanishing probability, we may work on the event that there are no ancestrally related vertices $u, v \in v(\mathrm{T}_n)$ such that both $\|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta}$ and $\|Y^{(u)}\|_\infty > n^{1/(4-\eta)-\delta}$.

By Skorokhod's representation theorem, we may work on a probability space where the convergence in Lemma 2.7.11 holds almost surely.

We now use Lemma 2.7.10 to study the asymptotic law of R_n^γ conditional on $(\mathbf{T}'_n, \mathbf{F}_n^{\mathrm{pr}})$. Lemma 2.7.10 implies that given $(\mathbf{T}'_n, \mathbf{F}^{\mathrm{pr}})$, we can obtain an object with the law of $\mathbf{T}_n$ by grafting each of the branching random walks in $\mathbf{F}_n^{\mathrm{pr}}$ onto uniformly random leaves in $\mathbf{T}'_n$. In fact, in order to obtain the (conditional) law of R_n^γ we only need to sample the positions of the vertices in $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ whose displacement vectors $Y^{(v)}$ satisfy that $\|Y^{(v)}\|_\infty \geq \gamma n^{1/(4-\eta)}$, since the trees of $\mathbf{F}_n^{\mathrm{pr}}$ attach to these vertices in exchangeable random order. We denote the branching random walks in $\mathbf{F}_n^{\mathrm{pr}}$ by $\mathbf{T}^{(v_1)}, \ldots, \mathbf{T}^{(v_{M_n})}$ (ordered according to the depth-first order of their roots $v_1, \ldots, v_{M_n} \in \mathrm{T}_n$). By symmetry we may assume that for $1 \leq j \leq M_n$, the largest and smallest displacement at the root of $\mathbf{T}^{(v_j)}$ (i.e., $Y^{(v_j,+)}$, and $Y^{(v_j,-)}$) are described by the j-th entry of $L_n^{\eta,\gamma}$.

We claim that as $n \to \infty$, $M_n \xrightarrow{\mathrm{d}} M$ for some finite, random variable M. Indeed, as $n \to \infty$, $n^{-1/(4-\eta)} L_n^{\eta,\gamma} \xrightarrow{\text{a.s.}} L^{\eta,\gamma}$. Furthermore, since $\gamma > 0$, almost surely $L^{\eta,\gamma}$ has finitely many non-zero terms and each non-zero entry of $n^{-1/(4-\eta)} L_n^{\eta,\gamma}$ is at ℓ^∞ distance at least γ from $(0, 0)$, there are finite random variables M and N such that $L_n^{\eta,\gamma}$ has M non-zero terms for all $n > N$ large enough; i.e., the number of vertices $v \in v(\mathrm{T}_n) \setminus \partial \mathrm{T}_n$ such that $\|Y^{(v)}\|_\infty \geq \gamma n^{1/(4-\eta)}$ is equal to M.

Now, for $k \geq 1$, let $\mathcal{L}'_n(k)$ denote the number of leaves in T'_n which are among the first k vertices in the depth-first order of the vertices of T_n. Then $\mathcal{L}'_n(k)$ is bounded from above by the number of down-steps of the Łukasiewicz path $W_n(k)$ of T_n by time k. It is bounded from below by this same number minus $|T_n \setminus T'_n|$ which is $o(n)$ by (2.7.13).

Therefore, by Lemma 2.8.1, as $n \to \infty$

$$\left(\frac{\mathcal{L}'_n(\lfloor nt \rfloor)}{n}\right)_{0 \leq t \leq 1} \xrightarrow{\mathrm{p}} (\mu_0 t)_{0 \leq t \leq 1},$$

so that the positions of M_n uniform leaves in T'_n in depth-first order converge upon rescaling by n^{-1} to M independent uniform samples from $[0,1]$, which we denote by $U_1, \ldots, U_M$ respectively. For all $1 \leq j \leq M_n$, we graft $\mathbf{T}^{(v_j)}$ (which has size $o(n)$ since T'_n has size $n - o(n)$ by (2.7.13)) onto the j-th such leaf of $\mathbf{T}'_n$, using the operation in Definition 2.7.8.

The branching random walk $\mathbf{T}^{(v_j)}$ contains exactly one vertex (namely the root) with displacement vector $\|Y^{(v)}\|_\infty > \gamma n^{1/(4-\eta)}$ (since we assumed that such vertices are not ancestrally related) and the largest and smallest displacements of this vertex are given by $L_n^{\eta,\gamma}(j)$. Therefore, asymptotically, $n^{-1/(4-\eta)} R_n^\gamma$ will contain a line segment from $(U_j, -Y_j^-)$ to $(U_j, -Y_j^+)$.

This implies that if $\eta = 0$

$$\left(\left(\frac{H'_n(nt)}{\sqrt{n}}, \frac{R'_n(nt)}{n^{1/4}} \right)_{0 \leq t \leq 1}, U\left(\frac{R_n^\gamma}{n^{1/4}}, \emptyset \right) \right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma} \mathbf{e}_t, \beta \sqrt{\frac{2}{\sigma}} \mathbf{r}_t \right)_{0 \leq t \leq 1}, U(0, \Xi^\gamma) \right),$$

and if $\eta \in (0,2)$

$$\left(\left(\frac{H'_n(nt)}{\sqrt{n}}, \frac{R'_n(nt)}{n^{1/(4-\eta)}} \right)_{0 \leq t \leq 1}, U\left(\frac{R_n^\gamma}{n^{1/(4-\eta)}}, \emptyset \right) \right) \xrightarrow{\mathrm{d}} \left(\left(\frac{2}{\sigma} \mathbf{e}_t, 0 \right)_{0 \leq t \leq 1}, U(0, \Xi^\gamma) \right).$$

The result then follows from (2.7.15). □

2.8 Standard results and remaining proofs

2.8.1 Standard results

In this section we provide standard results which we use throughout this work without proof. We start by stating a functional strong law of large numbers for sums of IID non-negative random variables that we use at multiple points in proofs of convergence of finite-dimensional distributions.

Lemma 2.8.1. *Let $X_1, X_2, \ldots$ be IID random variables with $X_1 \geq 0$ almost surely and $\mathbf{E}[X_1] = \mu < \infty$. Then, for any $a_n \uparrow \infty$,*

$$\left(\frac{1}{a_n} \sum_{i=1}^{\lfloor a_n t \rfloor} X_i, t \geq 0 \right) \xrightarrow{\text{a.s.}} (\mu t, t \geq 0)$$

uniformly on compact sets as $n \to \infty$.

The next result is a generalised local central limit theorem, from Theorem 13, Chapter VII of Petrov [94], which we use to prove tightness in Theorem 2.1.1.

Theorem 2.8.2 (Theorem 13, Chapter VII of Petrov [94])**.** *Let $(X_n)_{n\geq 1}$ be a sequence of* IID *integer-valued random variables. Suppose that $\mathbf{E}[X_1]=0$, $\mathbf{Var}\{X_1\}=\sigma^2>0$, $\mathbf{E}[|X_1|^3]<\infty$, and the maximal span of the distribution of X_1 is equal to 1. Let $S_n=\sum_{i=1}^n X_i$. Then,*

$$\sqrt{2\pi n}\sigma\mathbf{P}\{S_n=k\}=e^{-k^2/(2\sigma^2 n)}\left(1+\frac{1}{\sqrt{n}}\frac{\gamma_3}{6\sigma^3}\left(\frac{k^3}{\sigma^3 n^{3/2}}-\frac{3k}{\sigma\sqrt{n}}\right)\right)+o(n^{-1/2}),$$

uniformly in $k\in\mathbb{Z}$, where γ_3 is the third central moment of X_1.

The last result is a quantitative local central limit theorem proved in [5, Lemma 5.5] for $k=1$, which we use in the proof of Theorems 2.1.4 and 2.1.5. The generalisation to $k\geq 1$ is standard.

Lemma 2.8.3. *Fix $\eta,\beta>0$, $0<\gamma<1/2$ and $k\in\mathbb{N}$. Then, there exist constants $C=C(\eta,\beta,\gamma,k)$ and $M=M(\eta,\beta,\gamma,k)$ so that for all random variables X on $\mathbb{Z}_{\geq 0}$ that satisfy the following conditions:*

1. *the greatest common divisor of the support of X is 1;*
2. *$\mathbf{P}\{X=0\}>\gamma$ and $\mathbf{P}\{X=k\}>\gamma$;*
3. *$\mathbf{E}[X^2]<\eta$ and*
4. *$\mathbf{E}[X^3]<\beta$,*

it holds that for all $m>M$

$$\sup_{\ell\in\mathbb{Z}}\left|\sqrt{m}\mathbf{P}\left\{\sum_{i=1}^m X_i=\ell\right\}-\phi\left(\frac{\ell-m\mathbf{E}[X]}{\sqrt{\mathbf{Var}\{X\}m}}\right)\right|\leq\frac{C}{\sqrt{m}}.$$

where $X_1,X_2,\ldots,$ are IID *copies of X, and $\phi(t)=e^{-t^2/2}$ is the standard normal density.*

2.8.2 Supporting results from the introduction

Proof of Lemma 2.1.3. We argue by contradiction. Fix $T>0$. Without loss of generality, assume that $\pi(\{x\}\times\mathbb{R}_+)=\delta>0$ for some $x>0$. We show that this implies that $\pi((x/2,\infty)\times\mathbb{R}_+)>T$, which contradicts the requirement that $\pi((x/2,\infty)\times\mathbb{R}_+)<\infty$ because $T>0$ was chosen arbitrarily. Fix $0<\varepsilon<x/4$ small enough that $\delta\lfloor\frac{x}{8\varepsilon}\rfloor>T$ and $\pi(\{x-\varepsilon,x+\varepsilon\}\times\mathbb{R}_+)=0$. Define $A_0=(x-\varepsilon,x+\varepsilon)$, so that by [**A3**],

$$r^{4-\eta}\mathbf{P}\left\{\frac{1}{r}\max_{1\leq i\leq\xi}Y^+_{\xi,i}\in A_0\right\}\to\pi(A_0\times\mathbb{R}_+)\geq\delta\text{ as }r\to\infty.$$

Then, letting $J = \lfloor \frac{x}{8\varepsilon} \rfloor - 1$, for $j \in \{1, \ldots, J\}$, we can find $\theta_j \in (0,1]$ such that $A_j := \theta_j(x-\varepsilon, x+\varepsilon) \subset (x-(2j+1)\varepsilon, x-(2j-1)\varepsilon)$ and $\pi(\{\theta_j(x-\varepsilon), \theta_j(x+\varepsilon)\} \times \mathbb{R}_+) = 0$. By definition, $A_0, \ldots, A_J$ are pairwise disjoint, and by our choice for J, $\cup_{0 \le j \le J} A_j \subset (x/2, x+\varepsilon)$, so $\pi((x/2, \infty) \times \mathbb{R}_+) \ge \sum_{0 \le j \le J} \pi(A_j \times \mathbb{R}_+)$. Moreover, setting $r = \theta_j s$ in the above limit shows that

$$s^{4-\eta} \mathbf{P}\left\{\frac{1}{s} \max_{1 \le i \le \xi} Y^+_{\xi,i} \in \theta_j(x-\epsilon, x+\epsilon)\right\} \to \theta_j^{\eta-4} \pi(A_0 \times \mathbb{R}_+) \ge \delta \text{ as } s \to \infty.$$

But [**A3**] implies that

$$s^{4-\eta} \mathbf{P}\left\{\frac{1}{s} \max_{1 \le i \le \xi} Y^+_{\xi,i} \in \theta_j(x-\epsilon, x+\epsilon)\right\} \to \pi(A_j \times \mathbb{R}_+),$$

so $\pi((x/2, \infty) \times \mathbb{R}_+) \ge (J+1)\delta > T$, which implies the claim. □

Proof of Proposition 2.1.2. To ease notation, we write $n/2$ instead of $\lfloor n/2 \rfloor$ throughout the proof.

First observe that, by assumption, there exist $\varepsilon, \delta > 0$ such that, for $\xi_1, \ldots, \xi_n$ IID samples from μ,

$$\limsup_{n \to \infty} \mathbf{P}\left\{\max_{1 \le i \le n} \max_{1 \le j \le \xi_i} |Y_{\xi_i, j}| > \delta n^{1/4}\right\} > \varepsilon.$$

By the central limit theorem, we may pick K large enough that

$$\liminf_{n \to \infty} \mathbf{P}\left\{n/2 - Kn^{1/2} \le \sum_{i=1}^{n/2} \xi_i \le n/2 + Kn^{1/2}\right\} > 1 - \varepsilon/2,$$

so that by a union bound

$$\limsup_{n \to \infty} \mathbf{P}\left\{\max_{1 \le i \le n} \max_{1 \le j \le \xi_i} |Y_{\xi_i, j}| > \delta n^{1/4}, n/2 - Kn^{1/2} \le \sum_{i=1}^{n/2} \xi_i \le n/2 + Kn^{1/2}\right\} > \varepsilon/2.$$

Denote the event inside the probability by $\mathcal{E}_n$. We see that

$$\begin{aligned}
&\mathbf{P}\left\{\max_{1 \le i \le n} \max_{1 \le j \le D_i^n} |Y_{D_i^n, j}| > \delta n^{1/4}\right\} \\
&\quad \ge \mathbf{P}\left\{\max_{1 \le i \le n/2} \max_{1 \le j \le D_i^n} |Y_{D_i^n, j}| > \delta n^{1/4}, n/2 - Kn^{1/2} \le \sum_{i=1}^{n/2} D_i^n \le n/2 + Kn^{1/2}\right\} \\
&\quad = \frac{\mathbf{P}\left\{\mathcal{E}_n \cap \left\{\sum_{i=1}^n \xi_i = n-1\right\}\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \\
&\quad = \frac{\mathbf{E}\left[\mathbf{1}_{[\mathcal{E}_n]} \mathbf{P}\left\{\sum_{i=n/2+1}^n \xi_i = n - 1 - \sum_{i=1}^{n/2} \xi_i \mid \xi_1, \ldots, \xi_{n/2}, Y_{\xi_1}, \ldots, Y_{\xi_{n/2}}\right\}\right]}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}} \\
&\quad \ge \mathbf{P}\left\{\mathcal{E}_n\right\} \frac{\min_{n/2-1-Kn^{1/2} \le m \le n/2-1+Kn^{1/2}} \mathbf{P}\left\{\sum_{i=n/2+1}^n \xi_i = m\right\}}{\mathbf{P}\left\{\sum_{i=1}^n \xi_i = n-1\right\}}.
\end{aligned}$$

By the local central limit theorem, there exist constants $c, C > 0$ such that

$$\liminf_{n\to\infty} n^{1/2} \min_{n/2-1-Kn^{1/2}\le m\le n/2-1+Kn^{1/2}} \mathbf{P}\left\{\sum_{i=n/2+1}^{n} \xi_i = m\right\} > c$$

and

$$\limsup_{n\to\infty} n^{1/2}\mathbf{P}\left\{\sum_{i=1}^{n} \xi_i = n-1\right\} < C.$$

It follows that

$$\limsup_{n\to\infty} \mathbf{P}\left\{\max_{1\le i\le n} \max_{1\le j\le D_i^n} |Y_{D_i^n,j}| > \delta n^{1/4}\right\} \ge \frac{\varepsilon c}{2C} > 0,$$

as claimed. □

2.8.3 Measure change

For $n \ge 1$ let $\mathcal{S}_n$ denote the set of permutations of $[n]$. For $(k_1, \ldots, k_n) \in \mathbb{N}^n$, let $\Sigma = \Sigma_{(k_1,\ldots,k_n)}$ be the random permutation of $[n]$ with law given by

$$\mathbf{P}\{\Sigma = \sigma\} = \prod_{i=1}^{n} \frac{k_{\sigma(i)}}{\sum_{j=i}^{n} k_{\sigma(j)}}, \qquad \text{for } \sigma \in \mathcal{S}_n.$$

We call $(k_{\Sigma(1)}, \ldots k_{\Sigma(n)})$ the *size-biased random re-ordering* of $(k_1, \ldots, k_n)$. It will be convenient to extend this definition to vectors $(k_1, \ldots, k_n)$ that contain 0-valued entries. We start with a size-biased random re-ordering of the non-zero entries of $(k_1, \ldots, k_n)$ and then append to this the correct number of zeroes. Formally, if $(k_1, \ldots, k_n) \in \mathbb{Z}_{\ge 0}^n$ has $N \ge 0$ non-zero entries, let $\Sigma_{(k_1,\ldots,k_n)}$ be the random permutation of $[n]$ with

$$\mathbf{P}\left\{\Sigma_{(k_1,\ldots,k_n)} = \sigma\right\} = \frac{1}{(n-N)!} \prod_{i=1}^{N} \frac{k_{\sigma(i)}}{\sum_{j=1}^{N} k_{\sigma(j)}}, \tag{2.8.1}$$

for $\sigma \in \mathcal{S}_n$, and still refer to $(k_{\Sigma(1)}, \ldots k_{\Sigma(n)})$ as the size-biased random re-ordering of $(k_1, \ldots, k_n)$.

For a permutation $\sigma \in \mathcal{S}_n$ and $r \in \{0, 1, \ldots, n\}$ define

$$\tau_r(\sigma) = \begin{cases} \min\{j \in [n] : \sigma(j) \in [r]\} & \text{if } r \in [n], \\ n+1 & \text{if } r = 0\,. \end{cases}$$

As discussed in Section 2.4, the proof of Theorem 2.1.1 relies on establishing a change of measure, (2.4.5), which relates the size-biased random re-ordering of the positive entries of the degree sequence of T_n, and IID samples from the offspring distribution. The proofs of Theorems 2.1.4 and 2.1.5 rely on establishing a similar change of measure, which is a generalisation of (2.4.5) to the situation where instead of an IID sequence, the first r elements are non-zero and are fixed in advance; the whole sequence is conditioned to have sum $n-1$; and we consider the first m elements of the size-biased random reordering of the sequence. Specifically, let $m, n, r, s \in \mathbb{Z}_{\geq 0}$ with $m, r, s < n$, and μ be a distribution on $\mathbb{Z}_{\geq 0}$. For $k_1, \ldots, k_m \in \mathbb{N}$, we define

$$\Theta_\mu(k_1, \ldots, k_m) = \Theta_\mu^{n,r,s}(k_1, \ldots, k_m)$$
$$= \frac{\mathbf{P}\left\{X_{m+1} + \cdots + X_{n-r} = n - 1 - s - \sum_{i=1}^m k_i\right\}}{\mathbf{P}\left\{X_1 + \cdots + X_{n-r} = n - 1 - s\right\}} \cdot (\mathbf{E}\left[X_1\right])^m \cdot \prod_{i=1}^m \frac{n - r - i + 1}{n - 1 - \sum_{j=1}^{i-1} k_j} \tag{2.8.2}$$

if $k_1 + \cdots + k_m \leq n - 1 - s$, and otherwise $\Theta_\mu(k_1, \ldots, k_m) = 0$, where $(X_i, i \geq 1)$ are IID random variables with distribution μ. We note that when $r = s = 0$, and μ is a critical offspring distribution, we recover (2.4.5).

Proposition 2.8.4. *Fix $n, r, s \in \mathbb{Z}_{\geq 0}$ with $r, s < n$, and $d_1, \ldots, d_r \in \mathbb{N}$ with $\sum_{i=1}^r d_i = s$. Let μ be a distribution on $\mathbb{Z}_{\geq 0}$ and $(X_i, i \geq 1)$ be IID random variables with distribution μ. Further, let*

$$N = N_{n,r} = |\{i \in \{r+1, \ldots, n\} : X_i > 0\}|.$$

Let $\vec{Z} = (Z_1, \ldots, Z_n) = (d_1, \ldots, d_r, X_{r+1}, \ldots, X_n)$, and conditionally given $\vec{Z}$, let $\Sigma = \Sigma_{\vec{Z}}$ be given by (2.8.1). Finally, let $(\overline{X}_i, i \in [n])$ be IID samples from the size-biased distribution of X_1. Suppose that $\mathbf{E}\left[X_1\right] < \infty$. Then for any $m \in [n-r]$ and any function $f : \mathbb{N}^m \to \mathbb{R}$, if $\mathbf{P}\left\{X_{r+1} + \ldots + X_n = n - 1 - s)\right\} > 0$, then

$$\begin{aligned} &\mathbf{E}\left[f\left(Z_{\Sigma(1)}, \ldots, Z_{\Sigma(m)}\right) \mathbf{1}_{[N \geq m]} \mathbf{1}_{[\tau_r(\Sigma) > m]} \;\middle|\; X_{r+1} + \ldots + X_n = n - 1 - s\right] \\ &= \mathbf{E}\left[f(\overline{X}_1, \ldots, \overline{X}_m)\Theta_\mu(\overline{X}_1, \ldots, \overline{X}_m)\right], \end{aligned} \tag{2.8.3}$$

where $\Theta_\mu(\overline{X}_1, \ldots \overline{X}_m) = \Theta_\mu^{n,r,s}(\overline{X}_1, \ldots \overline{X}_m)$ is as in (2.8.2).

We observe that when $r \neq 0$, $\tau_r(\Sigma) > m$ implies that $N \geq m$ because all positive entries occur before zero-valued entries in the size-biased random reordering. However, when $r = 0$, the former event is vacuously true for all $m \in [n]$, but we still enforce that $N \geq m$ in (2.8.3). It follows that Proposition 2.4.3 is the special case when $r = 0$, $s = 0$ and $X_1, X_2, \ldots$ are IID samples from the offspring distribution μ.

Proof. In this proof, for $n \geq 1$, and $r \geq 1$, we let

$$[n]_r = \{(n_1, \ldots, n_r) \in \{1, \ldots, n\}^r \;:\; n_i \neq n_j \text{ for all } i \neq j\}.$$

Furthermore, for a set A we write A_r for the set of ordered sequences $(s_1, \ldots, s_r)$ of r distinct elements of A. We also let $\mu_i = \mathbf{P}\{X_1 = i\}$ for $i \in \mathbb{Z}_{\geq 0}$.

We first prove the proposition assuming that $\mu_0 = 0$; we will later generalise this by conditioning on the number of non-zero entries of $\vec{Z}$ and sampling a size-biased re-ordering of only these entries. When $\mu_0 = 0$, we have $\mathbf{P}\{N = n\} = 1$, so the indicator $\mathbf{1}_{[N \geq m]}$ in (2.8.3) equals 1 and may be ignored.

For $\sigma \in S_n$ we write $\vec{Z}_\sigma = (Z_{\sigma(1)}, \ldots, Z_{\sigma(n)})$ and $\sigma^{-1}[r] = (\sigma^{-1}(1), \ldots, \sigma^{-1}(r))$. Observe that for $m \in [n-r]$, we have the equality of events

$$\{\tau_r(\Sigma) > m\} = \left\{\Sigma^{-1}[r] \in ([n] \backslash [m])_r\right\}.$$

It is thus useful to determine the law of $(\vec{Z}_\Sigma, \Sigma^{-1}[r])$. Note that for any $\vec{k} = (k_1, \ldots, k_n) \in \mathbb{N}^n$ and $\vec{j} = (j_1, \ldots, j_r) \in [n]_r$, if $(\vec{k}, \vec{j})$ is in the support of $(\vec{Z}_\Sigma, \Sigma^{-1}[r])$ then $k_{j_i} = d_i$ for each $i \in [r]$. For such $(\vec{k}, \vec{j})$,

$$\begin{aligned}\mathbf{P}\left\{\vec{Z}_\Sigma = \vec{k}, \Sigma^{-1}[r] = \vec{j}\right\} &= \sum_{\sigma \in \mathcal{S}_n : \sigma^{-1}[r] = \vec{j}} \mathbf{P}\left\{\vec{Z}_\sigma = \vec{k}, \Sigma = \sigma\right\} \\ &= \prod_{i=1}^n \frac{k_i}{\sum_{j=i}^n k_j} \cdot \sum_{\sigma \in \mathcal{S}_n : \sigma^{-1}[r] = \vec{j}} \mathbf{P}\left\{\vec{Z}_\sigma = \vec{k}\right\}. \end{aligned} \tag{2.8.4}$$

Since we fixed $\sigma^{-1}[r]$, the sum (2.8.4) ranges over exactly $(n-r)!$ elements of $\mathcal{S}_n$ and each term of the sum is equal to

$$\prod_{j \in [n] \backslash \{j_1, \ldots, j_r\}} \mu_{k_j}.$$

Hence, for any $\vec{k} \in \mathbb{N}^n$ and $\vec{j} \in [n]_r$,

$$\mathbf{P}\left\{\vec{Z}_\Sigma = \vec{k}, \Sigma^{-1}[r] = \vec{j}\right\} = (n-r)! \left(\prod_{i=1}^r \mathbf{1}_{[k_{j_i} = d_i]}\right) \left(\prod_{j \in [n] \backslash \{j_1, \ldots, j_r\}} \mu_{k_j}\right) \left(\prod_{i=1}^n \frac{k_i}{\sum_{j=i}^n k_j}\right). \tag{2.8.5}$$

Now, fix $m \in [n-r]$ and $k_1, \ldots, k_m \in \mathbb{N}$. Note that it suffices to prove (2.8.3) when $f : \mathbb{N}^m \to \mathbb{R}$ has the form

$$f(z_1, \ldots, z_m) = \prod_{i=1}^m \mathbf{1}_{[z_i = k_i]}, \tag{2.8.6}$$

so we now restrict our attention to this case. Since

$$\sum_{i \in [n]} Z_{\Sigma(i)} = \sum_{i=1}^r d_i + \sum_{i=r+1}^n X_i = s + \sum_{i=r+1}^n X_i\,,$$

for any $k_1, \ldots, k_m \in \mathbb{N}$, by summing over the possible values of $Z_{\Sigma(m+1)}, \ldots, Z_{\Sigma(n)}$ we can use (2.8.5) to find that

$$\mathbf{P}\left\{(Z_{\Sigma(1)}, \ldots, Z_{\Sigma(m)}) = (k_1, \ldots, k_m), \tau_r(\Sigma) > m, \sum_{i=r+1}^{n} X_i = n-1-s\right\} \tag{2.8.7}$$
$$= \sum_{(k_{m+1},\ldots,k_n)\in\mathbb{N}^{n-m}} \sum_{\vec{j}\in([n]\setminus[m])_r} \mathbf{1}_{[\sum_{i=1}^n k_i = n-1]} \mathbf{P}\left\{\vec{Z}_\Sigma = (k_1, \ldots, k_n), \Sigma^{-1}[r] = \vec{j}\right\}$$
$$= (n-r)! \left(\prod_{i=1}^{m} \mu_{k_i}\right)\left(\prod_{i=1}^{m} \frac{k_i}{n-1-\sum_{j=1}^{i-1} k_i}\right)$$
$$\cdot \sum_{\substack{(k_{m+1},\ldots,k_n)\in\mathbb{N}^{n-m} \\ \vec{j}\in([n]\setminus[m])_r}} \mathbf{1}_{[\sum_{i=1}^n k_i = n-1]} \left(\prod_{i=1}^{r} \mathbf{1}_{[k_{j_i}=d_i]}\right)\left(\prod_{i\in([n]\setminus[m])\setminus\{j_1,\ldots,j_r\}} \mu_{k_i}\right)\left(\prod_{i=m+1}^{n} \frac{k_i}{\sum_{j=i}^{n} k_j}\right)$$

Using that $k\mu_k = \mathbf{P}\left\{\overline{X}_1 = k\right\}\mathbf{E}[X_1]$ for all $k \in \mathbb{N}$, writing $n' = n - m$, and re-indexing the above sum, this yields that (2.8.7) is equal to

$$\mathbf{P}\left\{(\overline{X}_1, \ldots, \overline{X}_m) = (k_1, \ldots, k_m)\right\}\mathbf{E}[X_1]^m \left(\prod_{i=1}^{m} \frac{n-r-i+1}{n-1-\sum_{j=1}^{i-1} k_j}\right)(n'-r)!$$
$$\cdot \sum_{\substack{(k'_1,\ldots,k'_{n'})\in\mathbb{N}^{n'} \\ \vec{j}\in[n']_r}} \mathbf{1}_{[\sum_{i=1}^{n'} k'_i = n-1-\sum_{i=1}^{m} k_i]} \left(\prod_{i=1}^{r} \mathbf{1}_{[k'_{j_i}=d_i]}\right)\left(\prod_{i\in[n']\setminus\{j_1,\ldots,j_r\}} \mu_{k'_i}\right) \cdot \prod_{i\in[n']} \frac{k'_i}{\sum_{j=i}^{n'} k'_j}.$$

Now, define $\vec{Z}' = (d_1, \ldots, d_r, X_{r+1}, \ldots, X_{n-m})$, and conditionally given $\vec{Z}'$, let $\Sigma' = \Sigma_{\vec{Z}'}$ be given by (2.8.1). Applying (2.8.5) to $\vec{Z}'$ and Σ', we thus find that (2.8.7) equals

$$\mathbf{P}\left\{(\overline{X}_1, \ldots, \overline{X}_m) = (k_1, \ldots, k_m)\right\}\prod_{i=1}^{m}\left(\frac{n-1-i+1}{n-1-\sum_{j=1}^{i-1} k_j}\mathbf{E}[X_1]\right)$$
$$\cdot \sum_{\substack{(k'_1,\ldots,k'_{n'})\in\mathbb{N}^{n'} \\ \vec{j}\in[n']_r}} \mathbf{1}_{[\sum_{i=1}^{n'} k'_i = n-1-\sum_{i=1}^{m} k_i]} \mathbf{P}\left\{\vec{Z}'_{\Sigma'} = \vec{k}, (\Sigma')^{-1}[r] = \vec{j}\right\}$$
$$= \mathbf{P}\left\{(\overline{X}_1, \ldots, \overline{X}_m) = (k_1, \ldots, k_m)\right\}\prod_{i=1}^{m}\left(\frac{n-1-i+1}{n-1-\sum_{j=1}^{i-1} k_j}\mathbf{E}[X_1]\right)$$
$$\cdot \mathbf{P}\left\{\sum_{i=1}^{n'} Z'_{\Sigma'(i)} = n-1-\sum_{i=1}^{m} k_i\right\}.$$

Finally, since the sum of the entries of $\vec{Z}'_{\Sigma'}$ is unaffected by the random reordering and is the same as $s+\sum_{i=1}^{n'-r+m} X_i = s+\sum_{i=1}^{n-(r+m)} X_i$, we deduce that (2.8.7) equals

$$\mathbf{P}\left\{(\overline{X}_1,\ldots,\overline{X}_m)=(k_1,\ldots,k_m)\right\}\prod_{i=1}^{m}\left(\frac{n-1-i+1}{n-1-\sum_{j=1}^{i-1}k_j}\mathbf{E}\left[X_1\right]\right)$$
$$\cdot\,\mathbf{P}\left\{\sum_{i=1}^{n-(r+m)} X_i = n-1-s-\sum_{i=1}^{m} k_i\right\}.$$

Dividing the above expression by $\mathbf{P}\left\{\sum_{i=1}^{n-r} X_i = n-1-s\right\}$ yields the statement when $\mu_0=0$, in the special case that f has the form given in (2.8.6), and thus for general f.

For the general case with $\mu_0>0$, we let $p=1-\mu_0$. Further, we let $\mathbf{X}_1,\mathbf{X}_2,\ldots$ be IID copies of X_1 conditioned to be positive. Notice that $\mathbf{E}\left[\mathbf{X}_1\right]=p^{-1}\mathbf{E}\left[X_1\right]$ and that the size-biased distributions of $\mathbf{X}_1$ and of X_1 are identical. We let $\widehat{\mathbf{X}}_1,\widehat{\mathbf{X}}_2,\ldots$ denote IID samples from the size-biased distribution of $\mathbf{X}_1$. Finally, fix $m'\geq 0$ and define

$$\vec{\mathbf{Z}}'=(\mathbf{Z}'_1,\ldots,\mathbf{Z}'_{m'+r})=(d_1,\ldots,d_r,\mathbf{X}_{r+1},\ldots,\mathbf{X}_{r+m'}),$$

and conditionally given $\vec{\mathbf{Z}}'$, let $\Sigma'=\Sigma_{\vec{\mathbf{Z}}'}$ be given by (2.8.1).

Now fix $m\in[n-r]$ with $m\leq m'$. For $k_1,\ldots,k_m\in\mathbb{N}$, we have that

$$\mathbf{P}\left\{(Z_{\Sigma(1)},\ldots,Z_{\Sigma(m)})=(k_1,\ldots,k_m),\tau_r(\Sigma)>m,\sum_{i=r+1}^{n}X_i=n-1-s\;\middle|\;N=m'\right\}$$
$$=\mathbf{P}\left\{(\mathbf{Z}'_{\Sigma'(1)},\ldots,\mathbf{Z}'_{\Sigma'(m)})=(k_1,\ldots,k_m),\ \tau_r(\Sigma')>m,\ \sum_{i=r+1}^{m'}\mathbf{X}_i=n-1-s\right\}. \tag{2.8.8}$$

By the proof of the case where $\mu_0=0$, if $k_1+\cdots+k_m\leq n-1-s$ this is equal to

$$\mathbf{P}\left\{(\overline{\mathbf{X}}_1,\ldots,\overline{\mathbf{X}}_m)=(k_1,\ldots,k_m)\right\}$$
$$\cdot\,\mathbf{P}\left\{\sum_{i=m+1}^{m'}\mathbf{X}_i=n-1-s-\sum_{i=1}^{m}k_i\right\}\prod_{i=1}^{m}\left(\frac{m'-i+1}{n-1-\sum_{j=1}^{i-1}k_j}\mathbf{E}\left[\mathbf{X}_1\right]\right),$$

and otherwise is equal to 0. For the remainder of the proof we may thus assume that $k_1+\cdots+k_m\leq n-1-s$. Since $\overline{\mathbf{X}}_1\overset{\mathrm{d}}{=}\overline{X}_1$ and $\mathbf{E}\left[\mathbf{X}_1\right]=p^{-1}\mathbf{E}\left[X_1\right]$ this is in turn equal to

$$\mathbf{P}\left\{(\overline{X}_1,\ldots,\overline{X}_m)=(k_1,\ldots,k_m)\right\}$$
$$\cdot\,\frac{1}{p^m}\mathbf{P}\left\{\sum_{i=m+1}^{m'}\mathbf{X}_i=n-1-s-\sum_{i=1}^{m}k_i\right\}\prod_{i=1}^{m}\left(\frac{m'-i+1}{n-1-\sum_{j=1}^{i-1}k_j}\mathbf{E}\left[X_1\right]\right). \tag{2.8.9}$$

It then follows from (2.8.8) and (2.8.9) that

$$\mathbf{P}\left\{(Z_{\Sigma(1)},\ldots,Z_{\Sigma(m)})=(k_1,\ldots,k_m),N\geq m,\tau_r(\Sigma)>m,\sum_{i=r+1}^{n}X_i=n-1-s\right\}$$
$$=\mathbf{P}\left\{(\overline{X}_1,\ldots,\overline{X}_m)=(k_1,\ldots,k_m)\right\}$$
$$\cdot\sum_{m'=m}^{n-r}\frac{\mathbf{P}\left\{N=m'\right\}}{p^m}\mathbf{P}\left\{\sum_{i=m+1}^{m'}\mathbf{X}_i=n-1-s-\sum_{i=1}^{m}k_i\right\}\prod_{i=1}^{m}\left(\frac{m'-i+1}{n-1-\sum_{j=1}^{i-1}k_j}\mathbf{E}\left[X_1\right]\right).\tag{2.8.10}$$

Notice now that $N\overset{\mathrm{d}}{=}\mathrm{Binomial}(n-r,p)$. So using the change of variable $\ell=m'-m$ and letting M be a Binomial$(n-(r+m),p)$, by routine algebra we obtain that (2.8.10) equals

$$\sum_{\ell=0}^{n-(r+m)}\mathbf{P}\left\{M=\ell\right\}\mathbf{P}\left\{\sum_{i=m'-\ell+1}^{m+\ell}\mathbf{X}_i=n-1-s-\sum_{i=1}^{m}k_i\right\}\cdot\prod_{i=1}^{m}\left(\frac{n-r-i+1}{n-1-\sum_{j=1}^{i-1}k_i}\mathbf{E}\left[X_1\right]\right)$$
$$=\sum_{\ell=0}^{n-(r+m)}\mathbf{P}\left\{M=\ell\right\}\mathbf{P}\left\{\sum_{i=1}^{\ell}\mathbf{X}_i=n-1-s-\sum_{i=1}^{m}k_i\right\}\cdot\prod_{i=1}^{m}\left(\frac{n-r-i+1}{n-1-\sum_{j=1}^{i-1}k_i}\mathbf{E}\left[X_1\right]\right)$$
$$=\mathbf{P}\left\{\sum_{i=1}^{M}\mathbf{X}_i=n-1-s-\sum_{i=1}^{m}k_i\right\}\prod_{i=1}^{m}\left(\frac{n-r-i+1}{n-1-\sum_{j=1}^{i-1}k_i}\mathbf{E}\left[X_1\right]\right).$$

Since $\sum_{i=1}^{M}\mathbf{X}_i\overset{\mathrm{d}}{=}\sum_{i=1}^{n-(r+m)}X_i$, dividing the above expression (which is equal to (2.8.10)) by $\mathbf{P}\left\{\sum_{i=1}^{n-r}X_i=n-1-s\right\}$ yields the result for the special case that f has the form given in (2.8.6), and thus for general f. □

The next proposition gives conditions under which the change of measure $\Theta_{\mu}^{n,r,s}$ appearing in (2.8.2) is asymptotically unimportant in the specific case when $m=\Theta(\sqrt{n})$ and $(X_i,i\geq 1)$ are IID samples from the offspring distribution μ conditioned to yield a displacement vector such that $\max_{1\leq j\leq X_i}|Y_{X_i,j}|\leq\gamma n^{1/(4-\eta)}$. This then allows us to use the measure change in the proofs of Theorems 2.1.4 and 2.1.5.

Lemma 2.8.5. *Let μ be a critical offspring distribution with variance $\sigma^2\in(0,\infty)$, and let $\nu=(\nu_k)_{k\geq 1}$ be such that [**A1**] holds and [**A3**] holds for a given measure π with $\eta\in[0,2)$. Fix $\gamma>0$. Let ξ denote a random variable with distribution μ, and for $n\geq 1$ let ξ^n be distributed as ξ, conditioned to not yield a displacement vector with $\max_{1\leq i\leq\xi^n}|Y_{\xi^n,i}|>\gamma n^{1/(4-\eta)}$. Further, let μ^n denote the distribution of ξ^n, and let $\bar{\xi}_1^n,\bar{\xi}_2^n,\ldots$ be IID samples from the size-biased law of ξ^n.*

Finally, fix $\varepsilon\in(0,1/6)$ and let $(r_n)_{n\geq 1}$ and $(s_n)_{n\geq 1}$ be sequences such that for all $n\geq 1$, $r_n<n^{\varepsilon}$, $s_n<n^{1/3+\varepsilon}$ and $n-1-s_n$ is in the support of $\sum_{i=r_n+1}^{n}\xi_i^n$.

Suppose that $m=\Theta(\sqrt{n})$. Then as $n\to\infty$,

$$\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n,\ldots,\bar{\xi}_m^n)\overset{\mathrm{p}}{\to}1,\tag{2.8.11}$$

and $(\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n,\ldots,\bar{\xi}_m^n))_{n\geq 1}$ *is a uniformly integrable sequence of random variables.*

The proof of Lemma 2.8.5 is very similar to that of Lemma 2.4.4. However, in this case instead of the standard local central limit theorem, we will require a quantitative local central limit theorem in order to get uniform estimates on local probabilities for the family of random variables $\{\xi^n, n \geq 1\}$.

Lemma 2.8.6. *Let μ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$, and let $\nu = (\nu_k)_{k\geq 1}$ be such that [**A1**] holds and [**A3**] holds for a given measure π with $\eta \in [0,2)$. Let $\gamma > 0$. Further, let ξ denote a random variable with distribution μ and for $n \geq 1$ let $\xi^n = (\xi_i^n, i \geq 1)$ be* IID *copies of ξ each conditioned to satisfy $\{\max_{1\leq i\leq \xi_j^n} |Y_{\xi_j^n,i}| \leq \gamma n^{1/(4-\eta)}\}$. Then there exist $C, N > 0$ and M such that for all $m, n > N$,*

$$\sup_{k\in\mathbb{Z}} \left| \sqrt{m}\mathbf{P}\left\{\sum_{i=1}^m \xi_i^n = k\right\} - \phi\left(\frac{k - m\mathbf{E}\left[\xi_1^n\right]}{\sqrt{\mathbf{Var}\left\{\xi_1^n\right\} m}}\right)\right| \leq \frac{C}{\sqrt{m}},$$

where $\phi(t) = e^{-t^2/2}$ is the standard normal density.

This lemma is immediate from Lemma 2.8.3 as soon as we show that the family $\{\xi^n,\ n \geq 1\}$ satisfies the conditions of that lemma. This is verified in Lemmas 2.8.7 and 2.8.8.

Lemma 2.8.7. *For all n sufficiently large, the support of ξ^n has greatest common divisor* 1*.*

Proof. By assumption, the support of ξ has greatest common divisor 1, so we can find an M such that the greatest common divisor of the support of ξ restricted to $\{0,\ldots,M\}$ is 1. Since $\gamma n^{1/(4-\eta)} > M$ for n sufficiently large, the result follows. □

Lemma 2.8.8. *As $n \to \infty$,*

$$\mathbf{E}\left[(\xi^n)^j\right] \to \mathbf{E}\left[\xi^j\right] \qquad \text{for } j = 1,2,3 \tag{2.8.12}$$

and

$$|\mathbf{E}\left[\xi^n\right] - 1| = O(n^{-2/3}). \tag{2.8.13}$$

Proof. For $j \in \{1,2,3\}$ we have

$$\begin{aligned}\mathbf{E}\left[(\xi^n)^j\right] &= \sum_{k=1}^\infty k^j \mathbf{P}\left\{\xi^n = k\right\} \\ &\leq \sum_{k=1}^\infty k^j \frac{\mathbf{P}\left\{\xi = k\right\}}{\mathbf{P}\left\{\max_{1\leq i\leq \xi} |Y_{\xi,i}| \leq \gamma n^{1/(4-\eta)}\right\}} = \left(1 + O\left(\frac{1}{n}\right)\right)\mathbf{E}\left[\xi^j\right],\end{aligned}$$

where the final equality follows by assumption [**A3**]. By the bounded convergence theorem, as $n \to \infty$,

$$\mathbf{E}\left[(\xi^n)^j\right] \geq \mathbf{E}\left[\xi^j\right] - \mathbf{E}\left[\xi^j \mathbf{1}_{[\max_{1\leq i\leq \xi} |Y_{\xi,i}|>\gamma n^{1/(4-\eta)}]}\right] \to \mathbf{E}\left[\xi^j\right],$$

where we have used assumption [**A3**] again. (2.8.12) follows.

To get the more precise lower bound for $j = 1$ in (2.8.13), observe that

$$\mathbf{E}\left[\xi \mathbf{1}_{[\max_{1\leq i\leq \xi} |Y_{\xi,i}|>\gamma n^{1/(4-\eta)}]}\right] \leq n^{1/3}\mathbf{P}\left\{\max_{1\leq i\leq \xi} |Y_{\xi,i}| > \gamma n^{1/(4-\eta)}\right\} + \mathbf{E}\left[\xi \mathbf{1}_{[\xi > n^{1/3}]}\right].$$

The first term on the right-hand side of this inequality is $O(n^{-2/3})$ by [**A3**]. Also, $\mathbf{E}\left[\xi^3\right] < \infty$ and so the second term is also $O(n^{-2/3})$, thus establishing (2.8.13). □

The last tool that we need to prove Lemma 2.8.5 is an upper bound on the total variation distance between $\bar{\xi}_1^n$ and $\bar{\xi}$ where $\bar{\xi}$, a sample from the size-biased law of ξ.

Lemma 2.8.9. *Let X be a random variable taking values in $\mathbb{N}$ such that $\mathbf{E}[X] \geq 0$ and $\mathbf{E}[X^3] < \infty$. Let $(\mathcal{E}_n)_{n\geq 1}$ be a sequence of events with $\mathbf{P}\{\mathcal{E}_n\} = 1 - O(1/n)$. Let X_n be distributed as X conditional on $\mathcal{E}_n$. Let $\overline{X}_n$ have the size-biased law of X_n and let $\overline{X}$ have the size-biased law of X. Then,*

$$d_{\mathrm{TV}}(\overline{X}_n, X) = \frac{1}{2}\sum_{k=1}^{\infty} \left|\mathbf{P}\left\{\overline{X}_n = k\right\} - \mathbf{P}\left\{\overline{X} = k\right\}\right| = O(n^{-2/3}).$$

Proof. By definition,

$$\mathbf{P}\left\{\overline{X}_n = k\right\} = \frac{k\mathbf{P}\{X = k, \mathcal{E}_n\}}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]}, \text{ and } \mathbf{P}\left\{\overline{X} = k\right\} = \frac{k\mathbf{P}\{X = k\}}{\mathbf{E}[X]}. \qquad (2.8.14)$$

Since $\mathbf{E}[X^3] < \infty$ we have that $\mathbf{P}\left\{X > n^{1/3}\right\} = o(n^{-1})$ as $n \to \infty$ and so Hölder's inequality yields that

$$\mathbf{E}\left[X\mathbf{1}_{[X>n^{1/3}]}\right] \leq \mathbf{E}\left[X^3\right]^{1/3} \mathbf{P}\left\{X > n^{1/3}\right\}^{2/3} = o(n^{-2/3}).$$

Next,

$$\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n^c]}\right] \leq n^{1/3}\mathbf{P}\{\mathcal{E}_n^c\} + \mathbf{E}\left[X\mathbf{1}_{[X>n^{1/3}]}\right] = O(n^{-2/3}),$$

so that $\mathbf{E}[X\mathbf{1}_{\mathcal{E}_n}] = \mathbf{E}[X] + O(n^{-2/3})$ and the difference between the denominators in (2.8.14) is $O(n^{-2/3})$. It follows that

$$\begin{aligned}
&\sum_{k=1}^{\infty}\left|\mathbf{P}\left\{\overline{X}_n = k\right\} - \mathbf{P}\left\{\overline{X} = k\right\}\right| \\
&\quad = \sum_{k=1}^{\infty}\left|\frac{k\mathbf{P}\left\{X = k,\ \mathcal{E}_n\right\}}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n^c]}\right]} - \frac{k\mathbf{P}\left\{X = k\right\}}{\mathbf{E}\left[X\right]}\right| \\
&\quad \le \frac{1}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]}\left(\sum_{k=1}^{\infty} k\mathbf{P}\left\{X = k, \mathcal{E}_n^c\right\}\right) + O(n^{-2/3}) \\
&\quad \le \frac{1}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]}\left(\sum_{k=1}^{n^{1/3}} k\mathbf{P}\left\{X = k, \mathcal{E}_n^c\right\} + \mathbf{E}\left[X\mathbf{1}_{[X>n^{1/3}]}\right]\right) + O(n^{-2/3}) \\
&\quad \le \frac{n^{1/3}\mathbf{P}\left\{\mathcal{E}_n^c\right\}}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]} + \frac{\mathbf{E}\left[X\mathbf{1}_{[X>n^{1/3}]}\right]}{\mathbf{E}\left[X\mathbf{1}_{[\mathcal{E}_n]}\right]} + O(n^{-2/3}) \\
&\quad = O(n^{-2/3}).
\end{aligned}$$

The first term on the right hand side of the above inequality is $O(n^{-2/3})$ since $\mathbf{P}\left\{\mathcal{E}_n^c\right\} = O(1/n)$. ☐

This lemma has the following corollary.

Corollary 2.8.10. *Let μ be a critical offspring distribution with variance $\sigma^2 \in (0,\infty)$, and let $\nu = (\nu_k)_{k\ge1}$ be such that [**A1**] holds and [**A3**] holds for a given measure π with $\eta \in [0,2)$. Fix $\gamma > 0$. Let ξ denote a random variable with distribution μ and let $\bar{\xi}_1, \bar{\xi}_2, \ldots$ be IID samples from the size-biased law of ξ. For $n \ge 1$ let ξ^n be distributed as ξ, conditioned to not yield a displacement vector with $\max_{1\le i\le\xi^n}|Y_{\xi^n,i}| > \gamma n^{1/(4-\eta)}$. Further, let μ^n denote the distribution of ξ^n, and let $\bar{\xi}_1^n, \bar{\xi}_2^n, \ldots$ be IID samples from the size-biased law of ξ^n. Then for $m = \Theta(\sqrt{n})$,*

$$d_{\mathrm{TV}}((\bar{\xi}_1^n, \ldots, \bar{\xi}_m^n), (\bar{\xi}_1, \ldots, \bar{\xi}_m)) = O(n^{-1/6}).$$

Proof. By [**A3**], $\bar{\xi}_1^n$ is obtained from $\bar{\xi}_1$ by conditioning on an event which occurs with probability $1 - O(1/n)$. Therefore, by Lemma 2.8.9, the total variation distance between $\bar{\xi}_1^n$ and $\bar{\xi}_1$ is $O(n^{-2/3})$. Since $m = \Theta(\sqrt{n})$, the conclusion follows. ☐

We now prove Lemma 2.8.5. Since the proof of this lemma is very similar to that of Lemma 2.4.4 we will be brief.

Proof of Lemma 2.8.5. As in the proof of Lemma 2.4.4, we may assume that there exists $t > 0$ such that $m/\sqrt{n} \to t$ as $n \to \infty$.

Suppose that $k_1, \dots, k_m \in \mathbb{Z}_{\geq 0}$. Then by almost identical techniques to those used to prove (2.4.6) (replacing the local central limit theorem by Lemma 2.8.6), we obtain that

$$\frac{\mathbf{P}\left\{\sum_{i=(r_n+m)+1}^{n} \xi_i^n = n-1-s_n-\sum_{i=1}^{m} k_i\right\}}{\mathbf{P}\left\{\sum_{i=r_n+1}^{n} \xi_i^n = n-1-s_n\right\}}$$
$$= \exp\left(-\left(\left(\frac{1+s_n-r_n+m\sigma^2+\sum_{i=1}^{m}(k_i-(1+\sigma^2))}{\sqrt{2\sigma^2(n-(r_n+m))}}\right)^2+o(1)\right)\right)+o(1). \tag{2.8.15}$$

Recall that, for $i \in [m]$, $\bar{\xi}_i$ is sample from the size-biased distribution of ξ. We claim that instead of substituting $\bar{\xi}_1^n, \dots, \bar{\xi}_m^n$ in the place of $k_1, \dots, k_m$ we can substitute $\bar{\xi}_1, \dots, \bar{\xi}_m$. Indeed, by Corollary 2.8.10, the total variation distance between $\bar{\xi}_1^n, \dots, \bar{\xi}_m^n$ and $\bar{\xi}_1, \dots, \bar{\xi}_m$ tends to 0 as $n \to \infty$. Therefore, by (2.4.8), we obtain that (2.8.15) tends to $\exp(-(t^2\sigma^2)/2)$ in probability as $n \to \infty$. This convergence is analogous to (2.4.8) in the proof of Lemma 2.4.4.

It remains to establish an analogue of (2.4.9), i.e.,

$$\prod_{i=1}^{m}\left(\frac{n-r_n-i+1}{n-1-\sum_{j=1}^{i-1}\bar{\xi}_j^n}\mathbf{E}\left[\xi^n\right]\right) = \mathbf{E}\left[\xi^n\right]^m \prod_{i=1}^{m}\left(\frac{n-r_n-i+1}{n-1-\sum_{j=1}^{i-1}\bar{\xi}_j}\right) \xrightarrow{\text{p}} \exp\left(\frac{t^2\sigma^2}{2}\right), \tag{2.8.16}$$

as $n \to \infty$.

By Lemma 2.8.8,

$$\mathbf{E}\left[\xi^n\right] = \mathbf{E}\left[\xi\right] + O(n^{-2/3}) = 1 + O(n^{-2/3}),$$

and so, since $m = (1+o(1))t\sqrt{n}$, we obtain that $\mathbf{E}\left[\xi^n\right]^m = 1+o(1)$. Therefore (2.8.16) follows from (2.4.9).

We now prove uniform integrability of the family $(\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}^n, \dots, \bar{\xi}_m^n))_{n\geq 1}$. Again, by the generalised Scheffé lemma [63, Theorem 5.12], since $\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n, \dots, \bar{\xi}_m^n) \xrightarrow{\text{p}} 1$ it suffices to show that $\mathbf{E}\left[\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n, \dots, \bar{\xi}_m^n)\right] \to 1$ as $n \to \infty$. By Proposition 2.8.4 with $f \equiv 1$,

$$\mathbf{E}\left[\Theta_{\mu^n}^{n,r_n,s_n}(\bar{\xi}_1^n, \dots, \bar{\xi}_m^n)\right] = \mathbf{P}\left\{N \geq m \ \tau_{r_n}(\Sigma) > m \ \middle| \ \sum_{i=r_n+1}^{n} \xi_i^n = n-1-s_n\right\}, \tag{2.8.17}$$

where $\Sigma = \Sigma_{\vec{Z}}$ with

$$\vec{Z} = (Z_1, \dots, Z_n) = (d_1, \dots, d_{r_n}, \xi_{r_n+1}^n, \dots, \xi_n^n)$$

such that $d_1 = s_n$, and $d_2, \ldots d_{r_n} = 0$. (Indeed, any fixed choice of $d_1, \ldots, d_{r_n}$ with $\sum_{i=1}^{r_n} d_i = s_n$ would suffice.) To see that the probability on the right-hand side of (2.8.17) tends to 1 as $n \to \infty$, first note that $N \overset{\mathrm{d}}{=} \mathrm{Binomial}(n - r_n, 1 - \mu_0)$ where $r_n < n^{\varepsilon}$. So even after conditioning on the event $\{\sum_{i=r_n+1}^{n} \xi_i^n = n - 1 - s_n\}$, which occurs with probability $O(n^{-1/2})$, there are $(1 + o_{\mathbf{P}}(1))(n - r_n)(1 - \mu_0)$ non-zero entries of $(\xi_{r_n+1}^n, \ldots, \xi_n^n)$. Therefore, to prove uniform integrability it remains to show that $\tau_{r_n}(\Sigma) = \omega_{\mathbf{P}}(\sqrt{n})$.

To see this, observe that for any $k \in [n]$,

$$\mathbf{P}\left\{\tau_{r_n}(\Sigma) = k + 1 \mid (Z_{\Sigma(1)}, \ldots, Z_{\Sigma(k)}),\ \tau_{r_n}(\Sigma) \geq k\right\} = \frac{s_n}{\sum_{i=k+1}^{n} Z_{\Sigma(i)}}.$$

Since $\vec{Z}$ contains $(1 + o_{\mathbf{P}}(1))(n - r_n)(1 - \mu_0) + 1$ positive entries, this denominator is $(1 + o_{\mathbf{P}}(1))(n - r_n)(1 - \mu_0) + 1$ uniformly over all $k \leq m = O(\sqrt{n})$, and all labeled random reorderings of $\vec{Z}$. Moreover, since $s_n = o(\sqrt{n})$ by assumption,

$$\mathbf{P}\left\{\tau_{r_n}(\Sigma) = k + 1 \mid \tau_{r_n}(\Sigma) \geq k\right\} = o(n^{-1/2}),$$

uniformly across all $k \leq m$. The claim follows by summing these probabilities over $k \leq m$, since by the above $\mathbf{P}\left\{\tau_{r_n}(\Sigma) > k\right\} = (1 - o(n^{-1/2}))^k$, and in particular for $T > 0$,

$$\mathbf{P}\left\{\tau_{r_n}(\Sigma) > T\sqrt{n}\right\} = (1 - o(n^{-1/2}))^{T\sqrt{n}}, \tag{2.8.18}$$

which tends to 1 as $n \to \infty$. □

2.8.4 Backstage at the hairy tour

To control the restrictions of the discrete snake introduced in the proofs of Theorems 2.1.1, 2.1.4 and 2.1.5 we require a couple of technical lemmas. The first of these results shows that if we truncate the displacements of the discrete snake by $n^{1/(4-\eta)-\delta}$, then the global moments agree with assumption [**A1**] in the limit.

Fix $\eta \in (0, 2]$, and $\delta \in (0, 1/(4 - \eta))$. For $n \geq 1$, and $k \geq 1$ let

$$Y_k^{n,\delta} = (Y_{k,1}^{n,\delta}, \ldots, Y_{k,k}^{n,\delta}) = \begin{cases} (Y_{k,1}, \ldots, Y_{k,k}) & \text{if } \max_{1 \leq j \leq k} |Y_{k,j}| \leq n^{1/(4-\eta)-\delta} \\ 0 & \text{else.} \end{cases}$$

Lemma 2.8.11. *It holds that*

$$\left|\mathbf{E}\left[Y_{\bar{\xi}, U_{\bar{\xi}}}^{n,\delta}\right]\right| = O\left((n^{1/(4-\eta)-\delta})^{1-2(4-\eta)/3}\right).$$

and furthermore as $n \to \infty$,

$$\mathbf{Var}\left(Y_{\bar{\xi}, U_{\bar{\xi}}}^{n,\delta}\right) \to \beta^2.$$

Proof. First, observe that by Hölder's inequality, there exists a constant $c > 0$ such that

$$\mathbf{P}\left\{\max_{1\le i\le\bar{\xi}} |Y_{\bar{\xi},i}| > y\right\} \le \mathbf{E}\left[\xi^3\right]^{1/3} \mathbf{P}\left\{\max_{1\le i\le\xi} |Y_{\xi,i}| > y\right\}^{2/3} \le cy^{-2(4-\eta)/3}.$$

Then, by global centering

$$\begin{aligned}
\left|\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right]\right| &= \left|\mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\mathbf{1}_{[\max_{1\le i\le\bar{\xi}}|Y_{\bar{\xi},i}|>n^{1/(4-\eta)-\delta}]}\right]\right| \\
&\le \int_0^\infty \mathbf{P}\left\{\left|Y_{\bar{\xi},U_{\bar{\xi}}}\mathbf{1}_{[\max_{1\le i\le\bar{\xi}}|Y_{\bar{\xi},i}|>n^{1/(4-\eta)-\delta}]}\right| > y\right\} dy \\
&\le n^{1/(4-\eta)-\delta}\mathbf{P}\left\{\max_{1\le i\le\bar{\xi}} |Y_{\bar{\xi},i}| > n^{1/(4-\eta)-\delta}\right\} \\
&\quad + \int_{n^{1/(4-\eta)-\delta}}^\infty \mathbf{P}\left\{\max_{1\le i\le\bar{\xi}} |Y_{\bar{\xi},i}| > y\right\} dy \\
&\le cn^{1/(4-\eta)-\delta}(n^{1/(4-\eta)-\delta})^{-2(4-\eta)/3} - \frac{c}{2(4-\eta)/3-1}\left[y^{-2(4-\eta)/3+1}\right]_{n^{1/(4-\eta)-\delta}}^{\infty} \\
&= O\left((n^{1/(4-\eta)-\delta})^{1-2(4-\eta)/3}\right),
\end{aligned}$$

as claimed.

As for the variance,

$$\begin{aligned}
\mathbf{Var}\left(Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right) &= \mathbf{E}\left[\left(Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right)^2\right] - \left(\mathbf{E}\left[Y^{n,\delta}_{\bar{\xi},U_{\bar{\xi}}}\right]\right)^2 \\
&= \mathbf{E}\left[Y^2_{\bar{\xi},U_{\bar{\xi}}}\mathbf{1}_{[\max_{1\le i\le\bar{\xi}}|Y_{\bar{\xi},i}|\le n^{1/(4-\eta)-\delta}]}\right] - \mathbf{E}\left[Y_{\bar{\xi},U_{\bar{\xi}}}\mathbf{1}_{[\max_{1\le i\le\bar{\xi}}|Y_{\bar{\xi},i}|\le n^{1/(4-\eta)-\delta}]}\right]^2 \\
&\to \mathbf{E}\left[Y^2_{\bar{\xi},U_{\bar{\xi}}}\right] = \beta^2,
\end{aligned}$$

as $n \to \infty$, by dominated convergence and the result for the mean. □

The above lemma pertains to snakes where the displacements which are above $n^{1/(4-\eta)-\delta}$ are all set to 0. The next lemma in this section will help us to understand the asymptotics of the head of the discrete snake where displacements which are *below* $n^{1/(4-\eta)-\delta}$ are set to 0. More specifically, we present a tail bound for the size of a set of marked vertices in random trees, which we apply in Proposition 2.5.5 where the marked vertices pertain to vertices $v \in v(\mathrm{T}_n) \setminus \partial\mathrm{T}_n$ for which $\|Y^{(v)}\|_\infty > n^{1/(4-\eta)-\delta}$.

Lemma 2.8.12. *Let* $\mathrm{d} = (d_1, \ldots, d_n)$ *be a degree sequence, fix* $\mathcal{B} \subset [n]$ *and write* $K = |\mathcal{B}|$*, and* $\Delta = \max_{1\le i\le n} d_i$*. Let* B_{d} *be the smallest distance between two vertices in* $\mathcal{B}$ *that are ancestrally related in* $T_{\mathrm{d}} = B(\Pi_{\mathrm{d}})$ *(with* $B_{\mathrm{d}} = \infty$ *if no vertices in* $\mathcal{B}$ *are ancestrally related). Then, for any* $b \ge 0$

$$\mathbf{P}\{B_{\mathrm{d}} \le b\} \le K\left(1 - \left(1 - \frac{K\Delta}{n-1-b\Delta}\right)^b\right).$$

Proof. It suffices to show the statement for integer b since for general b, $\mathbf{P}\{B_{\mathrm{d}} \le b\} = \mathbf{P}\{B_{\mathrm{d}} \le \lfloor b \rfloor\}$ and the upper bound is increasing in b.

Fix a degree sequence and a set $\mathcal{B}$. Without loss of generality, assume that $\mathcal{B} = \{n-K+1, \dots, n\}$.

For $v \in [n]$, let $p(v)$ be the parent of v in $\mathrm{T}_{\mathbf{d}}$ (with $p(v) = v$ if v is the root of $\mathrm{T}_{\mathbf{d}}$). Also set $p^0(v) = v$ and recursively for $k \ge 1$ define the k-th ancestor of v as $p^k(v) = p(p^{k-1}(v))$.

We will show that

$$\mathbf{P}\left\{\{p^1(n), \dots, p^b(n)\} \cap \{n-K+1, \dots, n-1\} = \emptyset\right\} \ge \left(1 - \frac{K\Delta}{n-1-b\Delta}\right)^b, \tag{2.8.19}$$

after which the statement follows by symmetry and the union bound.

We will prove (2.8.19) by induction. To ease notation, write $p^k = p^k(n)$ for $k \ge 0$. For (i,c) such that $i \in [n], c \in [d_i]$ write $\Pi_{\mathrm{d}}^{-1}(i,c)$ for the position of (i,c) in Π_{d}.

We will define a sequence of σ-algebras $(\mathcal{F}_k)_{k\ge 0}$ such that for each $k \ge 1$, $\mathcal{F}_k$ is the σ-algebra generated by the first k ancestors of n and the positions of their corresponding entries in Π_{d}. Let, $\mathcal{F}_0 = \sigma(\Pi_{\mathrm{d}}^{-1}(n,c) : c \in [d_n])$ contain the information on the position of vertex n in Π_{d}.

If $d_n = 0$ and $\{\Pi_{\mathrm{d}}^{-1}(n,c) : c \in [d_n]\} = \emptyset$ then n is the final vertex in the final path of the line-breaking construction, and the last entry of Π_{d} gives its parent. Thus, in this case, we reveal $\Pi_{\mathrm{d}}(n-1)$ and we have $\Pi_{\mathrm{d}}(n-1) = (p^1, c')$ for some $c' \in [d_{p^1}]$. Then, we reveal all other entries of the form $(p^1, c), c \in [d_{p^1}]$ in Π_{d} and this yields $\mathcal{F}_1$.

If $d_n > 0$, then set $m_0 = \min\{\Pi_{\mathrm{d}}^{-1}(n,c) : c \in [d_n]\}$. If $m_0 = 1$ then n is the root of $\mathbf{T}_{\mathrm{d}}$ so $p^\ell = n$ for all $\ell \ge 1$, and so we let $\mathcal{F}_\ell = \mathcal{F}_0$ for all $\ell \ge 1$. Otherwise, the entry before the first occurrence of an entry of the form $(n,c), c \in [d_n]$ in Π_{d} gives the parent of n so then we obtain $\mathcal{F}_1$ as follows. We reveal $\Pi_{\mathrm{d}}(m_0 - 1)$. In that case $\Pi_{\mathrm{d}}(m_0 - 1) = (p^1, c')$ for some $c' \in [d_{p^1}]$. Secondly, we reveal all other entries of the form $(p^1, c), c \in [d_{p^1}]$ in Π_{d} and this yields $\mathcal{F}_1$.

For $k \ge 1$, given $\mathcal{F}_k$, let

$$m_k = \min\{\Pi_{\mathrm{d}}^{-1}(p^k, c) : c \in [d_{p^k}]\}.$$

If $m_k = 1$ then p^k is the root of $\mathbf{T}_{\mathrm{d}}$ so $p^\ell = p^k$ and we take $\mathcal{F}_\ell = \mathcal{F}_k$ for all $\ell \ge k$. If $m_k > 1$, we obtain $\mathcal{F}_{k+1}$ as follows. First, we reveal $\Pi_{\mathrm{d}}(m_k - 1)$. In that case $\Pi_{\mathrm{d}}(m_k - 1) = (p^{k+1}, c')$ for some $c' \in [d_{p^{k+1}}]$. Secondly, we reveal all other entries of the form $(p^{k+1}, c), c \in [d_{p^{k+1}}]$ in Π_{d} and this yields $\mathcal{F}_{k+1}$.

Now, observe that, for $k \ge 0$, given $\mathcal{F}_k$, the unrevealed entries of Π_{d} occur in an order given by a uniformly random permutation. So given $\mathcal{F}_k$ if $m_k > 1$ the k-th ancestor of n is the first coordinate of a uniformly random sample from

$$\{(i,c) : i \in [n] \backslash \{p^0, \ldots, p^k\}, c \in [d_i]\}$$

and

$$\begin{aligned}
&\mathbf{P}\left\{p^{k+1} \in \{n-K+1, \ldots, n-1\} \;\middle|\; \mathcal{F}_k, \{p^1, \ldots, p^k\} \cap \{n-K+1, \ldots, n-1\} = \emptyset\right\} \\
&\quad = \frac{d_{n-K+1} + \cdots + d_{n-1}}{n-1-\sum_{i=0}^k d_{p^i}} \le \frac{K\Delta}{n-1-(k+1)\Delta}.
\end{aligned}$$

If $m_k = 1$ then the conditional probability above is 0 so the inequality also holds.

Therefore, we see inductively that

$$\begin{aligned}
\mathbf{P}\left\{\{p^1, \ldots, p^b\} \cap \{n-K+1, \ldots, n-1\} = \emptyset\right\} &\ge \prod_{k=1}^{b} \left(1 - \frac{K\Delta}{n-1-k\Delta}\right) \\
&\ge \left(1 - \frac{K\Delta}{n-1-b\Delta}\right)^b,
\end{aligned}$$

and so the result follows. □

Chapter 3

Temporal connectivity of random geometric graphs

This chapter is based on joint work with Anna Brandenberger (MIT), Serte Donderwinkel (University of Groningen), Céline Kerriou (Universität zu Köln) and Gábor Lugosi (Pompeu Fabra University) which appears in the preprint [20].

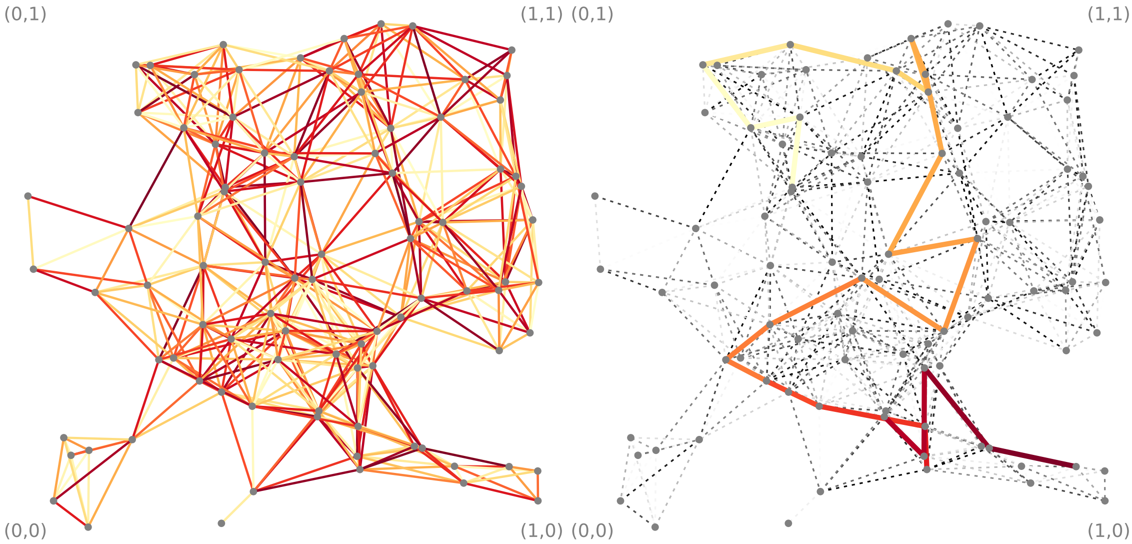

Figure 3.1: A temporal random geometric graph on $[0,1]^2$ where edges that are coloured darker arrive later (left), and its longest monotone increasing path (right).

3.1 Introduction

A *temporal graph* is an edge-labeled graph $\mathcal{G} = (G, \sigma)$ where the underlying graph $G = (V(G), E(G))$ is a finite simple graph and $\sigma : E(G) \to \{1, \dots, |E(G)|\}$ is an ordering of the edges. In this paper we consider random temporal graphs in which $\sigma(E(G))$ is a uniform random permutation. Equivalently, one can generate a random temporal graph using independent, uniform, labels: each edge $e \in E(G)$ is assigned a label $\tau_e \sim \mathrm{Unif}[0,1]$. For $e \in E(G)$ we say that τ_e is the *time-stamp* of e, and for $f \in E(G) \setminus \{e\}$, we say that e *precedes* f if $\tau_e \leq \tau_f$. We use this construction in the sequel, and denote the resulting graph as $\mathcal{G} = (G, (\tau_e)_{e\in E(G)})$.

The study of such graphs is largely motivated by the modeling of dynamic networks, see Tang et al. [104]. A temporal graph represents a network where interactions occur at particular times. In temporal networks one may study whether it is possible to transmit information, or an infectious disease, from one individual to another, as a chain of transmissions can only occur along interactions that are increasing in time.

Let $\mathcal{G} = (G, (\tau_e)_{e\in E(G)})$ be a temporal graph. For vertices $u, v \in V(G)$ we say that there is a *temporal path* from u to v, denoted $u \xrightarrow{\mathcal{G}} v$, if there exists a path from u to v comprised of non-decreasing edge time-stamps. Further, we let $\ell(u \xrightarrow{\mathcal{G}} v)$ denote the length of the shortest temporal path from u to v, that is, the minimal number of edges on any monotone increasing path from u to v. It is readily seen that the existence of a temporal path $u \xrightarrow{\mathcal{G}} v$ does not imply the existence of a temporal path $v \xrightarrow{\mathcal{G}} u$, and furthermore the existence of temporal paths $u \xrightarrow{\mathcal{G}} v$ and $v \xrightarrow{\mathcal{G}} w$ does not imply the existence of a temporal path $u \xrightarrow{\mathcal{G}} w$. Following the terminology of [17], we say that a vertex $u \in \mathcal{G}$ is a *temporal source* if there exist temporal paths $u \xrightarrow{\mathcal{G}} v$ for all $v \in \mathcal{G}$. Moreover, we call $\mathcal{G}$ *temporally connected* if every vertex of $\mathcal{G}$ is a temporal source.

The temporal random graph model can be seen as a random version of a well-known combinatorial problem posed by Chvátal and Komlós [34], which asks for the minimal value of the length of the longest monotone path, when considering the edge orderings of the complete graph on n vertices, henceforth denoted by K_n. This problem is well studied, with lower and upper bounds remaining far apart until recent years, see, e.g., Graham and Kleitman [49], Calderbank, Chung, and Sturtevant [28], Bucić et al. [26]. The average case with base graph $G = K_n$, was first considered by Lavrov and Loh [74], and Martinsson [85] later proved that with high probability, there exists a monotone path of length $n - 1$. When G is an Erdős–Rényi random graph, the resulting temporal graph is called the *random simple temporal graph*, first studied by Angel, Ferber, Sudakov and Tassion [12].

Temporal connectivity of random simple temporal graphs was studied in detail by Casteigts, Raskin, Renken and Zamarev [30] and Becker et al. [17]. In particular, it is shown in [30] that the threshold for temporal connectivity of random simple temporal graphs is $\sim 3 \log n/n$, which is just a factor of 3 larger than the connectivity threshold of the underlying Erdős-Rényi random graph. More recent works have studied the shortest and longest increasing paths between typical vertices (Broutin, Kamčev, and Lugosi [22]) and the size of the largest temporal clique (Atamanchuk, Devroye, and Lugosi [14, 17]).

In this paper, we initiate the study of *temporal random geometric graphs*. In this model the vertices of the underlying graph G_n correspond to randomly drawn points in a Euclidean space. Such graphs may be better suited for modeling epidemiological processes than random simple temporal graphs, as the geometry allows for spatial closeness (or closeness encoding similarity) of individuals to affect interaction probabilities. Our main results establish thresholds for temporal connectivity in certain temporal random geometric graphs.

For $d \geq 2$, let $\|\cdot\|$ denote the Euclidean distance on the d-dimensional unit torus, $[0,1]^d$. Let $K : [0,\infty) \to [0,1]$ be a non-increasing function. Further, let $n \in \mathbb{N}$ and $r_n > 0$. A *soft random geometric graph* $G_n = (\mathcal{X}_n, K, r_n)$ in dimension d is a graph on n uniform random points $\mathcal{X}_n$ in the unit torus $[0,1]^d$, such that for every pair of vertices $u, v \in \mathcal{X}_n$, an edge exists with probability

$$\mathbf{P}\left\{uv \in E(G_n) \mid \mathcal{X}_n\right\} = K\left(\frac{\|u-v\|}{r_n}\right).$$

The edges are conditionally independent given $\mathcal{X}_n$. When $K(x) = \mathbf{1}_{\{x \leq 1\}}$, G_n is called a *hard random geometric graph*. A *temporal random geometric graph* with n vertices is a temporal graph $\mathcal{G}_n = (G_n, (\tau_e)_{e \in E(G_n)})$ where $G_n = (\mathcal{X}_n, K, r_n)$ is a soft random geometric graph on n vertices. Our main result shows that under regularity conditions on K, with high probability, the graph becomes temporally connected when the radius r_n exceeds a constant multiple of $n^{-1/(d+1)}$, while the graph is temporally disconnected if r_n is smaller than another constant times $n^{-1/(d+1)}$.

Theorem 3.1.1. *Let $\mathcal{G}_n = (G_n, (\tau_e)_{e \in E(G_n)})$ be a temporal random geometric graph with $G_n = (\mathcal{X}_n, K, r_n)$ satisfying the following:*

[**A1**] *For $x > 1$, $K(x) \leq \beta x^{-d} e^{-2(x+1)\log(x+1)}$ for some $\beta > 0$.*

[**A2**] *There exists $\alpha > 0$ such that $K(x) \geq \alpha$ for all $x \leq 1$.*

Then, there exist constants $c_d, C_d > 0$ such that

(i) if $r_n \leq c_d n^{-1/(d+1)}$ for all n sufficiently large, then $\mathcal{G}_n$ is a.a.s. temporally disconnected,

(ii) if $r_n \geq C_d n^{-1/(d+1)}$ for all n sufficiently large, then $\mathcal{G}_n$ is a.a.s. temporally connected.

Observe that by monotonicity of temporal connectivity, Theorem 3.1.1 implies that (i) holds under just assumption [**A1**] and (ii) holds under just assumption [**A2**].

Note that the choice of $K(x) = \mathbf{1}_{\{x \leq 1\}}$ satisfies the regularity conditions [**A1**] and [**A2**], and therefore as a corollary of Theorem 3.1.1 we obtain a threshold for a.a.s. temporal connectivity of temporal hard random geometric graphs. The assumptions [**A1**] and [**A2**] ensure that there are not too many long edges, and there are sufficiently many vertices at graph distance $O(r_n)$ to any vertex in the soft random geometric graph.

Observe that the order of magnitude $n^{-1/(d+1)}$ of the critical radius is significantly larger than that of the critical radius for the connectivity of the underlying random geometric graph. Indeed, as it is well known from the theory of random geometric graphs, random geometric graphs become connected when the average degree of the graph is of the order $\log n$ (see [93]). On the other hand, our results show that temporal connectivity only occurs when the average degree becomes of the order of $n^{1/(d+1)}$. This is in stark contrast with Erdős-Rényi random graphs in which temporal connectivity and simple connectivity both occur when the average degree is of the order of $\log n$. The regime $r_n \approx n^{-1/(d+1)}$ is somewhat unusual, as we are not aware of any other natural phase transition happening in random geometric graphs in this range.

Theorem 3.1.1 establishes the order of magnitude of the temporal connectivity threshold. It is natural to conjecture that, similarly to the case of Erdős-Rényi graphs, there is a sharp threshold for temporal connectivity, that is, there exists a constant κ_d (depending on the kernel K) such that, for all $\varepsilon > 0$, if $r_n \leq (\kappa_d - \varepsilon)n^{-1/(d+1)}$, then $\mathcal{G}_n$ is a.a.s. temporally disconnected, while for $r_n \geq (\kappa_d + \varepsilon)n^{-1/(d+1)}$, $\mathcal{G}_n$ is a.a.s. temporally connected. Our techniques are not sufficiently strong to establish such results and we leave the problem of existence of a sharp threshold and the value of κ_d for challenging future research.

Becker et al. [17] introduce various definitions of temporal connectivity. In particular, for the case of random simple temporal graphs, they establish sharp thresholds for the probability of the existence of a temporal path between two typical vertices and for the probability of a typical vertex being a temporal source. It is an interesting question to study these probabilities, along with the probability of temporal connectivity. As it is apparent from our proofs, the critical radius for all these quantities is of the same order of magnitude, but we have no further information about their relationship.

The random geometric graph considered in this paper is defined on the unit torus. One can similarly set up the problem by considering the hypercube $[0, 1]^d$ with the usual Euclidean distance. Both setups are quite similar, and our choice makes it possible to avoid some technicalities arising from edge effects. We believe that is is quite straightforward to modify the proofs to extend the result of Theorem 3.1.1 to the Euclidean setting.

One may also consider the problem when $d = 1$. This case is considerably less complex, and a straightforward simplification of the proof of Theorem 3.1.1 shows that the theorem also holds in the one-dimensional case (i.e., the threshold for temporal connectivity is of the order of $n^{-1/2}$).

In Section 3.2 we prove the upper bound of Theorem 3.1.1, that is, we determine a threshold for when $\mathcal{G}_n$ is a.a.s. temporally disconnected. In Section 3.3 we prove the lower bound of Theorem 3.1.1 by mapping the construction of a temporal path to a directed percolation model.

3.2 Upper bound: $\mathcal{G}_n$ temporally disconnected

Let $\mathcal{G}_n = (G_n, (\tau_e)_{e \in E(G_n)})$ be a temporal random geometric graph with $G_n = (\mathcal{X}_n, K, r_n)$ satisfying assumptions [**A2**] and [**A1**]. In this section, we will prove that there exists $c > 0$ such that if $r_n \le cn^{-1/(d+1)}$ for all n sufficiently large, then a.a.s. $\mathcal{G}_n$ has a vertex that is not a temporal source (and in particular $\mathcal{G}_n$ is not temporally connected).

In fact, the following proposition shows that not only does there a.a.s. exist a vertex that is not a temporal source, but that in fact a.a.s. *no vertex* is a temporal source in this regime.

Proposition 3.2.1. *Suppose that $\mathcal{G}_n = (G_n, (\tau_e)_{e \in E(G_n)})$ is a temporal random geometric graph with $G_n = (\mathcal{X}_n, K, r_n)$ satisfying assumptions* [**A2**] *and* [**A1**]. *Then there exists $c > 0$ such that if $r_n \le cn^{-1/(d+1)}$ for all n sufficiently large, then*

$$\mathbf{P}\left\{\exists u \in \mathcal{X}_n \text{ such that } \forall v \in \mathcal{X}_n,\, u \xrightarrow{\mathcal{G}_n} v\right\} = o(1). \qquad (3.2.1)$$

Proof. Assume $r_n = cn^{-1/(d+1)}$ for some $c > 0$. We show that (3.2.1) holds, and the proposition then follows by monotonicity of temporal connectivity.

The event of having a temporal source is contained in the union of the following three events: all points $\mathcal{X}_n$ are contained in a ball of radius at most $\sqrt{d}(1 - r_n)/2$ around some point in $\mathcal{X}_n$, there exists a long temporal path of length at least $1/3r_n$, or there exists a short temporal path between two points at distance at least

$\sqrt{d}(1-r_n)/2$. More precisely,

$$
\begin{aligned}
&\mathbf{P}\left\{\exists u \in \mathcal{X}_n \text{ such that } \forall v \in \mathcal{X}_n, u \xrightarrow{\mathcal{G}_n} v\right\} \\
&\le \mathbf{P}\left\{\exists u \in \mathcal{X}_n \text{ such that } \forall v \in \mathcal{X}_n, \|u-v\| < \frac{\sqrt{d}}{2}(1-r_n)\right\} \\
&\quad + \mathbf{P}\left\{\exists u, v \in \mathcal{X}_n \text{ such that } u \xrightarrow{\mathcal{G}_n} v, \ell(u \xrightarrow{\mathcal{G}_n} v) \ge \frac{1}{3r_n}\right\} \\
&\quad + \mathbf{P}\left\{\exists u, v \in \mathcal{X}_n \text{ such that } \|u-v\| \ge \frac{\sqrt{d}}{2}(1-r_n), u \xrightarrow{\mathcal{G}_n} v, \ell(u \xrightarrow{\mathcal{G}_n} v) < \frac{1}{3r_n}\right\}.
\end{aligned}
\tag{3.2.2}
$$

Thus, it suffices to show that each of the terms on the right-hand side of (3.2.2) are $o(1)$.

Note that by applying a union bound to the first term of (3.2.2), we may consider each fixed vertex $u \in \mathcal{X}_n$ to be the point $(1/2, 1/2)$ without loss of generality. Therefore, the first term is at most

$$
\begin{aligned}
n \cdot \mathbf{P}\left\{\forall v \in \mathcal{X}_n, \left\|\left(\frac{1}{2}, \frac{1}{2}\right) - v\right\| < \frac{\sqrt{d}}{2}(1-r_n)\right\} &\le n \cdot \mathbf{P}\left\{\left[1-\frac{r_n}{2}, 1\right]^d \cap \mathcal{X}_n = \emptyset\right\} \\
&= n\left(1-\left(\frac{r_n}{2}\right)^d\right)^n \\
&\le \exp\left(\log n - n\left(\frac{r_n}{2}\right)^d\right) \\
&= o(1),
\end{aligned}
\tag{3.2.3}
$$

where the final equality holds as $r_n = cn^{-1/(d+1)} \gg (\log n/n)^{1/d}$ for any $c > 0$.

To bound the second term, we apply the first moment method. To this end, without loss of generality, assume that $1/3r_n \in \mathbb{N}$, let $k = 1/3r_n$, and let $N_n^{(k)}$ denote the number of temporal paths in $\mathcal{G}_n$ with at least $k-1$ edges. By choosing k ordered vertices and ensuring that the edges between them exist and have monotone increasing labels, it follows that

$$
\mathbf{E}\{N_n^{(k)}\} \le \binom{n}{k} k! \frac{1}{(k-1)!}\left(c'_d r_n^d + \int_{[0,1]^d \setminus \mathcal{B}(0,r_n)} K(\|y\|/r_n) dy\right)^{k-1},
$$

where $c'_d > 0$ is the volume of the unit ball in the d-dimensional torus. By our conditions on K, the integral in the above inequality is at most a constant multiple, $\gamma_d > 0$, of r_n^d. We note that both c'_d and γ_d may be uniformly bounded over all $d \ge 2$. Therefore, together with the bound $\binom{n}{k} \le \left(\frac{en}{k}\right)^k$, it follows that there exists $\lambda > 0$

such that

$$\begin{aligned}\mathbf{E}\left[N_n^{(k)}\right] &\le \lambda^{k-1}\left(\frac{en}{k}\right)^k k r_n^{d(k-1)}\\ &= \lambda^{k-1}\exp\left(k\log n + k - (k-1)\log k + d(k-1)\log(r_n)\right)\\ &\le \lambda^{k-1}\exp\left(k\log n + k - (k-1)\log k - d(k-1)\log k\right)\\ &\le \lambda^{k-1}\exp\left(k\log n + k\left(1-(d+1)\frac{k-1}{k}\log k\right)\right). \end{aligned} \tag{3.2.4}$$

Given our choice of $r_n = cn^{-1/(d+1)}$ for some $c > 0$, for n sufficiently large we have that $\log k = \log(1/3r_n) \ge \log(1/3c) + \log(n)/(d+1)$ and so (3.2.4) is at most

$$\lambda^{k-1}\exp\left(\log n + \left(1-(d+1)\log\left(\frac{1}{3c}\right)\right)k + (d+1)\log\left(\frac{1}{3c}\right)\right), \tag{3.2.5}$$

which is $o(1)$ for sufficiently small $c > 0$. Since any temporal path of length at least $k = 1/3r_n$ must include a sub-path of length $k-1$, combining (3.2.5) with the first moment method yields

$$\mathbf{P}\left\{\exists u, v \in \mathcal{X}_n \text{ such that } u \xrightarrow{\mathcal{G}_n} v, \ell(u \xrightarrow{\mathcal{G}_n} v) \ge \frac{1}{3r_n}\right\} = o(1). \tag{3.2.6}$$

From (3.2.3) and (3.2.6), we have that the first two terms in (3.2.2) are $o(1)$. Thus, it remains to prove that connectivity via short paths is unlikely for vertices which are geometrically far apart, i.e., to show that

$$\mathbf{P}\left\{\exists u, v \in \mathcal{X}_n : \|u-v\| \ge \frac{\sqrt{d}}{2}(1-r_n), u \xrightarrow{\mathcal{G}_n} v, \ell(u \xrightarrow{\mathcal{G}_n} v) < \frac{1}{3r_n}\right\} = o(1). \tag{3.2.7}$$

First, note that we can work under the event that for every edge $uv \in E(G_n)$, $\|u-v\| \le r_n \log n$, since the probability that there exists an edge of length greater than $r_n \log n$ is, by a union bound, at most

$$n^2\mathbf{P}\left\{uv \in E(G_n) \mid u, v \in \mathcal{X}_n\right\} \le n^2 K(\log n) = O\left(n^2 e^{-\log n \log\log n}\right) = o(1),$$

for u, v satisfying $\|u-v\| > r_n \log n$.

For sufficiently large n, we have $\sqrt{d}(1-r_n)/2 \ge 1/3$ so we can bound the left-hand side of (3.2.7) from above by the probability that there exists a temporal path $u \xrightarrow{\mathcal{G}_n} v$ with $M \in ((3r_n \log n)^{-1}, (3r_n)^{-1})$ edges, such that the sum of its edge lengths is at least $1/3$. We denote this event by A_n^M. For $L \in \mathbb{N}$ and $M \in ((3r_n \log n)^{-1}, (3r_n)^{-1})$, let $N_{M,L}$ denote the number of temporal paths $u \xrightarrow{\mathcal{G}_n} v$ with M edges with respective lengths $d_1, \dots, d_M$ such that $\sum_{i=1}^M \lceil d_i r_n^{-1}\rceil = L$. Then, by a union bound,

$$\mathbf{P}\left\{A_n^M\right\} \le \sum_{L \ge r_n^{-1}/3} \mathbf{E}\left[N_{M,L}\right].$$

To prove (3.2.7) it therefore suffices to show that $\sum_{L\geq(3r_n)^{-1}} \mathbf{E}[N_{M,L}] = o(1)$.

To this end, consider a path in G_n with M edges of respective lengths $d_1,\ldots,d_M$ such that $\sum_{i=1}^{M} d_i \geq 1/3$. These edge lengths correspond to a vector $(\ell_i)_{i\in[M]}$ where for all $i \in [M]$, $\ell_i := \lceil d_i r_n^{-1}\rceil \in \mathbb{N}$. With $L = \sum_{i=1}^{M} \ell_i$, it holds that $L \geq (3r_n)^{-1}$.

For a given L, the number of such vectors is bounded from above by

$$\binom{L-1}{M-1} \leq \left(\frac{eL}{M}\right)^M = e^{M(\log(L/M)+1)}.$$

Recall that the volume of the unit ball in $\mathbb{R}^d$, c'_d, is uniformly bounded above by λ. Fixing $(\ell_i)_{i\in M}$ and a set of distinct vertices $(v_i)_{i\in[M+1]}$, the probability that the temporal path $v_1 \xrightarrow{\mathcal{G}_n} v_{M+1}$ exists in G_n, i.e., that $v_i v_{i+1} \in E(G_n)$ for each $i \in [M]$, can be bounded from above by

$$\begin{aligned}
\prod_{i=1}^{M} \lambda \int_{(\ell_i-1)r_n}^{\ell_i r_n} K(\xi/r_n)\xi^{d-1}d\xi &= (\lambda r_n^d)^M \prod_{i=1}^{M}\left(\int_{\ell_i-1}^{\ell_i} K(x)x^{d-1}dx\right)\\
&= O\left((\lambda r_n^d)^M \prod_{i=1}^{M} e^{-2\ell_i\log(\ell_i)}\right)\\
&= O\left((\lambda r_n^d)^M \exp\left(-2L\sum_{i=1}^{M}\frac{\ell_i}{L}\log(\ell_i/L) - 2\sum_{i=1}^{M}\ell_i\log(L)\right)\right)\\
&= \ O\left((\lambda r_n^d)^M \exp(-2L\log(L/M))\right), \qquad (3.2.8)
\end{aligned}$$

where in the last step we use the fact that $-\sum_{i=1}^{M}\frac{\ell_i}{L}\log(\ell_i/L) \leq \log M$.

Union bounding over all choices of $(v_i)_{i\in[M+1]}$ and $(\ell_i)_{i\in[M]}$ we obtain

$$\mathbf{E}[N_{M,L}] = \ O\left(\binom{n}{M+1}(M+1)!\frac{1}{M!}(\lambda r_n^d)^M e^{M+M\log(L/M)-2L\log(L/M)}\right),$$

where we have used the fact that the paths counted in $N_{M,L}$ contain $M+1$ ordered vertices between which the edges exist and are ordered with monotone increasing edge labels. By similar computations as those used in (3.2.4) and the fact that $r_n = cn^{-1/(d+1)}$ we obtain that $\mathbf{E}[N_{M,L}]$ is

$$\begin{aligned}
&O\left(\exp\left((M+1)\log(n)+M(2+\log(\lambda)+d\log(c))+\frac{dM\log(n)}{d+1}+(M-2L)\log\left(\frac{L}{M}\right)\right)\right)\\
&= O\left(\exp\left(\log(n) + M(2+\log(\lambda)+d\log(c)+\log(3c)) + 2(M-L)\log\left(\frac{L}{M}\right)\right)\right)\\
&= O\left(\exp\left(\log(n) + M(2+\log(\lambda)+d\log(c)+\log(3c))\right)\left(\frac{L}{M}\right)^{2(M-L)}\right),
\end{aligned}$$

where the second line holds as $L \geq (3r_n)^{-1} = n^{1/(d+1)}/3c$.

Note that summing over $L \geq (3r_n)^{-1}$, the second term of the above yields

$$\sum_{L \geq (3r_n)^{-1}} \left(\frac{L}{M}\right)^{2(M-L)} \leq \sum_{L > M} \left(\frac{L}{M}\right)^{2(M-L)} = \sum_{k=1}^{\infty} \left(\frac{M+k}{M}\right)^{-2k} \leq \sum_{k=1}^{\infty} \left(1 + \frac{1}{M}\right)^{-2k} < M,$$

where the first inequality uses that $M < (3r_n)^{-1}$. Therefore,

$$\sum_{L \geq (3r_n)^{-1}} \mathbf{E}\left[N_{M,L}\right] = e^{(1+o(1))M(2+\log(\lambda)+\log(3c)+d\log(c))},$$

for $M \in ((3r_n \log n)^{-1}, (3r_n)^{-1})$. The result follows as this is $o(1)$ for sufficiently small $c > 0$. □

3.3 Lower bound: $\mathcal{G}_n$ temporally connected

In this section we show that there exists a constant $C_d > 0$ such that if $r_n \geq C_d n^{-1/(d+1)}$ for all n sufficiently large, then $\mathcal{G}_n$ is a.a.s temporally connected.

By monotonicity, to prove Theorem 3.1.1 3.1.1, it suffices to prove the statement for a hard random geometric graph where each edge is retained with probability $\alpha \in (0, 1]$. That is for $\mathcal{G}_n = (G_n, (\tau_e)_{e \in E(G_n)})$, with $G_n = (\mathcal{X}_n, K, r_n)$ such that $K(x) := \alpha \mathbf{1}_{\{x \leq 1\}}$. More specifically, we prove the following proposition.

Proposition 3.3.1. *Fix $\alpha \in (0, 1]$. For all $n \geq 1$ let $\mathcal{G}_n = (G_n, (\tau_e)_{e \in E(G_n)})$ be a temporal random geometric graph with $G_n = (\mathcal{X}_n, K, r_n)$ such that $K(x) = \alpha \mathbf{1}_{\{x \leq 1\}}$ for all $x \geq 0$. Then there exists $C_d > 0$ such that if for all n sufficiently large, $r_n \geq C_d n^{-1/(d+1)}$, then*

$$\forall u, v \in \mathcal{X}_n,\ u \xrightarrow{\mathcal{G}_n} v = 1 - o(1).$$

We present the proof in the specific case where $d = 2$, and discuss the extension to general dimensions in subsection 3.3.2.1. In this case, we henceforth simplify notation and let $C = C_2 > 0$. We also assume that for all n sufficiently large, $r_n = Cn^{-1/3}$ (since the case $r_n \geq Cn^{-1/3}$ follows by monotonicity).

To prove Proposition 3.3.1, we construct a temporal path between any two vertices $u, v \in \mathcal{X}_n$. First, we present the construction for the two furthest possible separated points on the torus $[0, 1]^2$, namely $u = (0, 0)$ and $v = (1/2, 1/2)$. An analogous construction then holds for any pair of points, and the result follows by a union bound over all $u, v \in \mathcal{X}_n$.

To construct a temporal path from $(0, 0)$ to $(1/2, 1/2)$, we split $[0, 1/2]^2$ into three regions, see Figure 3.3 for an illustration. Fix $\varepsilon \in (0, 1/3)$. In the first region (the red region in Figure 3.3), we construct many monotone increasing paths from $(0, 0)$ and ending at a point close to the diagonal $x + y = (4\sqrt{2})^{-1} n^{\varepsilon} r_n$. The monotone paths in this region are such that all edges have labels that are at most 1/4, and the label of the m-th edge in each path falls in an interval $T_m \subset [0, 1]$ with $|T_m| \approx n^{-\varepsilon}$. The size of each interval T_m is generous, and this generosity is such that if continued for too long, we would run out of admissible labels before reaching $(1/2, 1/2)$. Therefore, in the second region, (the white region in Figure 3.3), we build onto the paths constructed in the first region, if possible, along edges whose labels are more restricted, that is, ranging from 1/4 to 3/4, and falling in intervals of length $\approx n^{-1/3}$. This construction is such that with high probability, at least one of the paths from the first region can be extended to have an endpoint close to the diagonal $x + y = 1 - (4\sqrt{2})^{-1} r_n (1 + n^{\varepsilon}/4)$. In the third region, (the yellow region in Figure 3.3), we proceed similarly to in the first region, being generous once again with the edge labels. By a symmetric argument, the path which reached the diagonal $x + y = 1 - (4\sqrt{2})^{-1} r_n (1 + n^{\varepsilon}/4)$ can be extended to a path with endpoint at the point $(1/2, 1/2)$. Since the number of vertices of $\mathcal{X}_n$ lying in a given region of $[0, 1]^2$ with area $((4\sqrt{2})^{-1} r_n)^2$ is concentrated around its mean, $N = (C^2/32) n^{1/3}$, to formalize this argument, it is convenient to frame the construction in terms of an inhomogeneous directed percolation model on the dual grid lattice where the directed edges in the percolation model are open with probabilities given by (3.3.4).

3.3.1 From temporal graph to directed percolation on the dual grid lattice

Without loss of generality, assume that $((4\sqrt{2})^{-1} r_n)^{-1}, (2(4\sqrt{2})^{-1} r_n)^{-1} \in \mathbb{N}$. Partition $[0, 1/2]^2$ into boxes of side-length $\ell_n = (4\sqrt{2})^{-1} r_n$. Let $b_n = (2\ell_n)^{-1} - 1$ and for $i, j \in [b_n]_0 = \{0, 1, \ldots, b_n\}$, let

$$\square_{i,j} := [i\ell_n, (i+1)\ell_n] \times [j\ell_n, (j+1)\ell_n], \tag{3.3.1}$$

so that by our choice for ℓ_n, for $v \in \mathcal{X}_n \cap \square_{i,j}$ and $u \in \mathcal{X}_n \cap \square_{i',j'}$ with

$$(i', j') \in \{(i+1, j), (i-1, j), (i, j+1), (i, j-1)\}$$

it holds that $\mathbf{P}\{(u, v) \in E(G_n)\} = \alpha$.

In this section, we map the construction of a temporal path between the points $(0, 0)$ and $(1/2, 1/2)$ to a directed percolation on the dual grid lattice $\mathbb{L}_n = (\mathbb{V}_n, \mathbb{E}_n)$ with vertex set

$$\mathbb{V}_n = \{\square_{i,j} \; : \; i, j \in [b_n]_0\},$$

and directed edge set $\mathbb{E}_n$ being the union of the sets

$$\{(\square_{b_n,j}, \square_{b_n,j+1}) \ : \ j \in [b_n - 1]_0\},$$

$$\{(\square_{i,b_n}, \square_{i+1,b_n}) \ : \ i \in [b_n - 1]_0\},$$

and

$$\{(\square_{i,j}, \square_{i+1,j}), (\square_{i,j}, \square_{i,j+1}) \ : \ i, j \in [b_n - 1]_0\}.$$

See Figure 3.2 for an illustration of $\mathbb{L}_n$.

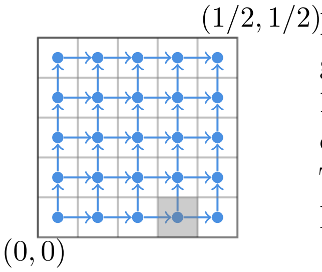

Figure 3.2: Directed grid lattice with $\ell_n = 1/10$ with vertices and directed edges in blue. The filled square highlights vertex $\square_{0,3}$.

The percolation on the finite probability space $(\Omega, \mathbb{P}, \mathcal{P}(\Omega))$, has configuration space $\Omega = \{0,1\}^{\mathbb{E}_n}$ where for $\omega \in \Omega$, and $\vec{e} \in \mathbb{E}_n$, $\omega(\vec{e}) = 1$ denotes the directed edge $\vec{e}$ being open. For any $\omega \in \Omega$, we say that the vertex $\square_{i,j} \in \mathbb{V}_n$ is connected to vertex $\square_{k,l}$, denoted $\square_{i,j} \to \square_{k,l}$, if there exists an open path in $\mathbb{E}_n$ from $\square_{i,j}$ to $\square_{k,l}$. That is, if there exists a sequence of open directed edges $\vec{e}_0 = (\square_{i,j}, \square_{i_0,j_0})$, $\vec{e}_1 = (\square_{i_0,j_0}, \square_{i_1,j_1}), \ldots, \vec{e}_m = (\square_{i_m-1,j_m-1}, \square_{k,l}) \in \mathbb{E}_n$.

Informally, the idea of the proof of Proposition 3.3.1 is to map the construction of a temporal random geometric graph $\mathcal{G}_n = (G_n, (\tau_e)_{e \in E(G_n)})$ to a directed percolation of $\mathbb{L}_n$. Fix $\varepsilon \in (0, 1/3)$. This directed percolation model is such that for $i+j \in R_1 \cup R_3$, an edge $(\square_{i,j}, \square_{i',j'}) \in \mathbb{E}_n$ is open if in $\mathcal{G}_n$, for each vertex $u \in \mathcal{X}_n \cap \square_{i,j}$ there exists a vertex $v \in \mathcal{X}_n \cap \square_{i',j'}$ such that the time-stamp τ_{uv} falls in a certain range $T_{i+j} \subset [0,1]$, with $|T_{i+j}| \approx n^{-\varepsilon}$. Crucially, if $i + j \in R_2$, then an edge $(\square_{i,j}, \square_{i',j'}) \in \mathbb{E}_n$ is open if in $\mathcal{G}_n$, for a fixed vertex $u \in \mathcal{X}_n \cap \square_{i,j}$ there exists a vertex $v \in \mathcal{X}_n \cap \square_{i',j'}$ such that $\tau_{uv} \in T_{i+j} \subset [0,1]$ with $|T_{i+j}| \approx n^{-1/3}$.

We define the ranges $(T_{i+j})_{i,j}$ as follows. For $c > 0$, $d \in \mathbb{R}$, $(a,b] \subset \mathbb{R}$, we write $d + c(a,b]$ for $(d + ca, d + cb]$. Then, taking n sufficiently large, for all $m \in [b_n]_0$, let

$$T_m := \begin{cases} \left(\frac{m}{n^\varepsilon}, \frac{m+1}{n^\varepsilon}\right] & m \in R_1, \\ \frac{1}{4} + \frac{C}{8\sqrt{2}} \left(\frac{1}{n^{1/3}}(m - \frac{n^\varepsilon}{4}), \frac{1}{n^{1/3}}(m - \frac{n^\varepsilon}{4} + 1)\right] & m \in R_2, \\ 1 - \frac{2b_n+1}{n^\varepsilon} + \left(\frac{m}{n^\varepsilon}, \frac{m+1}{n^\varepsilon}\right] & m \in R_3, \end{cases} \tag{3.3.2}$$

where $R_1 = [0, n^\varepsilon/4)$, $R_2 = [n^\varepsilon/4, 2b_n - n^\varepsilon/4)$ and $R_3 = [2b_n - n^\varepsilon/4, 2b_n]$. See Figure 3.3 for an illustration.

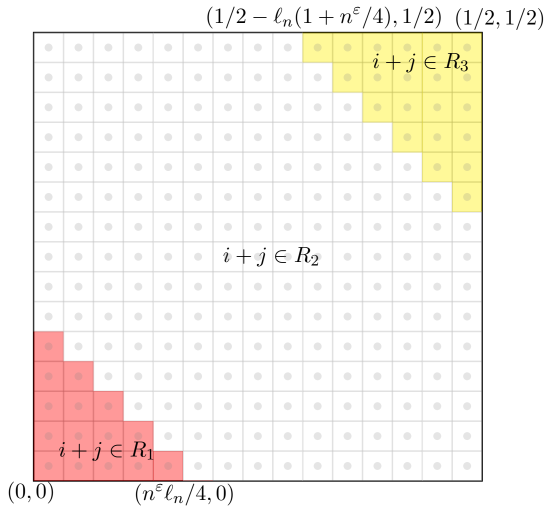

Figure 3.3: The intervals R_1, R_2, R_3 should be thought of as indicating "regions" in the grid in the sense that if, for example, $i+j\in R_1$, then $\square_{i,j}$ is a vertex in the portion of the grid which is highlighted in red above (and similarly $\square_{i,j}$ such that $i+j\in R_2$ is in white, and $\square_{i,j}$ for $i+j\in R_3$ in yellow).

Remark 3.3.2. *The sizes of the regions in (3.3.2) are chosen specifically to work well with the number of edges in $\mathcal{G}_n$ that have one endpoint in $\square_{i,j}$ and another in $\square_{i+1,j}\cup\square_{i,j+1}$. More specifically, given $\mathcal{X}_n$, the number of edges between a fixed vertex in $\mathcal{X}_n\cap\square_{i,j}$ and vertices in $\mathcal{X}_n\cap\square_{i+1,j}$ with time-stamp in $T_{i,j}$ has distribution* $\mathrm{Binomial}(|\mathcal{X}_n\cap\square_{i+1,j}|,\alpha|T_{i+j}|)$. *Since $|\mathcal{X}_n\cap\square_{i+1,j}|\overset{\mathrm{d}}{=}\mathrm{Binomial}(n,\mathrm{vol}(\square_{i+1,j}))$, the number of points in $\mathcal{X}_n\cap\square_{i+1,j}$ is highly concentrated around its mean, $n\cdot\mathrm{vol}(\square_{i+1,j})$. This implies that if $i+j\in R_2$, then in expectation a constant number of the edges between a fixed vertex in $\mathcal{X}_n\cap\square_{i,j}$ and vertices in $\mathcal{X}_n\cap(\square_{i+1,j}\cup\square_{i,j+1})$ have time-stamp in T_{i+j}. On the other hand, if $i+j\in R_1\cup R_3$, then the expected number of such edges with time-stamps in T_{i+j} is of the order of $n^{1/3-\varepsilon}$.*

We note that for all i,j, the number of vertices of $\mathcal{G}_n$ lying in a given box, $\square_{i+1,j}$, that is $|\mathcal{X}_n\cap\square_{i+1,j}|$ is concentrated around its mean $N=n\cdot\mathrm{vol}(\square_{i+1,j})=(C^2/32)n^{1/3}$. Fix $t\in(0,1)$ such that $(1+t)N,(1-t)N\in\mathbb{N}$. Let $(B_{\vec{e},k})_{\vec{e}\in\mathbb{E}_n,k\in[(1+t)N]}$ be independent random variables satisfying

$$B_{\vec{e},k}\overset{d}{=}\mathrm{Binomial}((1-t)N,\alpha|T_{i+j}|). \tag{3.3.3}$$

We consider the directed percolation model with edges $\vec{e}=(\square_{i,j},\square_{i',j'})\in\mathbb{E}_n$ open independently with probability $p_{\vec{e}}$, where

$$p_{\vec{e}}:=\begin{cases}\mathbf{P}\left\{B_{\vec{e},1}\geq 1\right\}^{(1+t)N} & \text{if } i+j\in R_1\cup R_3,\\ \mathbf{P}\left\{B_{\vec{e},1}\geq 1\right\} & \text{if } i+j\in R_2.\end{cases} \tag{3.3.4}$$

We will denote the law of the directed percolation model by $\mathbb{P}$.

Lemma 3.3.3. *Let $\varepsilon\in(0,1/3)$ with $n^\varepsilon/4\in\mathbb{N}$. The probability of having an open path from $\square_{0,0}$ to every box $\square_{i,j}$ with $i+j=n^\varepsilon/4$ in our directed percolation model* *is bounded from below by*

$$\mathbb{P}\left\{\bigcap_{i+j=\frac{n^\varepsilon}{4}}\{\square_{0,0}\to\square_{i,j}\}\right\}\geq 1-\exp\left(-\Omega(n^{1/3-\varepsilon})\right).$$

Proof. For all $\vec{e}\in\mathbb{E}_n$ such that $\vec{e}=(\square_{i,j},\square_{i',j'})$ with $i+j<n^\varepsilon/4$, we have that $|T_{i+j}|=n^{-\varepsilon}$ and so $(B_{\vec{e},k})_{k\in[(1+t)N]}$ are i.i.d. random variables with distribution Binomial$((1-t)N,\alpha n^{-\varepsilon})$. Therefore, recalling that $N=(C^2/32)n^{1/3}$, we obtain that

$$\begin{aligned}
p_{\vec{e}} &= \mathbf{P}\left\{B_{\vec{e},1}\geq 1\right\}^{(1+t)N}\\
&= \left(1-\mathbf{P}\left\{B_{\vec{e},1}<1\right\}\right)^{(1+t)N}\\
&\geq 1-(1+t)N\cdot\mathbf{P}\left\{B_{\vec{e},1}<1\right\}\\
&= 1-(1+t)N\exp\left(-\frac{\alpha}{n^\varepsilon}(1-t)N\right)\\
&= 1-\exp(-\Omega(n^{1/3-\varepsilon})). \qquad (3.3.5)
\end{aligned}$$

Each edge $\vec{e}\in\mathbb{E}_n$ is open independently with probability $p_{\vec{e}}$, and the events $\{\square_{0,0}\to\square_{i,j}\}_{i,j}$ are increasing in $\Omega=\{0,1\}^{\mathbb{E}_n}$. Therefore by the FKG inequality, see e.g. [61, Theorem 2.12],

$$\begin{aligned}
\mathbb{P}\left\{\bigcap_{i+j=\frac{n^\varepsilon}{4}}\{\square_{0,0}\to\square_{i,j}\}\right\} &\geq \prod_{i+j=\frac{n^\varepsilon}{4}}\mathbb{P}\{\square_{0,0}\to\square_{i,j}\}\\
&\geq \prod_{i+j=\frac{n^\varepsilon}{4}} p_{\vec{e}}^{\,n^\varepsilon/4},
\end{aligned}$$

where for the second inequality we have used the fact that $\mathbb{P}\{\square_{0,0}\to\square_{i,j}\}$ is at least the probability that a single path of length $n^\varepsilon/4-1$ is open. Combining this with (3.3.5) and the inequality $(1-x)^k\geq 1-xk$ for $x\in(0,1)$ yields the result. □

Lemma 3.3.4. *Let $\varepsilon\in(0,1/3)$ with $n^\varepsilon/4\in\mathbb{N}$, then the probability of having an open path from a box $\square_{i,j}$ with $i+j=n^\varepsilon/4$ to a box $\square_{k,l}$ with $k+l=2b_n-n^\varepsilon/4$ in our directed percolation model can be bounded from below by*

$$\mathbb{P}\left\{\bigcup_{i+j=\frac{n^\varepsilon}{4}}\ \bigcup_{k+l=2b_n-\frac{n^\varepsilon}{4}}\{\square_{i,j}\to\square_{k,l}\}\right\}\geq 1-\exp\left(-\Omega(n^\varepsilon\wedge n^{1/3-\varepsilon})\right).$$

Proof. We condition on the event

$$\mathcal{E}_n=\bigcap_{i+j=\frac{n^\varepsilon}{4}}\{\square_{0,0}\to\square_{i,j}\},$$

which by Lemma 3.3.3 occurs with probability $1-\exp(-\Omega(n^{1/3-\varepsilon}))$.

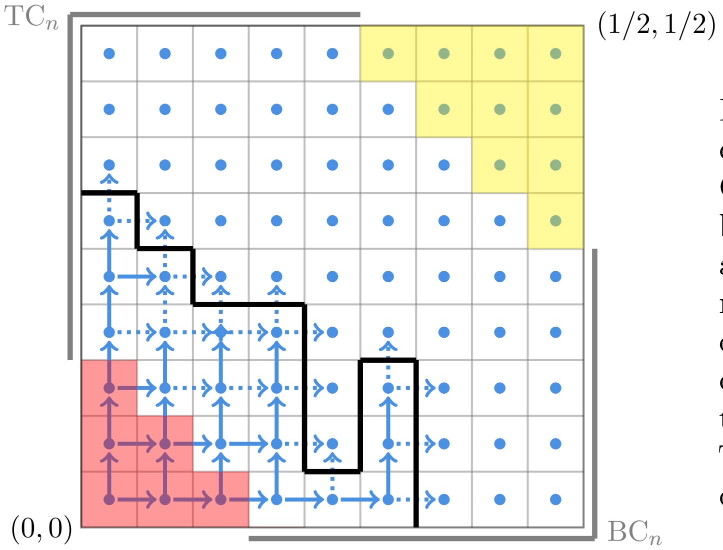

Figure 3.4: The connected component containing $\square_{0,0}$. Open edges are drawn as solid blue arrows while closed edges are drawn as dotted blue arrows. In black, the boundary of the connected component containing $\square_{0,0}$, as a path in the lattice with mesh size ℓ_n. The bold-face grey lines border the regions TC_n and BC_n.

On this event, if there is no open path between $\{\square_{i,j} : i+j = n^\varepsilon/4\}$ and $\{\square_{k,l} : k+l = 2b_n - n^\varepsilon/4\}$, then the connected component containing $\square_{0,0}$ and the vertices $\{\square_{i,j} : i+j = n^\varepsilon/4\}$ has a boundary which can be represented as a path in the lattice (rather than in the dual grid lattice) with mesh size ℓ_n, with one endpoint in the top left-hand corner region of the unit square,

$$\mathrm{TC}_n := \{(0, j\ell_n) : j \in \{n^\varepsilon/4, \dots, b_n+1\}\} \cup \{(i\ell_n, 1) : i \in \{0, \dots, b_n - n^\varepsilon/4\}\}$$

and another in bottom right-hand corner of the unit square,

$$\mathrm{BC}_n := \{(1, j\ell_n) : j \in \{0, \dots, b_n - n^\varepsilon/4\}\} \cup \{(i\ell_n, 0) : i \in \{n^\varepsilon/4, \dots, b_n+1\}\}.$$

See Figure 3.4 for an illustration.

Let $P_n^{(i,j)}$ be the number paths which form a boundary in the percolation model with one endpoint being the point $(i\ell_n, j\ell_n) \in \mathrm{TC}_n$, and the other endpoint being in BC_n. Then,

$$\mathbb{P}\left\{ \bigcup_{i+j=\frac{n^\varepsilon}{4}} \bigcup_{k+l=2b_n-\frac{n^\varepsilon}{4}} \{\square_{i,j} \to \square_{k,l}\} \,\middle|\, \mathcal{E}_n \right\} \geq 1 - \sum_{(i\ell_n, j\ell_n)\in\mathrm{TC}_n} \mathbb{P}\{P_n^{(i,j)} \geq 1\}. \quad (3.3.6)$$

To prove Lemma 3.3.4 it remains to show that the right-hand side of the above inequality can be bounded from below by $1-\exp(-\Omega(n^\varepsilon))$. To this end, let $(i\ell_n, j\ell_n) \in \mathrm{TC}_n$. We use the first moment method to bound $\mathbb{P}\{P_n^{(i,j)} \geq 1\}$. Each path that could form a boundary for the directed percolation has length $m \in \{n^\varepsilon/2, \dots, 2b_n(b_n+1)\}$, where we note that the upper bound is the total number of edges in the lattice and $2b_n(b_n+1) = \frac{16}{C^2}n^{2/3} - \frac{4\sqrt{2}}{C}n^{1/3}$. Lastly, in order for such a path to represent a boundary of the connected component, at least one half of its edges must intersect closed edges of $\mathbb{L}_n$ (viewing the path as from TC_n to BC_n, the edges that step down or right must intersect closed edges of $\mathbb{L}_n$), see again Figure 3.4. By (3.3.4) and (3.3.2), the directed edges intersecting this path are each closed independently with probability $1-\mathbf{P}\left\{B^{(n)} \geq 1\right\}$, where $B^{(n)}$ is a Binomial$((1-t)N, \alpha|T_{n^\varepsilon/4}|)$ distributed random variable where

$$|T_{n^\varepsilon/4}| = \frac{C}{8\sqrt{2}n^{1/3}}.$$

Since $N = \frac{C^2}{32}n^{1/3}$, it follows that

$$\mathbb{E}\left\{B^{(n)}\right\} = \frac{\alpha(1-t)N}{8\sqrt{2}n^{1/3}} = \frac{\alpha(1-t)C^3 n^{1/3}}{32\cdot 8\sqrt{2}n^{1/3}} = \frac{\alpha(1-t)C^3}{256\sqrt{2}}.$$

It follows that for all $\delta > 0$, for n sufficiently large,

$$\mathbf{P}\left\{B^{(n)} \geq 1\right\} \geq \mathbf{P}\left\{\text{Poisson}\left(\frac{\alpha(1-t)C^3}{256\sqrt{2}}\right) \geq 1\right\} - \delta,$$

and so

$$1 - \mathbf{P}\left\{B^{(n)} \geq 1\right\} \leq \exp\left(-\frac{\alpha(1-t)C^3}{256\sqrt{2}}\right) + \delta. \tag{3.3.7}$$

Let $\delta > 0$. Notice that to construct a boundary path, edge by edge starting from a point in TC_n, at each step there are at most 3 possibilities for the direction of the next added edge. Therefore, we obtain that for sufficiently large n,

$$\begin{aligned}
\mathbb{E}\left\{P_n^{(i,j)}\right\} &\leq \sum_{m=n^\varepsilon/2}^{2b_n(b_n+1)} 3^m (1-\mathbf{P}\{B \geq 1\})^{\lfloor m/2 \rfloor} &(3.3.8)\\
&\leq \sum_{m=n^\varepsilon/2}^{4b_n^2} 3^m \left(\exp\left(-\frac{\alpha(1-t)C^3}{256\sqrt{2}}\right) + \delta\right)^{\lfloor m/2 \rfloor} \\
&\leq \ell_n^{-2}\left(10\exp\left(-\frac{\alpha(1-t)C^3}{256\sqrt{2}}\right)\right)^{n^\varepsilon/4}, &(3.3.9)
\end{aligned}$$

where the final inequality follows from taking $\delta > 0$ sufficiently small. Taking $C > 0$ sufficiently large, so that

$$10\exp\left(-\frac{\alpha(1-t)C^3}{256\sqrt{2}}\right) < e^{-4}, \tag{3.3.10}$$

we obtain that $\mathbb{E}\{P_n^{(i,j)}\} = \exp(-\Omega(n^\varepsilon))$. The result then follows from (3.3.6) by an application of Markov's inequality and the fact that $\#\{(i\ell_n, j\ell_n) \in \mathrm{TC}_n\} = O(\ell_n^{-1}) = O(n^{1/3})$. $\square$

3.3.2 Proof of Proposition 3.3.1.

In this subsection we will show how Lemmas 3.3.3 and 3.3.4 imply that for $\varepsilon \in (0, 1/9)$,

$$\mathbf{P}\left\{(0,0) \xrightarrow{\mathcal{G}_n} (\frac{1}{2}, \frac{1}{2})\right\} \geq 1 - \exp(-\Omega(n^{\varepsilon})), \tag{3.3.11}$$

for $(\mathcal{G}_n)_{n\geq 1}$ as in Proposition 3.3.1 and where we work with a modified point process obtained by adding the points $(0,0)$ and $(1/2, 1/2)$ to $\mathcal{X}_n$. To see why Proposition 3.3.1 follows from (3.3.11), note that the maximal distance between two vertices in $\mathcal{G}_n$ is $\sqrt{2}/2$, for instance between $(0,0)$ and $(1/2, 1/2)$. Since we are working on the torus $[0,1]^2$, the same proof construction can be used for any pair of points $(x, y), (x', y') \in \mathcal{X}_n$ by translating and/or rotating the first and third regions (highlighted in red and yellow in Figure 3.5) such that they start at (x, y) and (x', y') respectively, and appropriately shrinking the second region (the white region in Figure 3.5), while keeping the box side-length $\ell_n = (4\sqrt{2})^{-1} r_n$ as before. If the first and third regions overlap (and so there is no white region in Figure 3.5), we shrink the first and third regions appropriately while keeping the same box side-length as in the original construction. See Figure 3.5.

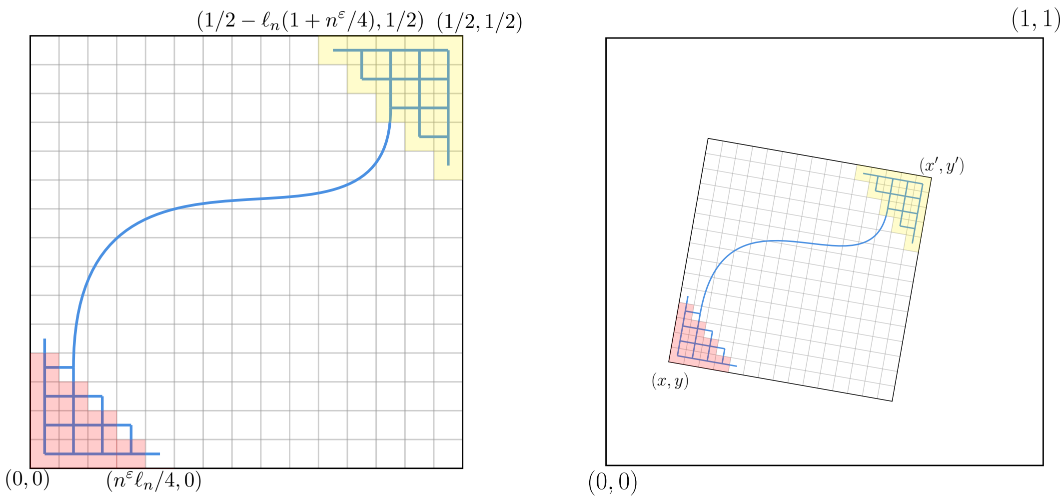

Figure 3.5: On the left, an illustration of the construction of $(0,0) \xrightarrow{\mathcal{G}_n} (\frac{1}{2}, \frac{1}{2})$. On the right, an illustration of the alteration of this construction for a temporal path between two arbitrary points (x, y) and (x', y').

A union bound over all possible pairs of vertices then gives

$$\mathbf{P}\left\{\forall u, v \in \mathcal{G}_n, u \xrightarrow{\mathcal{G}_n} v\right\} \geq 1 - n^2 \exp(-\Omega(n^{\varepsilon})), \tag{3.3.12}$$

proving Proposition 3.3.1 for $d = 2$.

Proof of Proposition 3.3.1 with $d = 2$. As discussed above, to prove Proposition 3.3.1, it suffices to prove (3.3.11). Recall $(\square_{i,j})_{i,j}$ from (3.3.1). We begin by defining A_n as the event that the number of vertices in each box $\square_{i,j}$, is well-concentrated. More precisely, for $t \in (0,1)$ let

$$A_n = A_n(t) := \left\{ \max_{i,j} ||\mathcal{X}_n \cap \square_{i,j}| - N| < tN \right\}, \tag{3.3.13}$$

recalling that $|\mathcal{X}_n \cap \square_{i,j}|$ is a Binomial$(n, \text{vol}(\square_{0,0}))$ random variable and $N = n\text{vol}(\square_{0,0}) = (C^2/32)n^{1/3}$. By [74, Corollary 2.3] it holds that for all $t \in (0,1)$, for $n \geq 1$,

$$\mathbf{P}\{A_n(t)^c\} \leq \frac{16n^{2/3}}{C^2} \exp\left(-\frac{(Ct)^2}{96} n^{1/3}\right) = \exp(-\Omega(n^{1/3})). \tag{3.3.14}$$

We therefore focus on proving that

$$\mathbf{P}\left\{(0,0) \xrightarrow{\mathcal{G}_n} (\frac{1}{2}, \frac{1}{2}) \mid A_n\right\}$$

can be bounded from below by the right-hand side of (3.3.11) for the remainder of this section. We lower bound this probability by that of having a temporal path from $(0,0)$ to $(\frac{1}{2}, \frac{1}{2})$ using only edges xy satisfying $x \in \mathcal{X}_n \cap \square_{i,j}$, $y \in \mathcal{X}_n \cap (\square_{i+1,j} \cup \square_{i,j+1})$ and $\tau_{xy} \in T_{i,j}$. Observe that by our choice of box side-length $\ell_n = (4\sqrt{2})^{-1} r_n$ we have that an edge xy with $x \in \mathcal{X}_n \cap \square_{i,j}$, $y \in \mathcal{X}_n \cap (\square_{i+1,j} \cup \square_{i,j+1})$ exists with probability α. Further remark that for all $i, j \in [b_n]_0$ and $\vec{e} = (\square_{i,j}, \square_{i',j'}) \in \mathbb{E}_n$,

$$\begin{aligned} &\mathbf{P}\{\forall x \in \mathcal{X}_n \cap \square_{i,j},\, \exists y \in \mathcal{X}_n \cap \square_{i',j'} \text{ s.t. } xy \in E(G_n), \tau_{xy} \in T_{i+j} | A_n\} \\ &\geq \mathbf{P}\left\{ \bigcap_{k \in [(1+t)N]} \{B_{\vec{e},k} \geq 1\} \right\}, \end{aligned} \tag{3.3.15}$$

and given $x \in \mathcal{X}_n \cap \square_{i,j}$,

$$\mathbf{P}\{\exists y \in \mathcal{X}_n \cap \square_{i',j'} \text{ s.t. } xy \in E(G_n), \tau_{xy} \in T_{i+j} | A_n\} \geq \mathbf{P}\{B_{\vec{e},1} \geq 1\}, \tag{3.3.16}$$

where we recall from (3.3.3) that $(B_{\vec{e},k})_{k\geq 1}$ are i.i.d. Binomial$((1-t)N, \alpha|T_{i,j}|)$ distributed random variables. Finally, we will ensure that if we can reach some vertex in the final box $\square_{b_n,b_n}$ with a temporal path using only time-stamps in $[0, 1-n^{-\varepsilon}]$, then we can complete the temporal path from $(0,0)$ to $(\frac{1}{2}, \frac{1}{2})$ using two further edges. To be precise, on the event A_n, given $x \in \mathcal{X}_n \cap \square_{b_n,b_n}$, the probability of having a temporal path from x to $(\frac{1}{2}, \frac{1}{2})$ only using edges with time-stamps in $(1-n^{-\varepsilon}, 1]$ is bounded from below by the probability of having $n^{1/3-2\varepsilon}$ vertices $y \in \mathcal{X}_n \cap \square_{b_n,b_n}$ such that $xy \in E(G_n)$ and $\tau_{xy} \in (1-n^{-\varepsilon}, 1-n^{-\varepsilon}/2]$ and such that one of these y

satisfies $y(\frac{1}{2},\frac{1}{2}) \in E(G_n)$ and $\tau_{y(\frac{1}{2},\frac{1}{2})} \in (1-n^{-\varepsilon}/2, 1]$. This probability can further be bounded from below by

$$\mathbf{P}\left\{\text{Binomial}((1-t)N, \alpha n^{-\varepsilon}/2) \geq n^{1/3-2\varepsilon}\right\} \times \mathbf{P}\left\{\text{Binomial}(n^{1/3-2\varepsilon}, \alpha n^{-\varepsilon}/2) \geq 1\right\}. \tag{3.3.17}$$

The mean of the Binomial in the first probability is of the order of $n^{1/3-\varepsilon}$, thus by a Chernoff bound, see [61, Theorem 2.1], this probability is at least $1-\exp(-\Omega(n^{1/3-\varepsilon}))$. The second term in (3.3.17) equals $1-(1-\alpha n^{-\varepsilon}/2)^{n^{1/3-2\varepsilon}}=1-\exp(-\Omega(n^{1/3-3\varepsilon}))$. Putting everything together, on the event A_n, given $x \in \mathcal{X}_n \cap \square_{b_n,b_n}$, the probability of having a temporal path from x to $(\frac{1}{2},\frac{1}{2})$ only using edges with time-stamps in $(1-n^{-\varepsilon}, 1]$ is bounded from below by $1-\exp(-\Omega(n^{1/3-3\varepsilon}))$.
From this, we can see that

$$\mathbf{P}\left\{(0,0) \xrightarrow{\mathcal{G}_n} (\frac{1}{2},\frac{1}{2}) \mid A_n\right\} \geq \left(1-\exp(-\Omega(n^{1/3-3\varepsilon}))\right) \mathbb{P}\left\{\square_{0,0} \to \square_{b_n,b_n}\right\}. \tag{3.3.18}$$

Since the edges $\vec{e} \in \mathbb{E}_n$ are open independently, $\mathbb{P}\{\square_{0,0} \to \square_{b_n,b_n}\}$ can be bounded from below by

$$\mathbb{P}\left\{\bigcap_{i+j=\frac{n^\varepsilon}{4}-1} \{\square_{0,0} \to \square_{i,j}\}\right\} \tag{3.3.19}$$

$$\times\, \mathbb{P}\left\{\bigcup_{i+j=\frac{n^\varepsilon}{4}-1} \bigcup_{k+l=2b_n-n^\varepsilon/4} \{\square_{i,j} \to \square_{k,l}\}\right\} \tag{3.3.20}$$

$$\times\, \mathbb{P}\left\{\bigcap_{i+j=2b_n-n^\varepsilon/4} \{\square_{i,j} \to \square_{b_n,b_n}\}\right\}. \tag{3.3.21}$$

Lemmas 3.3.3 and 3.3.4 yield a lower bound for (3.3.19) and (3.3.20). Further, since the time-stamp ranges in the lower left and upper right triangle regions, $\{\square_{i,j} : i+j \leq n^\varepsilon/4-1\}$ and $\{\square_{i,j} : i+j \geq 2b_n - n^\varepsilon/4\}$ respectively, all satisfy $|T_{i+j}| = n^{-\varepsilon}$, we can argue by symmetry that the lower bound in Lemma 3.3.3 also holds for (3.3.21), giving

$$\mathbb{P}\{\square_{0,0} \to \square_{b_n,b_n}\} \geq 1-\exp(-\Omega(n^{1/3-\varepsilon}))-\exp(-\Omega(n^\varepsilon)).$$

Together with (3.3.18) and (3.3.14) we conclude (3.3.11), proving Proposition 3.3.1. □

3.3.2.1 Extension to dimension $d > 2$

The proof of Proposition 3.3.1 in dimension $d > 2$ can be deduced from the dimension $d = 2$ case. Similarly to the case $d = 2$, it suffices to construct a likely path from the furthest possible separated points on the torus $[0,1]^d$, namely $\vec{0}, \vec{\frac{1}{2}}$ and then conclude by a union bound. For such a path, we only consider points that are inside a slab of width $r_n/2$. More precisely, in this slab the points $\vec{0}, \vec{\frac{1}{2}}$ are the opposite corners of a rotated copy $[0, \frac{\sqrt{d}}{2\sqrt{2}}]^2 \times \{0\}^{d-2}$. This slab has volume of order r_n^{d-2} and therefore, with high probability the number of points in this region is $\omega(nr_n^{d-2})$.

The result then follows from Lemmas 3.3.3 and 3.3.4 by projecting these $\omega(nr_n^{d-2})$ points onto a copy of $[0, \frac{\sqrt{d}}{2\sqrt{2}}]^2$ and noting that for any two vertices in the slab u, v, if their projections $\hat{u}$ and $\hat{v}$ are such that $\|\hat{u} - \hat{v}\| \leq r_n/2$, then $\|u - v\| \leq r_n$ and so

$$\mathbf{P}\left\{uv \in E(G_n) \mid \mathcal{X}_n\right\} = K\left(\frac{\|u-v\|}{r_n}\right) \geq K\left(\frac{\|\hat{u}-\hat{v}\|}{r_n/2}\right).$$

Chapter 4

Random walks on Coxeter interchange graphs

This chapter is based on joint work with Matthew Buckland (University of Lübeck), Brett Kolesnik (University of Warwick), and Tomasz Przybyłowski (University of Oxford). This chapter has been published in the *Electronic Journal of Probability* [27].

4.1 Introduction

A tournament is an orientation of a graph. We think of vertices as players and edges as games, the orientation of which indicates the winner. Tournaments are related to the geometry of the permutahedron Π_{n-1}, which is a classical polytope in discrete convex geometry. See, e.g., Stanley [102], Ziegler [109], and Kolesnik and Sanchez [67].

Classical combinatorics is related to the root system of type A_n. Coxeter combinatorics is concerned with extensions to the other roots systems of types B_n, C_n and D_n (and sometimes also the finite, exceptional types E_6, E_7, E_8, F_4 and G_2). For example, works by Galashin, Hopkins, McConville and Postnikov [43, 44] have investigated Coxeter versions of the chip-firing game (the sandpile model).

Recently, Kolesnik and Sanchez [69] introduced the Coxeter analogue of graph tournaments, which are associated with orientations of signed graphs, as in Zaslavsky [108], and the Coxeter permutahedra Π_Φ, recently introduced by Ardila, Castillo, Eur and Postnikov [13]. Coxeter tournaments involve collaborative and solitaire games, as well as the usual competitive games in graph tournaments.

In this work, we show (see Theorem 4.3.1 below) that random walks rapidly mix on the sets of Coxeter tournaments with given score sequence, that is, on the fibers of the Coxeter permutahedra Π_Φ. Informally, this means that the walk is close to uniform in the fiber after a short amount of time, yielding an efficient way to sample from this set of interest.

Many combinatorial properties of these structures remain mysterious. The purpose of this work is to explore the associated Coxeter interchange graphs, which encode their combinatorics, via random walks. These graphs, introduced by Kolesnik, Mitchell and Przybyłowski [68], generalize the interchange graphs introduced by Brualdi and Li [24]. Rapid mixing in the classical setting was established by Kannan, Tetali and Vempala [64] and McShine [89]. We recover these results by our general strategy.

Let us emphasize that even the classical interchange graphs appear to be difficult to describe in general. Indeed, Brualdi and Li [24, p. 151] state that they have "a rich and fascinating combinatorial structure and that much remains to be determined." Even counting the number of vertices is of "considerable interest and considerable difficulty" [24, p. 143].

Beginning with Spencer [101], and subsequent works by McKay [87], McKay and Wang [88], and Isaev, Iyer and McKay [56], asymptotic estimates for the number of vertices in the interchange graphs have been found only for fibers of points near the center of Π_{n-1}. In the other extreme, Chen, Chang and Wang [33] showed that, for certain points near the boundary of Π_{n-1}, the interchange graph is the classical hypercube.

The Coxeter interchange graphs are richer still. Therefore, in broad terms, we show in this work that random walks rapidly mix on a wide and intricate class of graphs.

Rapid mixing can sometimes be used to approximately count sets of interest. Roughly speaking, this is because the uniform measure π on a set S is related to the size of the set $\pi = 1/|S|$. See, e.g., Sinclair [100] for more details. The current work might serve as a first step towards developing efficient approximate counting schemes for the fibers of the Coxeter permutahedra Π_Φ.

As this work touches on a variety of subjects (combinatorics, geometry, algebra and probability), some preliminaries are required before we can state our results precisely. In Section 4.1.1, we discuss the literature related to tournaments and the standard permutahedron Π_{n-1} (of type A_{n-1}). Our results are discussed informally in Sections 4.1.2 and 4.1.3. Further background on tournaments, root systems, signed graphs and combinatorial geometry is in Section 4.2 and in the previous works in this series [67–69]. Our main result is stated formally in Section 4.3. See Sections 4.4, 4.5 and 4.6 for the proofs.

We hope that this work will serve as an invitation to step into the Coxeter "worlds" (of types B_n, C_n and D_n). We believe that Coxeter combinatorics is fertile ground, where algebraists, combinatorialists, geometers and probabilists can open new lines of fruitful communication. In particular, many problems in discrete probability likely have a Coxeter analogue, waiting to be discovered.

4.1.1 Context

A *tournament* is an orientation of the complete graph K_n, encoded as some $T = (w_{ij} : i > j)$ with all $w_{ij} \in \{0, 1\}$. Each edge $\{i, j\}$ in K_n is oriented as $i \to j$ if $w_{ij} = 1$ or $i \leftarrow j$ if $w_{ij} = 0$. We think of each edge as a game, directed away from the winner. The *win sequence*

$$\mathbf{w}(T) = \sum_{i>j} [w_{ij}\mathbf{e}_i + (1 - w_{ij})\mathbf{e}_j]$$

lists the total number of wins by each player, where $\mathbf{e}_i \in \mathbb{Z}^n$ are the standard basis vectors. We let $\mathbf{w}_n = (0, 1, \ldots, n-1)$ denote the *standard win sequence,* corresponding to the *transitive* (acyclic) tournament in which $w_{ij} = 1$ for all $i > j$. In a sense, $\mathbf{w}_n$ and its permutations are as "spread out" as possible.

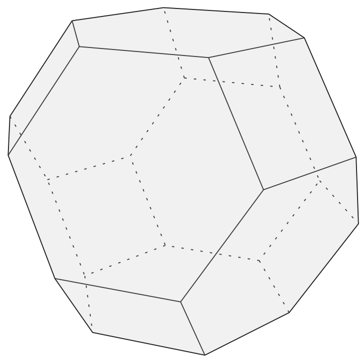

Figure 4.1: The permutahedron $\Pi_3 \subset \mathbb{R}^4$, projected into $\mathbb{R}^3$. Its 24 vertices correspond to the permutations of the standard win sequence $\mathbf{w}_4 = (0, 1, 2, 3)$.

Results by Rado [98] and Landau [73] imply that the set $\mathrm{Win}(n)$ of all win sequences is precisely the set of lattice points in the permutahedron Π_{n-1}, that is, $\mathrm{Win}(n) = \mathbb{Z}^n \cap \Pi_{n-1}$. We recall that Π_{n-1} is a classical polytope in discrete geometry (see, e.g., Ziegler [109]), obtained as the convex hull of $\mathbf{w}_n$ and its permutations, see Figure 4.1. By Stanley [102], win sequences are in bijection with spanning forests $F \subset K_n$. (The volume of Π_{n-1} is the number of spanning trees $T \subset K_n$.) See Postnikov [97] for generalizations.

It is convenient to make a linear shift

$$\Pi'_{n-1} = \Pi_{n-1} - \frac{n-1}{2}\mathbf{1}_n, \tag{4.1.1}$$

where $\mathbf{1}_n = (1, \ldots, 1) \in \mathbb{Z}^n$. Note that this re-centers the polytope at the origin $\mathbf{0}_n = (0, \ldots, 0) \in \mathbb{Z}^n$. The *score sequence*

$$\mathbf{s}(T) = \mathbf{w}(T) - \frac{n-1}{2}\mathbf{1}_n,$$

associated with the win sequence $\mathbf{w}(T)$ of a tournament T, is given by

$$\mathbf{s}(T) = \sum_{i>j}(w_{ij} - 1/2)(\mathbf{e}_i - \mathbf{e}_j). \tag{4.1.2}$$

This shift corresponds to awarding a $\pm 1/2$ point for each win/loss. We let $\text{Score}(n)$ denote the set of all possible score sequences.

Although the set $\text{Score}(n)$ has a simple, geometric description, the set $\text{Tour}(n, \mathbf{s})$, of tournaments with given score sequence $\mathbf{s}$, appears to be quite combinatorially complex.

Kannan, Tetali and Vempala [64] investigated simple random walk as a way of sampling from $\text{Tour}(n, \mathbf{s})$. However, rapid mixing was proved only for $\mathbf{s}$ sufficiently close to $\mathbf{0}_n$. McShine [89] established rapid mixing in time $O(n^3 \log n)$, for *all* $\mathbf{s} \in \text{Score}(n)$, by an elegant application of Bubley and Dyer's [25] method of path coupling, which was relatively new at the time. See Section 4.2.3 below for an overview. We note that Sarkar [99] has shown that mixing takes $\Omega(n^3)$ for some sequences.

More specifically, in [64, 89], the random walks are on the *interchange graphs* $\text{IntGr}(n, \mathbf{s})$ introduced by Brualdi and Li [24]. Any two tournaments with the same score sequence have the same number of copies of the *cyclic triangle* Δ_c (see Figure 4.5). Moreover, if some T contains a copy Δ of Δ_c, then the tournament $T * \Delta$, obtained by reversing the orientation of all edges in Δ, has the same score sequence as T. These observations are the key to exploring $\text{Tour}(n, \mathbf{s})$. The graph $\text{IntGr}(n, \mathbf{s})$ has a vertex $v(T)$ for each $T \in \text{Tour}(n, \mathbf{s})$ and two $v(T_1), v(T_2)$ are neighbours if $T_2 = T_1 * \Delta$ for some copy $\Delta \subset T_1$ of Δ_c. It can be shown that $\text{IntGr}(n, \mathbf{s})$ is connected. In this sense, Δ_c *generates* the set $\text{Tour}(n, \mathbf{s})$.

The permutahedron Π_{n-1} is related with the standard root system of type A_{n-1}, as the symmetric group S_n is the Weyl group of type A_{n-1}. We refer to, e.g., the standard text by Humphreys [55] for background on root systems. We recall that Killing [66] and Cartan [29] classified all (irreducible, crystallographic) root systems (up to isomorphism) as the infinite families A_{n-1}, B_n, C_n and D_n and the finite exceptional types E_6, E_7, E_8, F_4 and G_2.

Coxeter permutahedra Π_Φ, recently studied by Ardila, Castillo, Eur and Postnikov [13], are obtained by replacing the role of S_n in the definition of Π_{n-1} with the Weyl group W_Φ of a root system Φ. See, e.g., Figure 4.2.

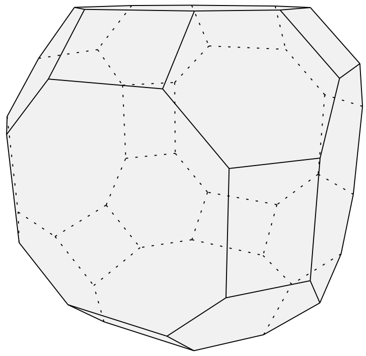

Figure 4.2: The Coxeter permutahedron of type C_3.

The previous works in this series [68, 69] studied the connection between the polytopes Π_Φ and *Coxeter tournaments,* which are related to orientations of signed graphs, as developed by Zaslavsky [106–108]. As mentioned above, these tournaments involve collaborative and solitaire games, as well as the usual competitive games in classical graph tournaments.

4.1.2 Purpose

In this work (see Theorem 4.3.1) we show that simple random walks mix rapidly on the *Coxeter interchange graphs* $\mathrm{IntGr}(\Phi, \mathbf{s})$. These graphs encode the combinatorics of the sets $\mathrm{Tour}(\Phi, \mathbf{s})$, of Coxeter tournaments with a given score sequence $\mathbf{s}$, and give structural information about the fibers of the Coxeter permutahedra Π_Φ. We focus on the non-standard types $\Phi = B_n$, C_n and D_n.

We also show (see Theorem 4.4.1) that all Coxeter interchange graphs are connected and we bound their diameter. In constructing our random walk couplings, we uncover various other fine, structural properties of the graphs $\mathrm{IntGr}(\Phi, \mathbf{s})$, and hence the sets $\mathrm{Tour}(\Phi, \mathbf{s})$, which might be of independent (algebraic, geometric, etc.) interest.

4.1.3 Discusssion

Path coupling is a powerful method for establishing rapid mixing (see Section 4.2.3). As already mentioned, path coupling was used in [89]. We will also use this method, however, the application in the Coxeter setting is significantly more delicate.

As discussed above, the interchange graphs in type A_{n-1} are generated by a single neutral tournament, namely, the cyclic triangle Δ_c. On the other hand, in the Coxeter setting, there are a number of other generators which play a role (see Figures 4.6, 4.7 and 4.5). A fascinating interplay arises, as these generators can interact in a variety of interesting ways. As such, the Coxeter interchange graphs are much richer in complexity. Likewise, the analysis of random walks on these structures is more involved.

The types increase in difficulty in order A_{n-1}, D_n, B_n, C_n. Type C_n is especially challenging, due to the presence of loops in some of the generators, which we call *clovers* (see Figure 4.7). These generators correspond to double edges in an interchange graph. In particular, a special type of structure, which we call a *crystal* (see Figure 4.25), can appear in type C_n interchange graphs. The crystal arises in C_3 when $\mathbf{s} = (2, 1, 1)$. Crystals can also be found as subgraphs in larger type C_n interchange graphs, for instance, in the *snare drum* in Figure 4.3, when $\mathbf{s} = (-1, 0, 1)$. In other examples, such as the *tambourine* in Figure 4.4, when $\mathbf{s} = (0, 0, 0)$ is the center of the polytope, there are no crystals, but interesting structure nonetheless.

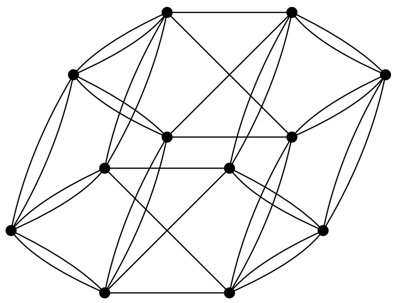

Figure 4.3: The snare drum interchange graph $\mathrm{IntGr}(C_3, \mathbf{s})$, when $\mathbf{s} = (-1, 0, 1)$, is the Cartesian product of a double edge and the crystal (see Figure 4.25).

The presence of crystals in type C_n interchange graphs leads to two main issues. The first is in extending certain natural couplings on various small subgraphs (the *extended networks* discussed in Sections 4.5.1 and 4.5.2) to a unified coupling on the entire interchange graph. To overcome this difficulty, we will prove a number of detailed combinatorial properties of the Coxeter interchange graphs. We classify the types of subgraphs (see Figures 4.14, 4.17 and 4.25), which together form the full graph, and study the ways in which they can intersect. For example, one crucial property (see Lemma 4.5.9) is that any two crystals can share at most one single edge. Without this property, it seems that a path coupling argument would not be possible.

The second issue caused by crystals is in obtaining a "contractive" (see Section 4.2.3) coupling. In applying path coupling, we will need to re-weight the graph metric in a specific way, which accounts for the occurrence of crystals. Loosely speaking, the choice of weights is related to the fact that, in our random walk couplings, crystals work like "switches," that convert single edges to double edges, and vice versa.

The overall coupling used to establish rapid mixing in the Coxeter setting is quite elaborate. See, e.g., Figures 4.28, 4.29 and 4.30 below. The classical type A_{n-1} result [64, 89] is a special case of the argument depicted in Figure 4.27.

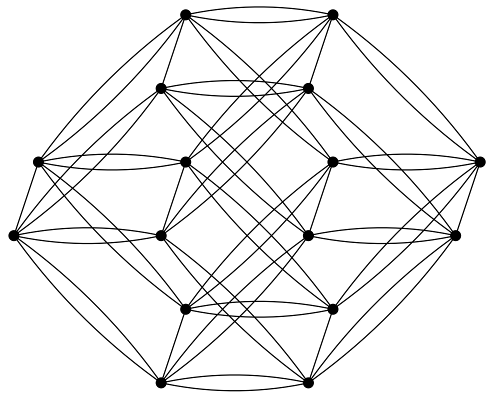

Figure 4.4: The tambourine interchange graph $\mathrm{IntGr}(C_3, \mathbf{s})$, when $\mathbf{s} = (0, 0, 0)$ is the center of the type C_3 permutahedron. This graph is the Cartesian product of a single edge and the cube of double edges.

4.2 Background

We refer to [13], Humphreys [55], Zaslavsky [106–108], and the previous works in this series [68, 69] for a detailed background on root systems, signed graphs and their connections to discrete geometry. In this section, we will only recall what is used in the current work.

4.2.1 Coxeter tournaments

A *signed graph* $\mathcal{S}$ on $[n] = \{1, 2, \ldots, n\}$ has a set of signed edges $E(\mathcal{S})$. The four possible types of edges are:

- *negative edges* e_{ij}^- between two vertices i and j,
- *positive edges* e_{ij}^+ between two vertices i and j,
- *half edges* e_i^h with only one vertex i, and
- *loops* e_i^ℓ at a vertex i.

We note that classical graphs G correspond to signed graphs $\mathcal{S}$ with only negative edges.

In this work, we focus on the *complete signed graphs* $\mathcal{K}_\Phi$ of types $\Phi = B_n$, C_n and D_n. These signed graphs contain all possible negative and positive edges $e_{ij}^\pm$. In type B_n (resp. C_n), all possible half edges e_i^h (resp. loops e_i^ℓ) are also included. We call a signed graph $\mathcal{S}$ a Φ-*graph* if $\mathcal{S} \subset \mathcal{K}_\Phi$.

Most of the results in the literature on classical (type A_{n-1}) graph tournaments restricts to the case that G is the complete graph K_n. We note that the signed graph $\mathcal{K}_{A_{n-1}}$ with all possible negative edges (and no other types of signed edges) corresponds to the classical complete graph K_n.

A *Coxeter tournament* $\mathcal{T}$ on a signed graph $\mathcal{S}$ is an orientation of $\mathcal{S}$. When $\mathcal{S}$ is unspecified, our default assumption will be that $\mathcal{S} = \mathcal{K}_\Phi$. More formally, $\mathcal{T} = (w_e : e \in E(\mathcal{S}))$, with all $w_e \in \{0, 1\}$. We think of each $e \in E(\mathcal{S})$ as a *game*, and w_e as indicating its outcome. (We think of $E(\mathcal{S})$ as having a natural ordering, so that $(w_e : e \in E(\mathcal{S}))$ holds all information about the orientation of $\mathcal{S}$ under $\mathcal{T}$. However, we could, somewhat pedantically, instead write $\mathcal{T} = \{(e, w_e) : e \in E(\mathcal{S})\}$.)

The *score sequence* is given by, cf. (4.1.2),

$$\mathbf{s}(\mathcal{T}) = \sum_{e \in E(\mathcal{S})} (w_e - 1/2)\mathbf{e},$$

where $\mathbf{e}$ is the vector corresponding to the signed edge e, given by $\mathbf{e}_{ij}^\pm = \mathbf{e}_i \pm \mathbf{e}_j$, $\mathbf{e}_i^h = \mathbf{e}_i$ and $\mathbf{e}_i^\ell = 2\mathbf{e}_i$. In other words:

- negative edges e_{ij}^- are *competitive games* in which one of i, j wins and the other loses a 1/2 point,
- positive edges e_{ij}^+ are *collaborative games* in which i, j both win or lose a 1/2 point,
- half edges e_i^h are *(half edge) solitaire games* in which i wins or loses a 1/2 point, and
- loops e_i^ℓ are *(loop) solitaire games* in which i wins or loses 1 point.

If $\mathbf{s}(\mathcal{T}) = \mathbf{0}_n$ we say that $\mathcal{T}$ is *neutral*.

We let $\mathbf{s}_\Phi$ denote the *standard score sequence* corresponding to the Coxeter tournament in which all $w_e = 1$. We note that, in some contexts, $\mathbf{s}_\Phi$ is called the *Weyl vector*. It is also the sum of the *fundamental weights* of the root system Φ. See, e.g., [52, 55] for more details.

As discussed in [13], the *Coxeter* Φ-*permutahedron* Π_Φ is the convex hull of the orbit of $\mathbf{s}_\Phi$ under the Weyl group W_Φ of type Φ. Thus $\mathbf{s}_\Phi$ is a distinguished vertex of Π_Φ. Note that the symmetric group S_n is the Weyl group of standard type $\Phi = A_{n-1}$ and the Weyl vector is $\mathbf{s}_n = \mathbf{w}_n - \frac{n-1}{2}\mathbf{1}_n$, so Π'_{n-1} (see (4.1.1) above) is the Φ-*permutahedron* of standard type $\Phi = A_{n-1}$.

In [69], we showed that Π_Φ is precisely the set of all possible *mean* score sequences of *random* Coxeter tournaments, thereby establishing a Coxeter analogue of a classical result of Moon [91]. The next work in this series [68] focused on deterministic Coxeter tournaments. The set Score(Φ) of all score sequences of Coxeter tournaments was classified, generalizing the classical result of Landau [73] discussed above.

The precise characterization of Score(Φ) is somewhat technical, involving a certain weak sub-majorization condition and additional parity conditions in types C_n and D_n. The proof is constructive, in that it shows how to build a Coxeter tournament with any given score sequence. See [68, Theorem 4] for more details.

4.2.2 Interchange graphs

The set Tour($\Phi, \mathbf{s}$) of all Coxeter tournaments on $\mathcal{K}_\Phi$ with given score sequence $\mathbf{s}$ was also investigated in [68]. Coxeter analogues of the interchange graphs IntGr($n, \mathbf{s}$) discussed above were introduced. Recall that the sets Tour($n, \mathbf{s}$) are generated by the cyclic triangle Δ_c. In the Coxeter setting, there are additional generators.

In all types B_n, C_n and D_n, in addition to Δ_c, we also require a *balanced triangle* Δ_b. In type B_n, there are also three *neutral pairs* Ω_1, Ω_2 and Ω_3. On the other hand, in type C_n, there are also two *neutral clovers* Θ_1 and Θ_2. See Figures 4.5, 4.6 and 4.7. In figures depicting Coxeter tournaments, we will draw:

- competitive games as edges directed away from their winner,
- collaborative games as solid/dotted lines if won/lost,
- half edge solitaire games as half edges directed away/toward from their (only) endpoint if won/lost, and
- loop solitaire games as solid/dotted loops if won/lost.

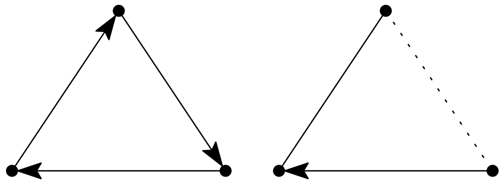

Figure 4.5: The cyclic and balanced triangles Δ_c and Δ_b are generators in all types B_n, C_n and D_n.

The *reversal* $\mathcal{T}^*$ of a Coxeter tournament $\mathcal{T} = (w_e : e \in E(\mathcal{S}))$ on $\mathcal{S}$ is obtained by reversing the outcome of all games in $\mathcal{T}$. That is, $\mathcal{T}^* = (w_e^* : e \in E(\mathcal{S}))$, where

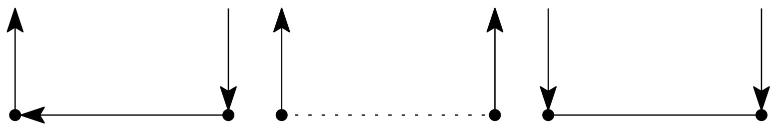

Figure 4.6: The neutral pairs Ω_1, Ω_2 and Ω_3 are additional generators in type B_n.

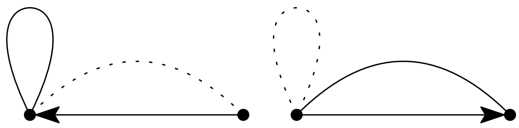

Figure 4.7: The neutral clovers Θ_1 and Θ_2 are additional generators in type C_n.

$w_e^* = 1 - w_e$. If $\mathcal{X} \subset \mathcal{T}$, we let $\mathcal{T} * \mathcal{X}$ denote the Coxeter tournament obtained from $\mathcal{T}$ by reversing the outcome of all games in $\mathcal{X}$. In particular, $\mathcal{T}^* = \mathcal{T} * \mathcal{T}$.

The *Coxeter interchange graph* $\mathrm{IntGr}(\Phi, \mathbf{s})$ has a vertex $v(\mathcal{T})$ for each $\mathcal{T} \in \mathrm{Tour}(\Phi, \mathbf{s})$. Vertices $v(\mathcal{T}_1), v(\mathcal{T}_2)$ are neighbours if $\mathcal{T}_2 = \mathcal{T}_1 * \mathcal{G}$, for some copy $\mathcal{G} \subset \mathcal{T}_1$ of a type Φ generator. If $\mathcal{G}$ is a neutral clover, we add a double edge, and otherwise we add a single edge.

For instance, the "snare drum," in Figure 4.3 above, is $\mathrm{IntGr}(C_3, \mathbf{s})$ when $\mathbf{s} = (-1, 0, 1)$.

The decision to represent clovers as double edges might seem arbitrary at first sight, however, there is a good reason. As it turns out, rather miraculously, this adjustment makes the type C_n interchange graphs degree regular. Furthermore, the degree of $\mathrm{IntGr}(\Phi, \mathbf{s})$ is related to distances in Π_Φ in the following way. Let $\|\mathbf{x}\|^2 = \sum_i x_i^2$ denote the squared length of $\mathbf{x} \in \mathbb{R}^n$.

Recall that $\mathrm{Score}(\Phi)$ is the set of all possible score sequences of Coxeter tournaments on $\mathcal{K}_\Phi$. As discussed above, this set is classified in [68].

Theorem 4.2.1 ([68])**.** *Let $\Phi = B_n$, C_n or D_n. Fix any $\mathbf{s} \in \mathrm{Score}(\Phi)$. Then the Coxeter interchange graph $\mathrm{IntGr}(\Phi, \mathbf{s})$ is regular, with degree given by*

$$d(\Phi, \mathbf{s}) = \frac{\|\mathbf{s}_\Phi\|^2 - \|\mathbf{s}\|^2}{2},$$

where $\mathbf{s}_\Phi$ is the standard score sequence.

In particular, $d(\Phi, \mathbf{s}) = O(n^3)$.

We observe that, as $\mathbf{s}$ moves closer to the center $\mathbf{0}_n$ of the polytope Π_Φ, the degree $d(\Phi, \mathbf{s})$ of $\mathrm{IntGr}(\Phi, \mathbf{s})$ increases. This is in line with the intuition that Coxeter tournaments with $\mathbf{s}$ closer to $\mathbf{0}_n$ (i.e., closer to being neutral) should contain more copies of the (neutral) generators.

Let us note that $\mathbf{s}$ with $\|\mathbf{s}\|^2 = \|\mathbf{s}_\Phi\|^2$ are precisely the vertices of Π_Φ. For such $\mathbf{s}$, we have $d(\Phi, \mathbf{s}) = 0$, in line with the fact there is a unique Coxeter tournaments with score sequence $\mathbf{s}$. Indeed, such a tournament is transitive, in the sense that it contains no copy of a neutral generator, and so its interchange graph is single isolated vertex.

In [68], we observed that such a result also holds for graph tournaments, in relation to the standard permutahedron, yielding a geometric interpretation of the classical result (see, e.g., Moon [92]) that any two tournaments with the same win/score sequence have the same number of cyclic triangles.

In closing, let us emphasize the the neutral generators in Figures 4.5, 4.6 and 4.7 were identified in [68]. Theorem 4.2.1, proved therein, identifies the degree of the interchange graphs. However, in the current work, we will show (see Theorem 4.4.1 below) that the interchange graphs are connected. It is this result that justifies calling these structures "generators," in the sense that the entire space $\mathrm{Tour}(\Phi, \mathbf{s})$ is obtained by iteratively reversing copies of these specific neutral structures.

4.2.3 Path coupling

We recall that an aperiodic, irreducible discrete-time Markov chain (X_n) on a finite state space Ω has a unique *equilibrium* π on S such that, for all $x, y \in \Omega$,

$$p_n(x, y) = \mathbf{P}\{X_n = y \mid X_0 = x\} \to \pi(y),$$

as $n \to \infty$. We note that $\pi(y)$ is the asymptotic proportion of time spent at state $y \in \Omega$. The maximal total variation distance from π by time n,

$$\tau(n) = \max_{x \in \Omega} \frac{1}{2} \sum_{y \in S} |p_n(x, y) - \pi(y)|,$$

is non-increasing. The *mixing time* is defined as

$$t_{\mathrm{mix}} = \inf\{n \geq 0 : \tau(n) \leq 1/4\}.$$

A Markov chain is said to be *rapidly mixing* if t_{mix} is bounded by a polynomial in $\log |\Omega|$.

Path coupling was introduced by Bubley and Dyer [25]. See, e.g., Aldous and Fill [10, Sec. 12.1.12] or Levin, Peres and Wilmer [79, Sec. 14.2] for reformulations of the original result that are closer in appearance to that of the following.

Consider a connected graph $G = (V, E)$. The *graph distance* $\delta(x, y)$ is the minimal number of edges in a path between x and y. The *diameter* is $D = \max_{x,y \in V} \delta(x, y)$ is the maximal length of such a path in G.

Definition 4.2.2. *We say that $G = (V, E)$ is weighted by w if each edge $\{u, v\} \in E$ is assigned some weight $w(u, v) \geq 1$. The weighted distance $w(x, y)$ is the minimal total weight path between x and y. Likewise, $D_w = \max_{x,y \in V} w(x, y)$ is the weighted diameter.*

The usual graph distance δ corresponds to the w for which $w(u, v) = 1$ for all $\{u, v\} \in E$.

Theorem 4.2.3 (Path coupling, [25])**.** *Consider a Markov chain (X_n) on a connected graph $G = (V, E)$, weighted by w. Suppose that, for some $\alpha > 0$, for each $\{x', x''\} \in E$ there is a coupling (X_1', X_1'') with $(X_0', X_0'') = (x', x'')$ so that*

$$\mathbf{E}\left[w(X_1', X_1'')\right] \leq (1 - \alpha) w(x', x'').$$

Then (X_n) mixes in time $t_{\text{mix}} = O(\alpha^{-1} \log D_w)$.

Often this result is applied with $w = \delta$, and, indeed, this will suffice for us in types B_n and D_n. In this case, path coupling has the intuitive interpretation that if the chain is "contractive" in expectation, then it is rapidly mixing.

On the other hand, in the more complicated type C_n, we will select a careful re-weighting w that takes into account some of the more intricate features in the interchange graphs of this type.

Finally, note that the type C_n interchange graphs are, in fact, *multi*graphs. Specifically, some pairs of vertices (corresponding to clover reversals) are joined by double edges, as in Figures 4.3 and 4.4 above. This is for technical convenience, as it makes the graph regular, and thereby our coupling procedure easier to explain. We note that Theorem 4.2.3 still applies, since a Markov chain (Y_n) on a multigraph M with some double edges is equivalent to the Markov chain (X_n) on the graph G, obtained by collapsing each double edge in M into a single edge, and combining the two associated edge crossing probabilities.

4.3 Main result

Our main result shows that random walks rapidly mix on the Coxeter interchange graphs.

Theorem 4.3.1. *Let $\Phi = B_n$, C_n or D_n. Fix any $\mathbf{s} \in \text{Score}(\Phi)$. Then lazy simple random walk $(\mathcal{T}_n : n \geq 0)$ on $\text{IntGr}(\Phi, \mathbf{s})$ mixes in time $t_{\text{mix}} = O(n^3 \log n)$ if $\Phi = B_n$ or D_n, and in time $t_{\text{mix}} = O(n^4 \log n)$ if $\Phi = C_n$.*

Rapid mixing in type A_{n-1}, proved in [64, 89], follows as a special case of our proof of this result in type D_n.

We note that the classification of Score(Φ) in [69] is constructive, which allows us to initialize the random walk in the first place.

In fact, we will prove sharper bounds (see Theorem 4.6.1 and 4.6.3 below). In types B_n and D_n, we will show that $t_{\text{mix}} = O(d \log n)$, where d is the degree (see Theorem 4.2.1 above) of the interchange graph $\text{IntGr}(\Phi, \mathbf{s})$. The result above follows, since $d = O(n^3)$. In type C_n, we will show that $t_{\text{mix}} = O(\gamma d \log n)$, where γ is a certain quantity (see Lemma 4.5.10) satisfying $\gamma \leq \min\{d, 2n\}$. We call γ the *maximal crystal degree* of the interchange graph. Roughly speaking, it is maximal number of crystals, all containing the same double edge. This quantity is related to the re-weighting w that we will use in applying Theorem 4.2.3 in type C_n. We note that re-weighting arguments have been used before, e.g., in the work of Wilson [105].

4.4 Connectivity

Before proving Theorem 4.3.1, we will first establish the following combinatorial result, giving further structural information (beyond its regularity, given by Theorem 4.2.1) about the Coxeter interchange graphs.

Theorem 4.4.1. *Let $\Phi = B_n$, C_n or D_n. Fix any $\mathbf{s} \in \text{Score}(\Phi)$. The Coxeter interchange graph $\text{IntGr}(\Phi, \mathbf{s})$ is connected and its diameter $D = O(n^2)$.*

This result is a corollary of Lemma 4.4.9 (the "reversing lemma") proved at the end of this section. A number of preliminaries are required. First, in the next subsection, we will find a way of encoding Coxeter tournaments as special types of directed graphs.

4.4.1 Z-frames

Oriented signed graphs were studied by Zaslavsky [108]. From this point of view (see [108, Fig. 1]), each oriented signed edge in an oriented signed graph $\mathcal{S}$ is the union of at most two directed half edges. We modify this idea, by adding a named endpoint to each half-edge, which we call a *match*. This will allow us to prove certain structural facts using graph theory techniques. We call such a structure a *Z-frame*.

Definition 4.4.2. *A Z-frame $\mathbf{Z}$ is a directed, bipartite multigraph on disjoint sets of players V and matches M, such that every match has degree 1 or 2.*

This concept is fairly general, and not all Z-frames correspond to a Coxeter tournament $\mathcal{T}$ on some $\mathcal{S} \subset \mathcal{K}_\Phi$. However, each such $\mathcal{T}$ has a unique representation as a Z-frame $\mathbf{Z}(\mathcal{T})$. We think of $\mathbf{Z}(\mathcal{T})$ as revealing the "inner directed graph structure" of $\mathcal{T}$. Players in $\mathbf{Z}(\mathcal{T})$ correspond to vertices in $\mathcal{T}$. Recall that each game in $\mathcal{T}$ corresponds to an oriented signed edge. Each such game is associated with a match in $\mathbf{Z}(\mathcal{T})$.

We will think of edges directed away/toward players $v \in V$ as positively/negatively charged. (That being said, positive/negative edges in a Z-frame should not be confused with positive/negative edges in a Coxeter tournament.)

We say that $\mathbf{Z}$ is *neutral* if all players $v \in V$ have net zero charge, i.e., $\deg^+(v) - \deg^-(v) = 0$, where $\deg^\pm(v)$ is the number of positive/negative edges incident to v. We put $\deg(v) = \deg^+(v) + \deg^-(v)$. Note that, if $\mathbf{Z}$ is *neutral* then $\deg(v)$ is even.

Definition 4.4.3. *Let $\mathcal{T}$ be a Coxeter tournament on a signed graph $\mathcal{S}$ on $[n]$. Let $\mathbf{Z}(\mathcal{T})$ be the Z-frame on $V = [n]$ and $M = \{m_e \colon e \in \mathcal{S}\}$, with the following directed edges:*

- *If a competitive game $e_{ij}^- \in \mathcal{S}$ between players i and j is won (resp. lost) by i in $\mathcal{T}$, we include two directed edges $i \to m_{ij}^- \to j$ (resp. $i \leftarrow m_{ij}^- \leftarrow j$).*
- *If a collaborative game $e_{ij}^+ \in \mathcal{S}$ between players i and j is won (resp. lost) in $\mathcal{T}$, we include two directed edges $i \to m_{ij}^+ \leftarrow j$ (resp. $i \leftarrow m_{ij}^+ \to j$).*
- *If a half edge solitaire game $e_i^h \in \mathcal{S}$ by player i is won (resp. lost) in $\mathcal{T}$, we include one directed edge $i \to m_i^h$ (resp. $i \leftarrow m_i^h$).*
- *If a loop solitaire game $e_i^\ell \in \mathcal{S}$ by player i is won (resp. lost) in $\mathcal{T}$, we include two directed edges $i \rightrightarrows m_i^\ell$ (resp. $i \leftleftarrows m_i^\ell$).*

See Figure 4.8 for an example of a Coxeter tournament $\mathcal{T}$ and its corresponding Z-frame $\mathbf{Z}(\mathcal{T})$.

The reason for the above terminology is that directed edges in $\mathbf{Z}(\mathcal{T})$ with a positive/negative charge correspond to positive/negative contributions to the score sequence $\mathbf{s}(\mathcal{T})$. Note that, in competitive/collaborative games, the two charges are opposing/aligned (regardless of the outcome of the game). See Figure 4.1.

4.4.2 Decomposing Z-frames

Our aim is to decompose neutral Coxeter tournaments into irreducible neutral parts. In this section, we will do this for Z-frames in general.

Recall that a *trail* is a walk in which no edge is visited twice.

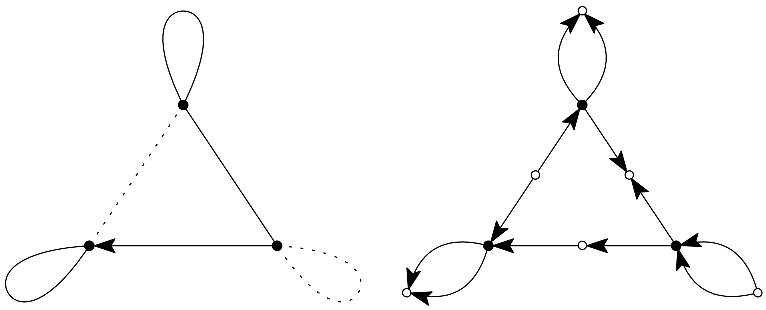

Figure 4.8: A C_3-tournament $\mathcal{T}$ and its Z-frame $\mathbf{Z}(\mathcal{T})$. Players/matches are drawn as black/white dots.

Table 4.1: Signed edges e in a signed graph $\mathcal{S}$ can be oriented in one of two ways $w_e \in \{0, 1\}$ by a Coxeter tournament $\mathcal{T}$. The contribution to the score sequence $\mathbf{s}(\mathcal{T})$ and representation in the Z-frame $\mathbf{Z}(\mathcal{T})$ are given below.

e	w_e	$\mathbf{s}$	$\mathbf{Z}$
e_{ij}^-	1	$+\mathbf{e}_{ij}^-/2 = +\mathbf{e}_i/2 - \mathbf{e}_j/2$	$i \to m_{ij}^- \to j$
	0	$-\mathbf{e}_{ij}^-/2 = -\mathbf{e}_i/2 + \mathbf{e}_j/2$	$i \leftarrow m_{ij}^- \leftarrow j$
e_{ij}^+	1	$+\mathbf{e}_{ij}^+/2 = +\mathbf{e}_i/2 + \mathbf{e}_j/2$	$i \to m_{ij}^+ \leftarrow j$
	0	$-\mathbf{e}_{ij}^+/2 = -\mathbf{e}_i/2 - \mathbf{e}_j/2$	$i \leftarrow m_{ij}^+ \to j$
e_i^h	1	$+\mathbf{e}_i^h/2 = +\mathbf{e}_i/2$	$i \to m_i^h$
	0	$-\mathbf{e}_i^h/2 = -\mathbf{e}_i/2$	$i \leftarrow m_i^h$
e_i^ℓ	1	$+\mathbf{e}_i^\ell/2 = +\mathbf{e}_i$	$i \rightrightarrows m_i^\ell$
	0	$-\mathbf{e}_i^\ell/2 = -\mathbf{e}_i$	$i \leftleftarrows m_i^\ell$

Definition 4.4.4. *We call a trail* $\mathbf{T}$ *of (directed) edges in a Z-frame* closed *if it starts and ends at the same vertex, and otherwise we call it* open*. A trail is* neutral *if the two consecutive edges at each player $v \in V$ have opposite charges. The length ℓ of a trail is its number of matches.*

Note that the edges along a trail in a Z-frame do not repeat, and are connected to each other in alternation by a player/match. Also note that neutral open trails start and end at distinct final matches (the left and rightmost matches along the trail). See Figure 4.9.

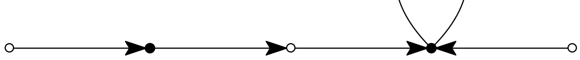

Figure 4.9: An open neutral trail of length $\ell = 4$.

Lemma 4.4.5. *Any neutral Z-frame* $\mathbf{Z}$ *can be decomposed into an edge-disjoint union* $\mathbf{Z} = \bigcup_i \mathbf{T}_i$ *of closed neutral trails and open neutral trails, such that no two open trails have a common final match.*

Proof. Let $v \in V$. Since $\mathbf{Z}$ is neutral, $\deg_+(v) = \deg_-(v)$. Therefore we may pair each edge directed away from v with an edge directed toward v. Note that any such pair is a neutral trail. We let $\mathcal{P}_v$ denote the set of all such pairs. Then $\bigcup_{v\in V} \mathcal{P}_v$ is a decomposition of $\mathbf{Z}$ into an edge-disjoint union of neutral trails. Select a decomposition $\mathbf{Z} = \bigcup_i \mathbf{T}_i$ with the minimal number of neutral trails. Open trails in this decomposition cannot have a common final match, as otherwise they could be concatenated to form a single longer trail, contradicting minimality. □

Let $\mathbf{Z}$ be a neutral Z-frame. We say that $\mathbf{Z}$ is *reducible* if it contains a non-empty neutral $\mathbf{Z}' \subsetneq \mathbf{Z}$. Otherwise, $\mathbf{Z}$ is *irreducible*.

Each vertex v in a neutral trail has even degree, since $\deg^+(v) = \deg^-(v)$ and $\deg(v) = \deg^+(v) + \deg^-(v)$. As discussed, vertices v in a neutral $\mathbf{Z}$ have even $\deg(v)$. Our next result observes that if $\mathbf{Z}$ is irreducible, then all $\deg(v) \leq 4$.

Lemma 4.4.6. *Let* $\mathbf{Z}$ *be an irreducible neutral Z-frame. Then* $\mathbf{Z}$ *is a neutral trail and* $\deg(v) \in \{0, 2, 4\}$ *for all* $v \in V$.

Proof. By Lemma 4.4.5, $\mathbf{Z}$ is a neutral trail, and so all $\deg(v)$ are even. If v is an isolated vertex that plays no solitaire games, then $\deg(v) = 0$. Otherwise, if $\mathbf{Z}$ is non-trivial, we will argue that $\deg(v) \in \{2, 4\}$ are the only possibilities. To see this, start at any v along the trail, and then follow the trail. Consider the charges of the edges incident to v, in the order in which they are visited by the trail. If the first and second charges are opposing, then the trail is complete with $\deg(v) = 2$. Otherwise, suppose they are both positive (the other case is symmetric). Since $\mathbf{Z}$ is neutral, the 3rd charge is negative. Finally, consider the 4th charge. If it were positive, then the sub-trail between the third negative edge incident to v and the forth positive edge incident to v would be neutral. Therefore, the 4th charge is negative, and so the trail is complete with $\deg(v) = 4$. □

4.4.3 Reversing Coxeter tournaments

Applying the results of the previous section, we obtain the following result for Z-frames $\mathbf{Z}(\mathcal{T})$ of Coxeter tournaments $\mathcal{T}$.

Lemma 4.4.7. *Let* $\Phi = B_n$, C_n *or* D_n. *Let* $\mathbf{Z}(\mathcal{T})$ *be the Z-frame of a neutral Coxeter tournament* $\mathcal{T}$ *on a signed* Φ*-graph* $\mathcal{S} \subset \mathcal{K}_\Phi$. *Consider a decomposition* $\mathbf{Z}(\mathcal{T}) = \bigcup_i \mathbf{T}_i$

into neutral trails given by Lemma 4.4.5. If $\Phi = C_n$ or D_n then all trails $\mathbf{T}_i$ are closed. If $\Phi = B_n$ then possibly some $\mathbf{T}_i$ are open.

Proof. This result follows by noting that if $\mathcal{T}$ is of type C_n or D_n then all matches in $\mathbf{Z}(\mathcal{T})$ are degree 2. Therefore, there are no open neutral trails in the decomposition, so each $\mathbf{T}_i$ is closed. On the other hand, in type B_n it is possible to have open and closed trails, since in this case possibly some matches are degree 1. □

Note that a tournament $\mathcal{T}$ on a signed graph $\mathcal{S}$ is neutral if and only if its Z-frame $\mathbf{Z}(\mathcal{T})$ is neutral. Naturally, we call such a $\mathcal{T}$ *irreducible* if $\mathbf{Z}(\mathcal{T})$ is irreducible, and *reducible* otherwise.

Lemma 4.4.8. *Suppose that $\mathcal{S}$ is a D_n-graph in which all players $v \in V$ have degree four. Then any neutral tournament $\mathcal{T}$ on $\mathcal{S}$ is reducible.*

Proof. The proof is by contradiction. Suppose that $\mathcal{T}$ on $\mathcal{S}$ is neutral and that $\mathbf{Z}(\mathcal{T})$ is irreducible. By Lemma 4.4.5, $\mathbf{Z}(\mathcal{T})$ is in fact a single neutral trail, which for convenience we will denote by $\mathbf{T}$. Following the consecutive edges of $\mathbf{T}$, we can find a closed trail $\mathbf{T}' \subset \mathbf{T}$ which:

- starts and ends at some player v,
- visits no player $u \neq v$ more than once along the way,
- and is neutral everywhere except at v.

Let us assume that the two directed edges in $\mathbf{T}'$ incident to v are positive, since the other case is symmetric.

Consider the extension $\mathbf{T}' \cup \mathbf{T}''$ of $\mathbf{T}'$, as it departs v via some negatively charged directed edge, until it eventually returns to some player $u \in \mathbf{T}'$ for the first time. Note that $u \neq v$, as else, since $\mathbf{T}$ is irreducible, it would follow that $\mathbf{T} = \mathbf{T}' \cup \mathbf{T}''$, and then that there are degree 2 vertices in $\mathbf{T}$ along $\mathbf{T}'$. Since $\mathbf{T}'$ is neutral at u, the edges in $\mathbf{T}'$ incident to u have opposing charges. Let $\mathbf{T}' = \mathbf{T}'_+ \cup \mathbf{T}'_-$, where $\mathbf{T}'_\pm \subset \mathbf{T}'$ is the trail between u and v which includes the positive/negative edge in $\mathbf{T}'$ incident to u, as in Figure 4.10. To conclude, consider the charge of the directed edge in $\mathbf{T}''$ along which u is revisited. To obtain the required contradiction, note that if this charge is positive (resp. negative) then $\mathbf{T}'_- \cup \mathbf{T}''$ (resp. $\mathbf{T}'_+ \cup \mathbf{T}''$) is a neutral trail. □

Finally, we turn our attention to the main result of this section.

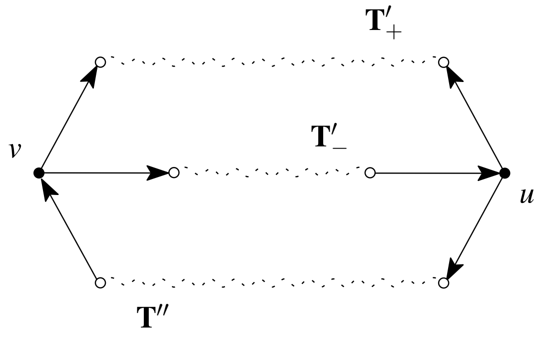

Figure 4.10: In this instance, the directed edge in $\mathbf{T}''$ along which u is revisited is positive, so $\mathbf{T}'_- \cup \mathbf{T}''$ is a neutral trail. If it were negative, then $\mathbf{T}'_+ \cup \mathbf{T}''$ would be neutral. Curved dotted lines represent the continuation of a trail.

Lemma 4.4.9 (Reversing lemma)**.** *Let $\Phi = B_n$, C_n or D_n. Let $\mathcal{T}$ be a tournament on the complete signed graph $\mathcal{K}_\Phi$, and let $\mathcal{I} \subset \mathcal{T}$ be a neutral sub-tournament. If $\mathcal{I}$ has $\ell \geq 3$ games then $\mathcal{I}$ can be reversed in a series of at most $\ell - 2$ type Φ generator reversals.*

The assumption that $\mathcal{T}$ is on $\mathcal{K}_\Phi$ is crucial, since not all of the games in the generators used to reverse $\mathcal{I}$ will be in $\mathcal{I}$ itself. However, after the series of reversals, only the games in $\mathcal{I}$ will have been reoriented, i.e., all games in $\mathcal{T} \setminus \mathcal{I}$ will be restored to their initial orientations.

Proof. Without loss of generality we may assume that $\mathcal{I}$ is irreducible. In this case, by Lemma 4.4.5, the Z-frame $\mathbf{Z}(\mathcal{I})$ is a single neutral trail. For simplicity, we will speak of $\mathcal{I}$ as a tournament and trail interchangeably.

We first address the simplest case of open neutral trails, which appears only in type B_n. We proceed by induction on the length ℓ. The smallest open trails, with $\ell = 3$, are the neutral pairs Ω_i (as in Figure 4.6) themselves, which are clearly reversible in a single reversal. For longer open neutral trails $\mathcal{I}$, consider any (half edge) solitaire game h in $\mathcal{T}$ played by some u in $\mathcal{I}$, which is not an endpoint of the trail. Using h, we can first reverse either the part of the trail that is to the "left" of u or else the part which is to the "right." We can do this using the inductive hypothesis, as one of the parts is neutral and both have length smaller than ℓ. Then, using the reversal of this game h^*, we can reverse the other part of the trail in turn. See Figure 4.11 for an example.

Next, we turn to the case of closed neutral trails. Recall that closed trails do not have solitaire half edge games, so from this point on we assume that $\Phi = C_n$ or D_n. If $\mathcal{I}$ is a generator, then the result clearly holds. If it is not, then let $k \geq 3$ be the

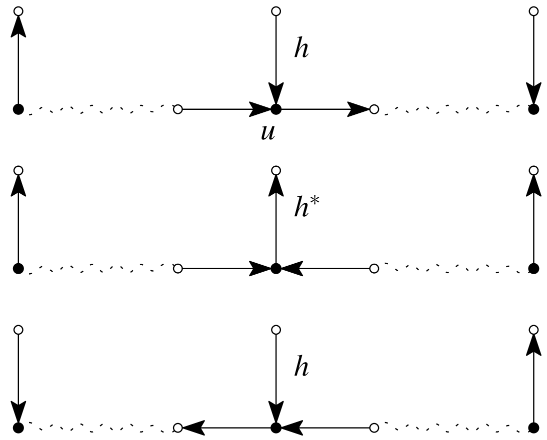

Figure 4.11: In type B_n: Reversing an open neutral trail of length $\ell > 3$, using a (half edge) solitaire game h played by some u in the "middle" of the trail. In this example, from top to bottom, we first reverse the "right" side of $\mathcal{I}$ using h, and then the "left" side using h^*.

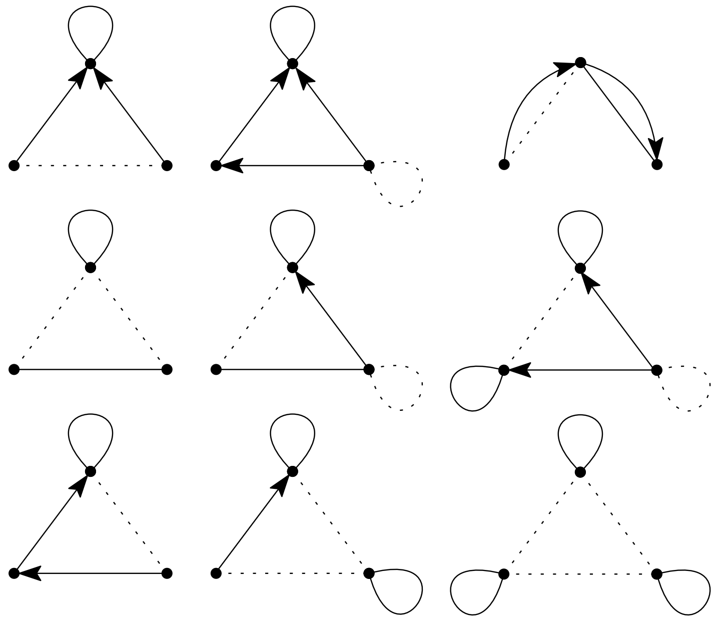

Figure 4.12: Up to symmetry, it suffices to consider the above neutral tournaments on three vertices. Note that each of these tournaments with ℓ games can be reversed in $\ell - 2$ steps.

number of vertices in $\mathcal{I}$. If $k = 3$, one can verify directly that the result holds. See Figure 4.12.

For $k \geq 4$, we aim to find a pair of vertices i, j in $\mathcal{I}$ which do not play a game with each other in $\mathcal{I}$. Once we find such a pair, the argument is similar to the case of open trails above; the difference being that, in the case of closed trails, we will either first reverse the "top/bottom" (instead of the "left/right") of the trail, and then the other side in turn, as in Figure 4.13.

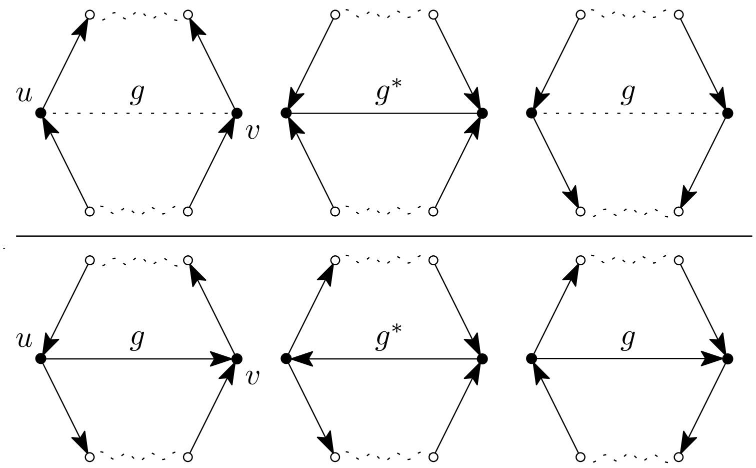

Figure 4.13: Reversing $\mathcal{I}$ when its Z-frame is a closed neutral trail. *Above:* If the charges of the edges incident to u and v are "aligned," we use the collaborative game g between u, v to reverse $\mathcal{I}$. *Below:* Otherwise, if they are "unaligned," we use the competitive game between u and v.

To this end, suppose that every pair of vertices in $\mathcal{I}$ plays at least one game with each other. As $k \geq 4$, this implies that the degree of every player in $\mathcal{I}$ is at least 3. Moreover, if some v in $\mathcal{I}$ plays a loop game, then $\deg(v) \geq 5$. However, as $\mathcal{I}$ is neutral and irreducible, this would contradict Lemma 4.4.6. Thus, $\mathcal{I}$ is a 4-regular D_n-graph. But this is also impossible, by Lemma 4.4.8. Therefore, for all $k \geq 4$, there exists a pair i, j of players which do not play a game with each other, and this concludes the proof. □

Finally, using the reversing lemma, we will prove the main result of this section.

Proof of Theorem 4.4.1. Let $\mathcal{T}, \mathcal{T}' \in \mathrm{Tour}(\Phi, \mathbf{s})$. The distance between $\mathcal{T}$ and $\mathcal{T}'$ is the smallest number of the generator reversals which transforms $\mathcal{T}$ into $\mathcal{T}'$. The games in the difference $\mathcal{D} = \mathcal{T} \setminus \mathcal{T}'$ are precisely those which need to be reversed. Since $\mathbf{s}(\mathcal{T}) = \mathbf{s}(\mathcal{T}')$, it follows that $\mathcal{D}$ is neutral. Therefore, by Lemma 4.4.9, there is a path from $v(\mathcal{T})$ to $v(\mathcal{T}')$ in $\mathrm{IntGr}(\Phi, \mathbf{s})$ of length $O(n^2)$. Hence $\mathrm{IntGr}(\Phi, \mathbf{s})$ is connected and its diameter $D = O(n^2)$. □

4.5 Interchange networks

In this section, for ease of exposition, we will speak of Coxeter tournaments $\mathcal{T}$ and their corresponding vertices $v(\mathcal{T})$ in IntGr interchangeably.

Definition 4.5.1. *For $\mathcal{T}_1, \mathcal{T}_2 \in \mathrm{Tour}(\Phi, \mathbf{s})$ at distance two in $\mathrm{IntGr}(\Phi, \mathbf{s})$, we define the interchange network $N(\mathcal{T}_1, \mathcal{T}_2)$ to be the union over all paths of length two between $\mathcal{T}_1, \mathcal{T}_2$.*

Note that each path of length two between such $\mathcal{T}_1, \mathcal{T}_2$ corresponds to a way of reversing the *difference* $\mathcal{D} = \mathcal{T}_1 \setminus \mathcal{T}_2$ between $\mathcal{T}_1$ and $\mathcal{T}_2$.

4.5.1 Classifying networks

In this section, we will classify the possibilities for $(N, \mathcal{D})$. As we will see, this is the key to applying path coupling (Theorem 4.2.3 above) in Section 4.6 below. In the classical case of type A_{n-1} there is only one possibility for N (a "single diamond"), and this is the reason why such a simple contractive coupling (as in Figure 4.27 below) is possible. As it turns out, this continues to hold in types B_n and D_n, but the underlying reasons are more complicated. Type C_n, on the other hand, is significantly more complex, as then the structure of N can take various other forms.

It can be seen that any two distinct generators $\mathcal{G}_1 \neq \mathcal{G}_2$ are either *disjoint* $\mathcal{G}_1 \cap \mathcal{G}_2 = \emptyset$ or else have exactly one game g in common $\mathcal{G}_1 \cap \mathcal{G}_2 = \{g\}$. In this case, we say that $\mathcal{G}_1, \mathcal{G}_2$ are *adjacent*.

Note that if a path of length two from $\mathcal{T}_1$ to $\mathcal{T}_2$ passes through some $\mathcal{T}_{12}$, then there are two generators $\mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{T}_{12}$ such that $\mathcal{T}_i = \mathcal{T}_{12} * \mathcal{G}_i$, for $i \in \{1, 2\}$. In this way, every such path of length two from $\mathcal{T}_1$ to $\mathcal{T}_2$ is determined by a *midpoint* $\mathcal{T}_{12}$ and a pair of generators $\mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{T}_{12}$.

There are three possible networks when $\mathcal{G}_1, \mathcal{G}_2$ are disjoint. We call these the *single*, *double* and *quadruple diamonds*. See Figure 4.14. Recall that double edges in IntGr correspond to neutral clover reversals. All other types of reversals (neutral triangles and pairs) are represented as single edges.

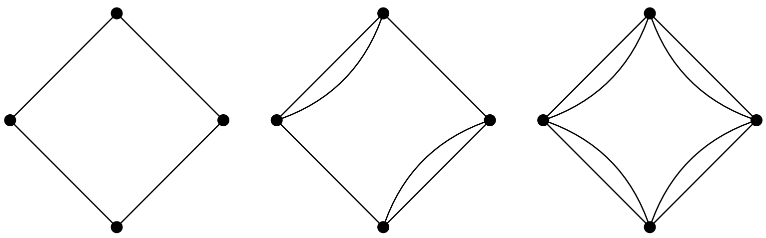

Figure 4.14: *Left to right:* The single, double and quadruple diamond interchange networks.

The following result classifies the types of networks $N(\mathcal{T}_1, \mathcal{T}_2)$ when $\mathcal{G}_1, \mathcal{G}_2$ are disjoint.

Lemma 4.5.2. *Suppose that there is a path of length two between $\mathcal{T}_1, \mathcal{T}_2$ in* $\mathrm{IntGr}(\Phi, \mathbf{s})$ *that passes through midpoint $\mathcal{T}_{12}$, with associated generators $\mathcal{G}_i \subset \mathcal{T}_{12}$ such that $\mathcal{T}_i = \mathcal{T}_{12} * \mathcal{G}_i$. Suppose that $\mathcal{G}_1, \mathcal{G}_2$ are disjoint. Then if exactly zero, one or two of the*

$\mathcal{G}_i$ *are clovers then the network* $N(\mathcal{T}_1, \mathcal{T}_2)$ *is a single, double or quadruple diamond, respectively.*

Proof. Clearly, there are exactly two paths from $\mathcal{T}_1$ to $\mathcal{T}_2$. These paths correspond to reversing the disjoint generators $\mathcal{G}_1^*, \mathcal{G}_2 \subset \mathcal{T}_1$ in series, in one of the two possible orders. See Figure 4.15. □

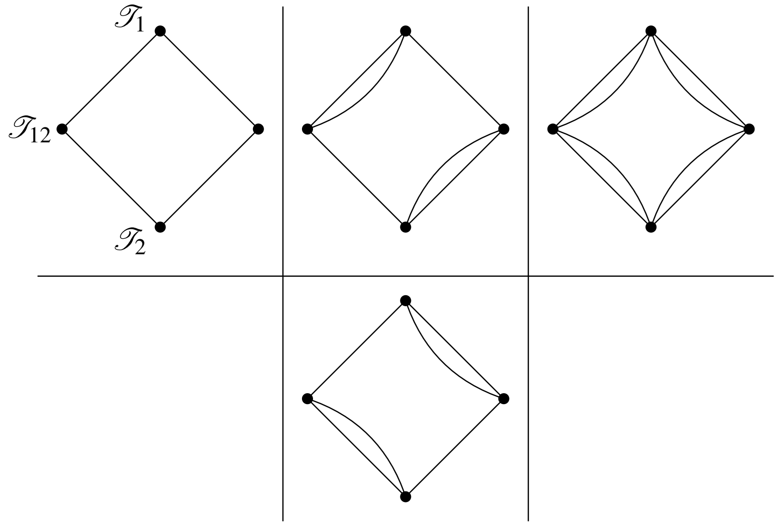

Figure 4.15: *Left to right:* $N(\mathcal{T}_1, \mathcal{T}_2)$ when exactly zero, one or two of the disjoint $\mathcal{G}_1, \mathcal{G}_2$ is a clover (single, double and quadruple diamonds). In each cell, the leftmost vertices are $\mathcal{T}_1$, $\mathcal{T}_{12}$ and $\mathcal{T}_2$, as in the first cell.

The case that $\mathcal{G}_1, \mathcal{G}_2$ are adjacent is more involved. It is useful to note that, in this case, the difference $\mathcal{D} = \mathcal{T}_1 \setminus \mathcal{T}_2$ is a neutral tournament with exactly four games on either three or four vertices. Even so, there are a number of cases to consider, and the key to a concise argument is grouping symmetric cases together. For a Coxeter tournament $\mathcal{T}$, we define its *projection graph* $\pi(\mathcal{T})$ to be the graph obtained by changing each:

- oriented negative/positive edge (i.e., competitive/collaborative game) into an undirected edge,
- oriented half edge (i.e., half edge solitaire game) into an undirected half edge,
- oriented loop (i.e., loop solitaire game) into an undirected loop.

There are a number of ways that two generators $\mathcal{G}_1, \mathcal{G}_2$ can be adjacent. However, there are only four possibilities for their projected difference $\pi(\mathcal{D})$. We call these the *square*, *tent*, *fork* and *hanger*, as in Figure 4.16. The following observation is essentially self-evident, and can be verified by an elementary case analysis. We omit the proof.

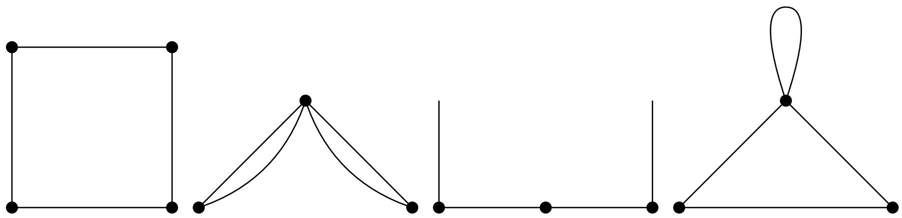

Figure 4.16: *Left to right:* The square, tent, fork, and hanger.

Lemma 4.5.3. *Assume the same set up as Lemma 4.5.2, except that instead $\mathcal{G}_1, \mathcal{G}_2$ are adjacent. Let $\mathcal{D} = \mathcal{T}_1 \setminus \mathcal{T}_2$. Then:*

1. *If $\mathcal{G}_1, \mathcal{G}_2$ are neutral triangles on four/three vertices then $\pi(\mathcal{D})$ is a square/tent.*
2. *If $\mathcal{G}_1, \mathcal{G}_2$ are neutral clovers then $\pi(\mathcal{D})$ is a tent.*
3. *If $\mathcal{G}_1$ is a neutral pair and $\mathcal{G}_2$ is a neutral pair or triangle then $\pi(\mathcal{D})$ is a fork.*
4. *If $\mathcal{G}_1$ is a neutral clover and $\mathcal{G}_2$ is a neutral triangle then $\pi(\mathcal{D})$ is a hanger.*

In addition to the single and double diamond networks in Figure 4.14, there are two additional networks that can occur when $\mathcal{G}_1, \mathcal{G}_2$ are adjacent. We call these the *split* and *heavy diamonds*, see Figure 4.17.

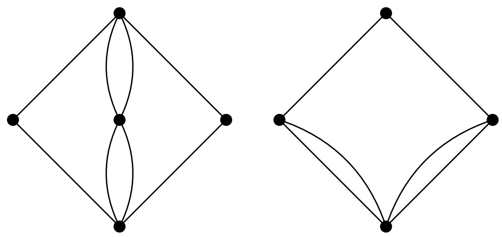

Figure 4.17: *Left to right:* The split and heavy diamond interchange networks.

The following result classifies the types of networks $N(\mathcal{T}_1, \mathcal{T}_2)$ when $\mathcal{G}_1, \mathcal{G}_2$ are adjacent.

Lemma 4.5.4. *Assume the same set up as Lemma 4.5.3 (with $\mathcal{G}_1, \mathcal{G}_2$ are adjacent). Let $\mathcal{D} = \mathcal{T}_1 \setminus \mathcal{T}_2$ and $N = N(\mathcal{T}_1, \mathcal{T}_2)$. Then:*

1. *If $\Phi = B_n$ or D_n, then N is a single diamond.*
2. *If $\Phi = C_n$ and $\pi(\mathcal{D})$ is a square, then N is a single diamond.*
3. *If $\Phi = C_n$ and $\pi(\mathcal{D})$ is a tent, then N is a split diamond.*
4. *If $\Phi = C_n$ and $\pi(\mathcal{D})$ is a hanger, then N is a double or heavy diamond.*

Proof. **Case 1a.** We start with the simplest case that $\Phi = B_n$ or D_n and $\mathcal{G}_1, \mathcal{G}_2$ are adjacent neutral triangles on four vertices, so that $\pi(\mathcal{D})$ is a square. Note that, to reverse $\mathcal{D}$ in two steps, we must reverse exactly two edges in $\mathcal{D}$ in each step. As such, no neutral pairs will be involved in reversing $\mathcal{D}$ in two steps. There are exactly two ways to reverse $\mathcal{D}$. For each pair of "antipodal" vertices in $\mathcal{D}$, consider the two games played between the pair. Exactly one of the two games g allows us to reverse the games in $\mathcal{D}$ on one "side" of g. Then, in turn, we can use g^* to reverse the other two games in $\mathcal{D}$. See Figure 4.18 for all the possible cases of $\mathcal{D}$.

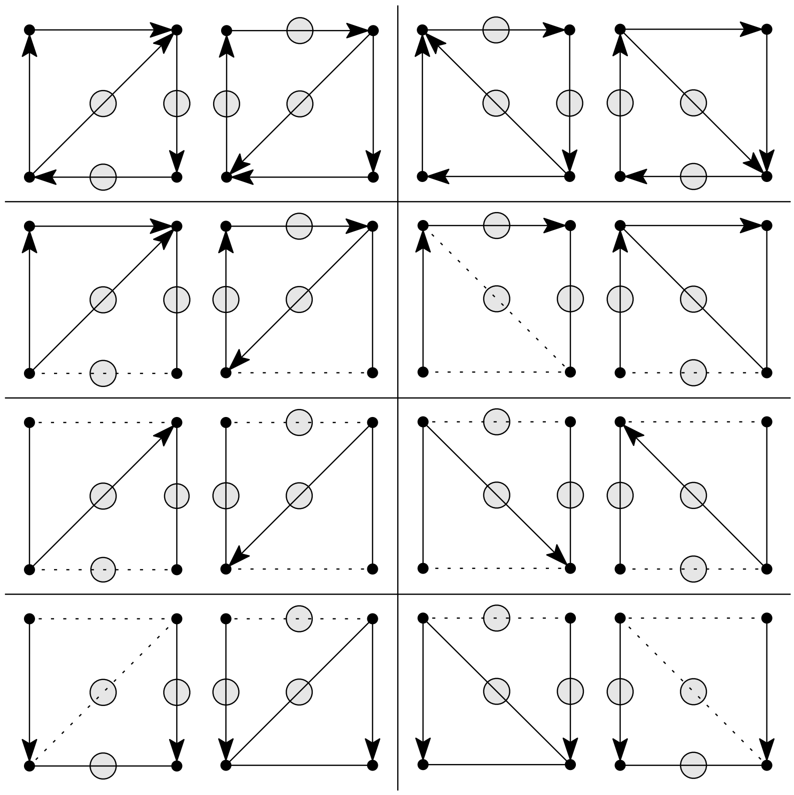

Figure 4.18: Reversing when $\pi(\mathcal{D})$ is a square in types B_n, C_n or D_n. Each row corresponds to one of the four possible configurations of $\mathcal{D}$. In each cell, the shaded circles indicate how to reverse $\mathcal{D}$ in two steps, using one of the games between an "antipodal" pair of players along $\mathcal{D}$. Note that each row has exactly two cells, as there is always exactly two ways to reverse $\mathcal{D}$ in two steps. The "middle" game (not in the square itself) is reversed twice, so returned to its original orientation.

Case 1b. Suppose that $\Phi = B_n$ or D_n and $\mathcal{G}_1, \mathcal{G}_2$ are adjacent neutral triangles on three vertices, so that $\pi(\mathcal{D})$ is a tent. In this case, both of the (competitive and collaborative) games between the "base" vertices lead to a way of reversing $\mathcal{D}$. See Figure 4.19.

Case 1c. Suppose that $\Phi = B_n$ and that one of $\mathcal{G}_1, \mathcal{G}_2$ is a neutral pair and the other is an adjacent neutral pair or triangle, so that $\pi(\mathcal{D})$ is a fork. Then the half

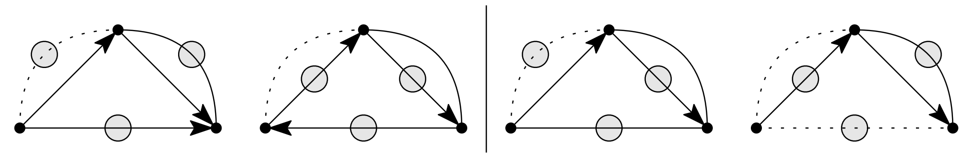

Figure 4.19: Reversing when $\pi(\mathcal{D})$ is a tent formed by two neutral triangles in types B_n and D_n. This figure has only one row, as there is only one possibility for $\mathcal{D}$.

edge game played by the “middle” vertex and exactly one of the games between the “base” vertices lead to ways of reversing $\mathcal{D}$. See Figure 4.20.

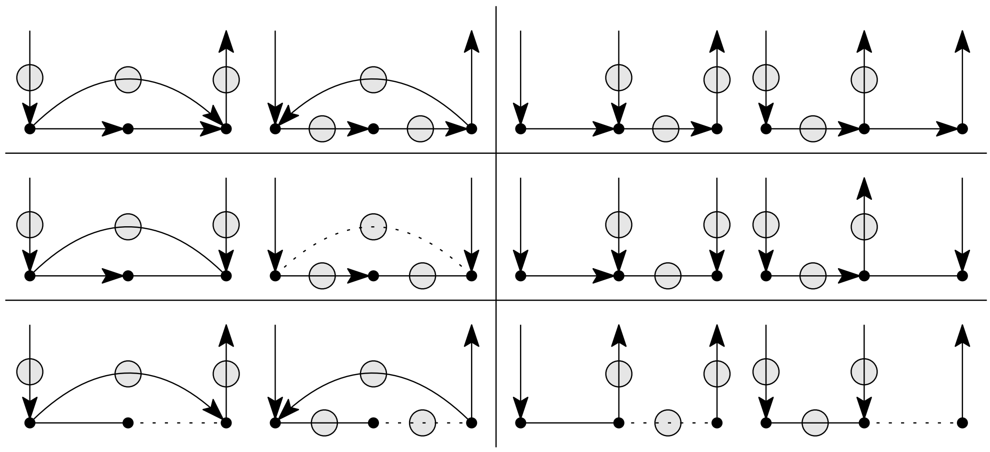

Figure 4.20: Reversing when $\pi(\mathcal{D})$ is a fork in type B_n. Each row corresponds to a possible configuration for $\mathcal{D}$.

By Cases 1a–c, statement (1) follows, that is, in types B_n and D_n the network N is always a single diamond, as in Figure 4.21.

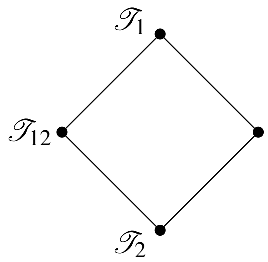

Figure 4.21: When $\mathcal{G}_1, \mathcal{G}_2$ are adjacent, the only possible $N(\mathcal{T}_1, \mathcal{T}_2)$ in types B_n and D_n is a diamond.

Case 2. In type C_n, the case that $\pi(\mathcal{D})$ is a square follows by the same argument as in types B_n and D_n. Indeed, recall that any reversal of $\mathcal{D}$ in two steps will involve reversing exactly two games in $\mathcal{D}$ in each step. Therefore, no clovers will be involved in such a reversal of $\mathcal{D}$, and so once again N is a single diamond, yielding statement (2).

Case 3a. Suppose that $\Phi = C_n$ and that $\pi(\mathcal{D})$ is a tent formed by two adjacent neutral triangles on three vertices. Then by Case 1b, N contains a single diamond. However, using the loop game ℓ played by the "middle" vertex, we obtain an additional path of length two between $\mathcal{T}_1, \mathcal{T}_2$. We can use ℓ to reverse the two games on one "side" of the tent. Then, in turn, we can use ℓ^* to reverse the other two games. See Figure 4.22. Hence N is a split diamond in this case.

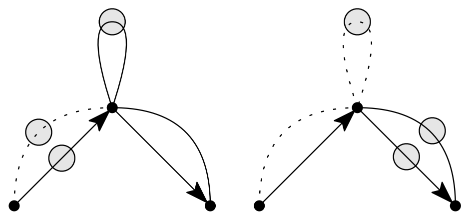

Figure 4.22: The additional way of reversing when $\pi(\mathcal{D})$ is a tent in type C_n.

Case 3b. Suppose that $\Phi = C_n$ and that $\pi(\mathcal{D})$ is a tent formed by two adjacent neutral clovers. Then $\mathcal{D}$ is on three vertices, and so by Case 1b, we find that N is a split diamond, once again.

By Cases 3a–b, statement (3) follows. The difference between Cases 3a and 3b are depicted in the first column of Figure 4.23.

Case 4. Finally, suppose that $\Phi = C_n$ and that $\pi(\mathcal{D})$ is a hanger. We will argue that this case corresponds to the second and third columns in Figure 4.23.

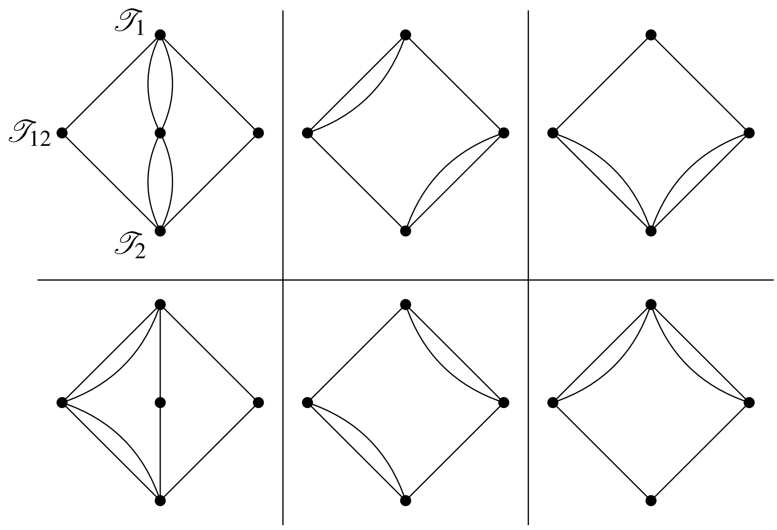

Figure 4.23: The additional (other than the single diamond) possible $N(\mathcal{T}_1, \mathcal{T}_2)$ in type C_n, when $\mathcal{G}_1, \mathcal{G}_2$ are adjacent. *From left to right:* Split, double and heavy diamonds. In each cell, the leftmost vertices are $\mathcal{T}_1$, $\mathcal{T}_{12}$ and $\mathcal{T}_2$, as in the first cell.

Note that, in this case, exactly one of $\mathcal{G}_1, \mathcal{G}_2$ is a neutral clover and the other is an adjacent neutral triangle. Suppose that the loop game ℓ in $\mathcal{D}$ is played by vertex

x and that the other two vertices in $\mathcal{D}$ are y, z. Note that any way of reversing $\mathcal{D}$ in two steps will involve reversing ℓ exactly once, and so each path of length two from $\mathcal{T}_1$ to $\mathcal{T}_2$ will contain exactly one double edge.

The four cases in the second and third columns in Figure 4.23 can be seen by considering the other games played between x, y and x, z that are not in $\mathcal{D}$. Depending on their outcomes, each such game either creates a clover with loop ℓ at x or else forms a neutral triangle together with the two "opposite" games in $\mathcal{D}$. After this clover/triangle is reversed, the triangle/clover, which was not initially, becomes present. See Figure 4.24.

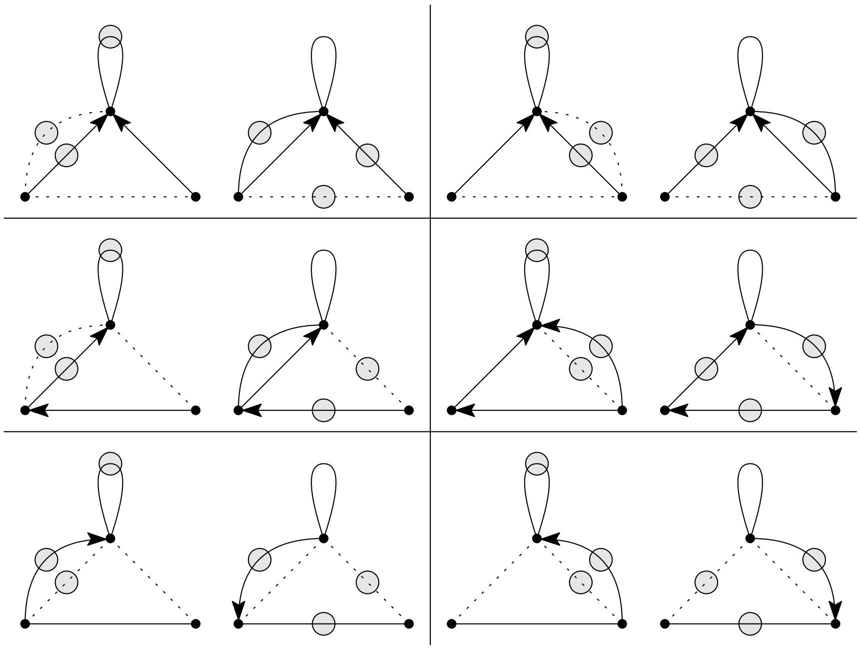

Figure 4.24: Reversing when $\pi(\mathcal{D})$ is a hanger in type C_n. Each row corresponds to a possible configuration for $\mathcal{D}$.

The proof is complete. □

4.5.2 Extended networks

Lemmas 4.5.2 and 4.5.4 above classify the types of interchange networks $N(\mathcal{T}_1, \mathcal{T}_2)$. Recall that such a network contains all paths of length two between $\mathcal{T}_1, \mathcal{T}_2$.

Definition 4.5.5. *We define the extended interchange network $\hat{N}(\mathcal{T}_1, \mathcal{T}_2)$ to be the union of $N(\mathcal{T}_1', \mathcal{T}_2')$ over all "antipodal" pairs $\mathcal{T}_1', \mathcal{T}_2'$ in $N(\mathcal{T}_1, \mathcal{T}_2)$ at distance two.*

Single, double and quadruple diamonds are "stable," in the sense that $\hat{N} = N$. In contrast, split and heavy diamond networks extend to a type of structure, which we call a *crystal*. See Figure 4.25.

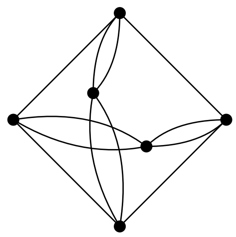

Figure 4.25: The crystal extended interchange network.

Remark 4.5.6. *All of the interchange networks that we have described, except the single diamond, can be found in Figure 4.3 above. This demonstrates how interchange networks can overlap and mesh together to form the interchange graph of a given score sequence.*

Lemma 4.5.7. *Suppose that $\mathcal{T}_1, \mathcal{T}_2$ are at distance two in* $\mathrm{IntGr}(\Phi, \mathbf{s})$*. Let $N = N(\mathcal{T}_1, \mathcal{T}_2)$ and $\hat{N} = \hat{N}(\mathcal{T}_1, \mathcal{T}_2)$.*

1. *If N is a single, double or quadruple diamond, then the extended interchange network $\hat{N} = N$.*

2. *Otherwise, if N is a split or heavy diamond, then the extended interchange network $\hat{N}$ is a crystal.*

Proof. Statement (1) is clear, and can be seen by inspection. On the other hand, statement (2) follows by repeated application of Lemma 4.5.4, considering the various antipodal pairs in N.

Case 1. If N is a split diamond, as in the first column of Figure 4.23, then consider $\mathcal{T}_1', \mathcal{T}_2'$ in N that are incident to only single edges in N. By Lemma 4.5.4, it follows that $N(\mathcal{T}_1', \mathcal{T}_2')$ is a split diamond, and therefore $\hat{N}$ is a crystal.

Case 2. If N is a heavy diamond, as in the third column of Figure 4.23, then consider $\mathcal{T}_1', \mathcal{T}_2'$ in N, each of which incident to exactly one single edge and one double edge in N. By Lemma 4.5.4, it follows that $N(\mathcal{T}_1', \mathcal{T}_2')$ is a split diamond, and therefore there is a path of length two between them consisting of two single edges in $\hat{N}$. Then, applying Lemma 4.5.4, once again, but this time to the midpoint $\mathcal{T}_{12}'$ along this path and $\mathcal{T}_3$ in N that is incident to two single edges in N, we find that $\hat{N}$ is a crystal, as claimed. □

Recall that two generators are either disjoint or have exactly one game in common. A similar property holds for extended networks.

Lemma 4.5.8. *Any two distinct extended networks $\hat{N} \neq \hat{N}'$ are either edge-disjoint or have exactly one single or double edge in common.*

Proof. By Lemma 4.5.7 there are only four types of extended networks. By an elementary case analysis, it can be seen that any two antipodal (distance two) vertices $\mathcal{T}_1, \mathcal{T}_2$ in an extended network $\hat{N}$ give rise to the same extended network $\hat{N}$. That is, $\hat{N} = \hat{N}(\mathcal{T}_1, \mathcal{T}_2)$, for any such $\mathcal{T}_1, \mathcal{T}_2$. From this observation the result follows, since if two extended networks $\hat{N}, \hat{N}'$ share at least three vertices, then they necessarily have at least one antipodal pair of vertices in common. ☐

4.5.3 Properties of crystals

In this section, we obtain two key properties of crystals, which will play a crucial role in the type C_n couplings discussed in Section 4.6.2 below.

First, we note that crystals cannot share a single edge. We will use this, together with Lemma 4.5.8, to extend natural couplings on networks to a coupling on the full interchange graph.

Lemma 4.5.9. *Suppose that $\hat{N} \neq \hat{N}'$ are distinct crystals in an interchange network* $\mathrm{IntGr}(C_n, \mathbf{s})$*. Then $\hat{N}$, $\hat{N}'$ are either edge-disjoint or share a double edge. That is, no such $\hat{N} \neq \hat{N}'$ share a single edge.*

Proof. By the proof of Lemmas 4.5.4 and 4.5.7, it can be seen that each crystal is associated with three players. Each double edge in the crystal corresponds to reversing a neutral clover involving two of them, and each single edge corresponds to reversing a neutral triangle involving all three.

By Lemma 4.5.8, it suffices to show that two crystals $\hat{N} \neq \hat{N}'$ cannot share a single edge. To see this, simply note that otherwise both tournaments joined by this single edge would contain a tournament on three players with three neutral triangles, which is impossible. See Figure 4.26. ☐

The previous result shows that single edges can be in at most one crystal. Double edges, on the other hand, can be in more than one. The following result gives an upper bound on this number, which is related to the re-weighting w of the graph metric, discussed in Section 4.6.2, under which our coupling will be contractive.

Recall that $d = d(C_n, \mathbf{s})$ is the degree of $\mathrm{IntGr}(C_n, \mathbf{s})$.

Lemma 4.5.10. *Any given double edge in an interchange network* $\mathrm{IntGr}(C_n, \mathbf{s})$ *is contained in at most* $\min\{d, 2n\}$ *crystals.*

Proof. Consider a double edge between some $\mathcal{T}_1, \mathcal{T}_2$. By Lemma 4.5.8, each crystal containing it corresponds to an additional double edge or two single edges incident to $\mathcal{T}_1$. It follows that there are at most $(d-2)/2$ such crystals.

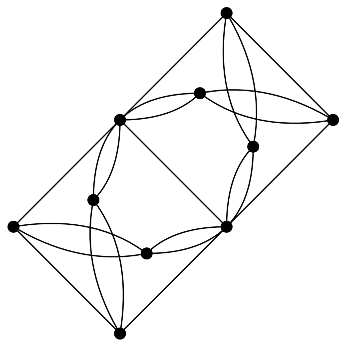

Figure 4.26: It is impossible for two (distinct) crystals to share a single edge, as depicted above, since there are at most two neutral triangles on any given three players in a tournament. However, each of the two "middle" vertices in this figure are incident to three single edges.

The second bound is somewhat more complicated. Recall, as noted in the proof of Lemma 4.5.9, that each crystal is associated with three players. Suppose that a double edge between some $\mathcal{T}_1, \mathcal{T}_2$ is associated with a neutral clover involving players i, j. We claim that, for any other player k, there are at most two crystals associated with i, j, k. To see this, observe that, if there were three, then one of $\mathcal{T}_1, \mathcal{T}_2$ would be incident to four single edges in these crystals. However, this would imply that in one of $\mathcal{T}_1, \mathcal{T}_2$ there are four neutral triangles on i, j, k, which is impossible. Indeed, as noted in the proof of Lemma 4.5.9, there can be at most two. Therefore, there are at most $2(n-2)$ crystals containing any given double edge. (In fact, the upper bound $n-2$ can be proved, but involves a more careful analysis.) □

4.6 Rapid mixing

Using the results of the previous section, we show that simple random walk on any given $\mathrm{IntGr}(\Phi, \mathbf{s})$ is rapidly mixing. The idea is to first define couplings on extended networks $\hat{N}$. We then argue that these couplings are compatible, and extend to a full coupling.

In types B_n and D_n, rapid mixing then follows by Theorem 4.2.3, using the standard weighting $w = \delta$ given by the graph distance δ in $\mathrm{IntGr}(\Phi, \mathbf{s})$. In type C_n, we will need to select a special re-weighting $w \neq \delta$, accounting for the presence of crystals in the interchange graphs of this type.

For a tournament $\mathcal{T} \in \mathrm{Tour}(\Phi, \mathbf{s})$, we let $\mathcal{E}(\mathcal{T})$ denote the set of edges in $\mathrm{IntGr}(\Phi, \mathbf{s})$ incident to $v(\mathcal{T})$.

4.6.1 Coupling in B_n and D_n

We begin with the simplest cases of types B_n and D_n. These types are the most straightforward, since then all networks $N(\mathcal{T}_1, \mathcal{T}_2)$ are single diamonds, and no re-weighting of the graph metric is necessary.

Theorem 4.6.1. *Let $\Phi = B_n$ or D_n. Fix any $\mathbf{s} \in \mathrm{Score}(\Phi)$. Then lazy simple random walk $(\mathcal{T}_n : n \geq 0)$ on the Coxeter interchange graph* $\mathrm{IntGr}(\Phi, \mathbf{s})$ *is rapidly mixing in time $t_{\mathrm{mix}} = O(d \log n)$.*

Proof. Let $\Phi = B_n$ or D_n. Consider two copies of lazy simple random walk $(\mathcal{T}_n')$ and $(\mathcal{T}_n'')$ on $\mathrm{IntGr}(\Phi, \mathbf{s})$, started from neighboring $\mathcal{T}_0', \mathcal{T}_0'' \in \mathrm{Tour}(\Phi, \mathbf{s})$. Then $\mathcal{T}_0'' = \mathcal{T}_0' * \mathcal{G}$ for some type Φ generator $\mathcal{G} \subset \mathcal{T}_0'$. In this sense, the random walks start at distance 1.

We will construct a contractive coupling of $\mathcal{T}_1', \mathcal{T}_1''$, such that the expected distance between $\mathcal{T}_1', \mathcal{T}_1''$ is strictly less than 1, for every choice of $\mathcal{T}_0', \mathcal{T}_0''$. In fact, in this coupling, the $\mathcal{T}_1', \mathcal{T}_1''$ will coincide with probability $1/d$, and otherwise remain at distance 1.

More specifically, we couple $\mathcal{T}_1', \mathcal{T}_1''$ using the natural bijection ψ from $\mathcal{E}(\mathcal{T}_0')$ to $\mathcal{E}(\mathcal{T}_0'')$, which fixes the edge $\{\mathcal{T}_0', \mathcal{T}_0''\}$ and pairs "opposite" edges in each single diamond containing $\{\mathcal{T}_0', \mathcal{T}_0''\}$. By Lemmas 4.5.2 and 4.5.4, for each type Φ generator $\mathcal{G}' \neq \mathcal{G} \subset \mathcal{T}_0'$, the network $N(\mathcal{T}_0'', \mathcal{T}_0' * \mathcal{G}')$ is a single diamond. As such, there are exactly two paths of length two from $\mathcal{T}_0''$ to $\mathcal{T}_0' * \mathcal{G}'$. One such path passes through $\mathcal{T}_0'$. We let

$$\psi(\{\mathcal{T}_0', \mathcal{T}_0' * \mathcal{G}'\}) = \{\mathcal{T}_0'', \mathcal{T}_0'' * \mathcal{G}''\} \tag{4.6.1}$$

be the first edge along the other such path. By Theorem 4.2.1 and Lemmas 4.5.7 and 4.5.8, ψ is a bijection from $\mathcal{E}(\mathcal{T}_0')$ to $\mathcal{E}(\mathcal{T}_0'')$.

Finally, we define the coupling of $\mathcal{T}_1', \mathcal{T}_1''$ as follows. Let $\{\mathcal{T}_0', \mathcal{T}_0' * \mathcal{G}'\}$ be a uniformly random edge in $\mathcal{E}(\mathcal{T}_0')$ and r_0 a Bernoulli(1/2) random variable (i.e., a fair "coin flip"). If $\mathcal{G}' = \mathcal{G}$, we put $\mathcal{T}_1' = \mathcal{T}_1'' = \mathcal{T}_0'$ if $r_0 = 0$ and $\mathcal{T}_1' = \mathcal{T}_1'' = \mathcal{T}_0''$ if $r_0 = 1$. In this case, the coupling contracts. On the other hand, if $\mathcal{G}' \neq \mathcal{G}$ we put $\mathcal{T}_1' = \mathcal{T}_0'$ and $\mathcal{T}_1'' = \mathcal{T}_0''$ if $r_0 = 0$, and $\mathcal{T}_1' = \mathcal{T}_0' * \mathcal{G}'$ and $\mathcal{T}_1'' = \mathcal{T}_0'' * \mathcal{G}''$ if $r_0 = 1$, where $\mathcal{G}''$ is given by the bijection ψ in (4.6.1). See Figure 4.27.

In this coupling, $\mathcal{T}_1' = \mathcal{T}_1''$ with probability $1/d$. Otherwise, they remain at distance 1. By Theorem 4.4.1, the diameter of $\mathrm{IntGr}(\Phi, \mathbf{s})$ is $D = O(n^2)$. Therefore, by Theorem 4.2.3, the mixing time is bounded by $O(d \log n)$. $\square$

Remark 4.6.2. *Rapid mixing for classical (type A_{n-1}) tournaments follows as a special case of the argument above.*

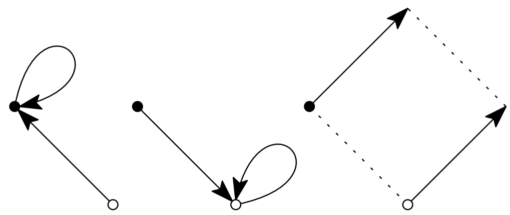

Figure 4.27: A contractive coupling in B_n and D_n. The black and white vertices represent the starting positions $\mathcal{T}_0', \mathcal{T}_0''$. In the configuration at right, corresponding to $\mathcal{G}' \neq \mathcal{G}$, the loops have been omitted, since in this case the walks are either both lazy or not.

4.6.2 Coupling in C_n

Finally, we investigate mixing in type C_n.

Recall that, in types B_n and D_n, the coupling was determined by an edge pairing, given by a bijection ψ from $\mathcal{E}(\mathcal{T}_0')$ to $\mathcal{E}(\mathcal{T}_0'')$, where $\mathcal{T}_0', \mathcal{T}_0''$ are neighboring tournaments in $\mathrm{IntGr}(\Phi, \mathbf{s})$. The coupling in type C_n is also determined by such a ψ, however, since there are a number of different networks in type C_n, the pairing is more involved.

The fact (see Lemma 4.5.9) that distinct crystals cannot share a single edge is crucial. Otherwise it would not be possible to extend couplings on extended networks to a full coupling. Roughly speaking, this is because (see Case 1b in the proof of Theorem 4.6.3 below) single edges $\{\mathcal{T}_0', \mathcal{T}_0''\}$ in a crystal will need to be paired with one of the edges in a double edge of the same crystal. As such, if there were two crystals with the same single edge, a bijective pairing would not be possible.

Furthermore, there is an additional complication in type C_n. As it turns out, the pairing ψ does not lead to a contractive coupling, with respect to the graph distance in $\mathrm{IntGr}(C_n, \mathbf{s})$. The problem concerns the case that the initial starting positions $\mathcal{T}_0', \mathcal{T}_0''$ are joined by a single edge in a crystal. In this case, the natural coupling is only "neutral" (i.e., with $\alpha = 0$ in Theorem 4.2.3), rather than contractive.

There are (at least) three ways to overcome this difficulty, leading to increasingly better bounds on the mixing time. The first way is to apply Bordewich and Dyer's [19] path coupling without contraction, leading to an upper bound $O(dn^4) = O(n^7)$. It is also possible to apply path coupling at time $t = 2$, since at this point the coupling (with respect to the usual graph metric) becomes contractive. In doing so, the key is to observe that, in the problematic case that $\mathcal{T}_0', \mathcal{T}_0''$ are joined by a single edge in a crystal, if $\mathcal{T}_1'$ stays within the crystal then the pairing ψ (described below) selects a $\mathcal{T}_1''$ in the crystal such that $\mathcal{T}_1', \mathcal{T}_1''$ are now joined by a double edge. This argument leads to an upper bound of $O(d^2 \log n) = O(n^6 \log n)$.

We will present a third strategy, by re-weighting the metric, which yields a better bound.

Recall (see Lemma 4.5.9) that single edges in $\mathrm{IntGr}(C_n, \mathbf{s})$ are contained in at most one crystal. Double edges, on the other hand, can be contained in more than one. We define the *crystal degree* of a double edge to be the number of crystals containing it. We let $\gamma = \gamma(C_n, \mathbf{s})$ denote the *maximal crystal degree,* over all double edges. By Lemma 4.5.10, we have that $\gamma \le \min\{d, 2n\}$, where $d = d(C_n, \mathbf{s})$ is the degree of $\mathrm{IntGr}(C_n, \mathbf{s})$.

We will assume throughout that $\gamma > 0$. Indeed, if $\gamma = 0$, then there are no crystals in the interchange graph. In this case, a straightforward modification of the proof of Theorem 4.6.1 (using the standard graph metric) shows that $t_{\mathrm{mix}} = O(d \log n)$.

We will prove the following result, using the weighting w that puts $w = 1$ on each edge in a double edge and $w = 1 + 1/\gamma$ on each single edge.

To be clear, this choice of w re-weights the graph distance between neighboring vertices joined by single edges, but not those joined by double edges. Specifically, if u, v are joined by a single edge then $w(u, v) = 1 + 1/\gamma$, and if u, v are joined by a double edge then each edge is given weight 1, and so the weighted distance between u, v remains $w(u, v) = 1$ (see Definition 4.2.2).

Theorem 4.6.3. *Let* $\Phi = C_n$*. Fix any* $\mathbf{s} \in \mathrm{Score}(\Phi)$*. Then lazy simple random walk* $(\mathcal{T}_n : n \ge 0)$ *on the Coxeter interchange graph* $\mathrm{IntGr}(C_n, \mathbf{s})$ *is rapidly mixing. If there are no crystals in* $\mathrm{IntGr}(C_n, \mathbf{s})$ *(when* $\gamma = 0$*) then* $t_{\mathrm{mix}} = O(d \log n)$*. Otherwise, we have* $t_{\mathrm{mix}} = O(\gamma d \log n)$*.*

In particular, this result implies that $t_{\mathrm{mix}} = O(n^4 \log n)$.

Proof. As discussed, let us assume that $\gamma > 0$, as otherwise a simple adaptation of the general reasoning in types B_n and D_n (the proof of Theorem 4.6.1) shows that $t_{\mathrm{mix}} = O(d \log n)$.

Consider two copies of lazy simple random walk $(\mathcal{T}'_n)$ and $(\mathcal{T}''_n)$ on $\mathrm{IntGr}(C_n, \mathbf{s})$, started from neighbours $\mathcal{T}'_0, \mathcal{T}''_0 \in \mathrm{Tour}(C_n, \mathbf{s})$. Let $\mathcal{G}$ be such that $\mathcal{T}''_0 = \mathcal{T}'_0 * \mathcal{G}$.

The first step is to obtain an edge pairing ψ, which will associate a $\mathcal{T}''_1$ to each possible $\mathcal{T}'_1$. Then we will show that this coupling is contractive, with respect to the re-weighting $w = 1 + 1/\gamma$ on single edges and $w = 1$ on each edge in double edges. The key in this regard will be the classification of extended networks, established in Section 4.5.

By Lemma 4.5.9, there are three cases to consider:

1. $\mathcal{G} = \Delta$ is a neutral triangle, and the edge $\{\mathcal{T}'_0, \mathcal{T}''_0\}$ is in

    (a) no crystal,

(b) exactly one crystal.

2. $\mathcal{G} = \Theta$ is a neutral clover, and the double edge between $\mathcal{T}_0', \mathcal{T}_0''$ is in $\gamma' \leq \gamma$ crystals.

In these cases, we will construct couplings with the following properties:

- In Case 1a, either $\mathcal{T}_1' = \mathcal{T}_1''$, or else $\mathcal{T}_1', \mathcal{T}_1''$ are again joined by a single edge.
- In Case 1b, either $\mathcal{T}_1', \mathcal{T}_1''$ are joined by a double edge in the crystal, or else $\mathcal{T}_1', \mathcal{T}_1''$ are joined by a single edge.
- In Case 2, either $\mathcal{T}_1', \mathcal{T}_1''$ are joined by a single edge in some crystal containing the double edge between $\mathcal{T}_0', \mathcal{T}_0''$, or else $\mathcal{T}_1', \mathcal{T}_1''$ are joined by a double edge.

Note that, under these couplings, crystal networks work like "switches," in that they move single edges to double edges, and vice versa. Also note that, it is Cases 1b and 2 in which the choice of w is crucial. We put weight $w = 1 + 1/\gamma$ on single edges so that, as we will see, the couplings in these cases are contractive.

Case 1a. Suppose that $\mathcal{G} = \Delta$ is a neutral triangle, and that the single edge $\{\mathcal{T}_0', \mathcal{T}_0''\}$ is not contained in a crystal. Then, by Lemmas 4.5.2 and 4.5.4, all extended networks $\hat{N}$ containing $\{\mathcal{T}_0', \mathcal{T}_0''\}$ are single and double diamonds. As such, it is only slightly more complicated to construct a contractive coupling in this case, than it was in types B_n and D_n above. We proceed as depicted in Figure 4.28 (cf. Figure 4.27).

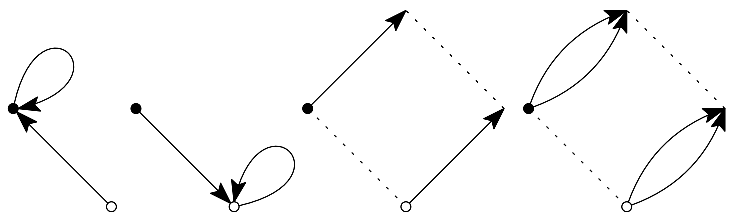

Figure 4.28: *Case 1a:* A contractive coupling, when the starting positions (black and white vertices) are joined by a single edge, which is not in a crystal.

Once again (as in the proof of Theorem 4.6.1), using Theorem 4.2.1 and Lemmas 4.5.7 and 4.5.8, we find a bijection ψ from from $\mathcal{E}(\mathcal{T}_0')$ to $\mathcal{E}(\mathcal{T}_0'')$ that fixes the edge $\{\mathcal{T}_0', \mathcal{T}_0''\}$ and pairs "opposite" edges in single and double diamonds containing $\{\mathcal{T}_0', \mathcal{T}_0''\}$. and so correspond to the same generator.

For each edge $\{\mathcal{T}_0', \mathcal{T}_0' * \mathcal{G}'\}$ in a single diamond $\hat{N}$ containing $\{\mathcal{T}_0', \mathcal{T}_0''\}$, we let

$$\psi(\{\mathcal{T}_0', \mathcal{T}_0' * \mathcal{G}'\}) = \{\mathcal{T}_0'', \mathcal{T}_0'' * \mathcal{G}''\} \tag{4.6.2}$$

be the "opposite" edge in $\hat{N}$.

Likewise, for each double edge, consisting of two copies $\{\mathcal{T}_0', \mathcal{T}_0' * \mathcal{G}'\}^{(i)}$, with $i \in \{1, 2\}$, of the same edge in a double diamond network $\hat{N}$ containing $\{\mathcal{T}_0', \mathcal{T}_0''\}$, we let

$$\psi(\{\mathcal{T}_0', \mathcal{T}_0' * \mathcal{G}'\}^{(i)}) = \{\mathcal{T}_0'', \mathcal{T}_0'' * \mathcal{G}''\}^{(i)}, \tag{4.6.3}$$

with $i \in \{1, 2\}$, be the "opposite" edges in $\hat{N}$.

We couple $\mathcal{T}_1', \mathcal{T}_1''$ as follows. Let $\{\mathcal{T}_0', \mathcal{T}_0' * \mathcal{G}'\}$ be a uniformly random edge in $\mathcal{E}(\mathcal{T}_0')$ and r_0 a Bernoulli(1/2). (Note that, this is a uniformly random edge, not generator. Indeed, neutral clovers $\mathcal{G}' = \Theta$ corresponding to double edges are twice as likely to be selected as neutral triangles $\mathcal{G}' = \Delta$.) If $\mathcal{G}' = \mathcal{G}$, we put $\mathcal{T}_1' = \mathcal{T}_1'' = \mathcal{T}_0'$ if $r_0 = 0$ and $\mathcal{T}_1' = \mathcal{T}_1'' = \mathcal{T}_0''$ if $r_0 = 1$. On the other hand, if $\mathcal{G}' \neq \mathcal{G}$, we put $\mathcal{T}_1' = \mathcal{T}_0'$ and $\mathcal{T}_1'' = \mathcal{T}_0''$ if $r_0 = 0$ and $\mathcal{T}_1' = \mathcal{T}_0' * \mathcal{G}'$ and $\mathcal{T}_1'' = \mathcal{T}_0'' * \mathcal{G}''$ if $r_0 = 1$, where $\mathcal{G}''$ is given by the bijection ψ, defined in (4.6.2) or (4.6.3) above.

Note that $\mathcal{T}_1' = \mathcal{T}_1''$ with probability $1/d$. Otherwise, $\mathcal{T}_1', \mathcal{T}_1''$ are again joined by a single edge. As such

$$\mathbf{E}\left[w(\mathcal{T}_1', \mathcal{T}_1'')\right] = (1 - 1/d)w(\mathcal{T}_0', \mathcal{T}_0''). \tag{4.6.4}$$

Case 1b. Suppose that $\mathcal{G} = \Delta$ is a neutral triangle, and that the single edge $\{\mathcal{T}_0', \mathcal{T}_0''\}$ is contained in *exactly* one crystal. Once again, by Lemma 4.5.8, all single and double diamonds and the one crystal containing $\{\mathcal{T}_0', \mathcal{T}_0''\}$ are otherwise edge-disjoint. In this case, the bijection can no longer fix $\{\mathcal{T}_0', \mathcal{T}_0''\}$, as in Case 1a. Rather, we will need to use this edge in a non-trivial way in order to define the edge pairing within the crystal.

The bijection ψ, in this case, is defined in the same way as in Case 1a for the edges in each single and double diamond. On the other hand, for the edges in the crystal, we define ψ as indicated in Figure 4.29. That is, the two single edges in the crystal incident to $\mathcal{T}_0'$ (one of which is $\{\mathcal{T}_0', \mathcal{T}_0''\}$) are paired with the double edges in the crystal incident to $\mathcal{T}_0''$, and vice versa. By Theorem 4.2.1, ψ is a bijection from $\mathcal{E}(\mathcal{T}_0')$ to $\mathcal{E}(\mathcal{T}_0'')$. We stress here that ψ pairs edges, not generators, and it is critical, in this case, that there is only one crystal containing $\{\mathcal{T}_0', \mathcal{T}_0''\}$ (since it can only be paired once).

To couple $\mathcal{T}_1', \mathcal{T}_1''$, we let $\{\mathcal{T}_0', \mathcal{T}_0' * \mathcal{G}'\}$ be a uniformly random edge in $\mathcal{E}(\mathcal{T}_0')$ and r_0 a Bernoulli(1/2). If $r_0 = 0$, we put $\mathcal{T}_1' = \mathcal{T}_0'$ and $\mathcal{T}_1'' = \mathcal{T}_0''$. If $r_0 = 1$, we put $\mathcal{T}_1' = \mathcal{T}_0' * \mathcal{G}'$ and $\mathcal{T}_1'' = \mathcal{T}_0'' * \mathcal{G}''$, where $\mathcal{G}''$ is given by the bijection ψ.

In this case, with probability $2/d$ the pair $\mathcal{T}_1', \mathcal{T}_1''$ remains in the crystal, but are now joined by a double edge. Otherwise, $\mathcal{T}_1', \mathcal{T}_1''$ are again joined by a single edge.

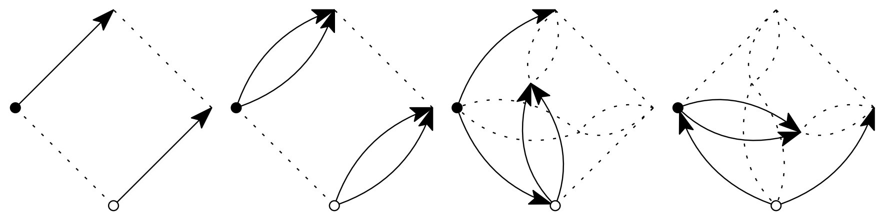

Figure 4.29: *Case 1b:* A neutral coupling, when the starting positions (black and white vertices) are joined by a single edge, which is in exactly one crystal. Note that, if the walks stay in the crystal, they move to a pair joined by a double edge.

Therefore,

$$\begin{aligned}\mathbf{E}\left[w(\mathcal{T}_1', \mathcal{T}_1'')\right] &= (1-2/d)(1+1/\gamma) + 2/d \\ &= \left[1 - \frac{2}{d(1+\gamma)}\right] w(\mathcal{T}_0', \mathcal{T}_0''), \qquad (4.6.5)\end{aligned}$$

since $w(\mathcal{T}_0', \mathcal{T}_0'') = 1 + 1/\gamma$.

Case 2. Finally, suppose that $\mathcal{G} = \Theta$ is a neutral clover. Suppose that $\{\mathcal{T}_0', \mathcal{T}_0''\}$ is contained in $\gamma' \leq \gamma$ crystals.

By Lemmas 4.5.2 and 4.5.4, all extended networks $\hat{N}$ containing $\{\mathcal{T}_0', \mathcal{T}_0''\}$ are double and quadruple diamonds and crystals. As in the previous cases, we define ψ in this case by pairing "opposite" edges in the double and quadruple diamonds. In this case, ψ fixes the two edges in the double edge between $\mathcal{T}_0', \mathcal{T}_0''$. Note that if $\{\mathcal{T}_0', \mathcal{T}_0''\}$ is in a crystal, then one of $\mathcal{T}_0', \mathcal{T}_0''$ is incident to two single edges in the crystal and the other is incident to a double edge $\neq \{\mathcal{T}_0', \mathcal{T}_0''\}$ in the crystal. We define ψ on each such crystal by pairing these edges, as indicated in Figure 4.30. Once again, applying Theorem 4.2.1, we see that ψ is a bijection from $\mathcal{E}(\mathcal{T}_0')$ to $\mathcal{E}(\mathcal{T}_0'')$.

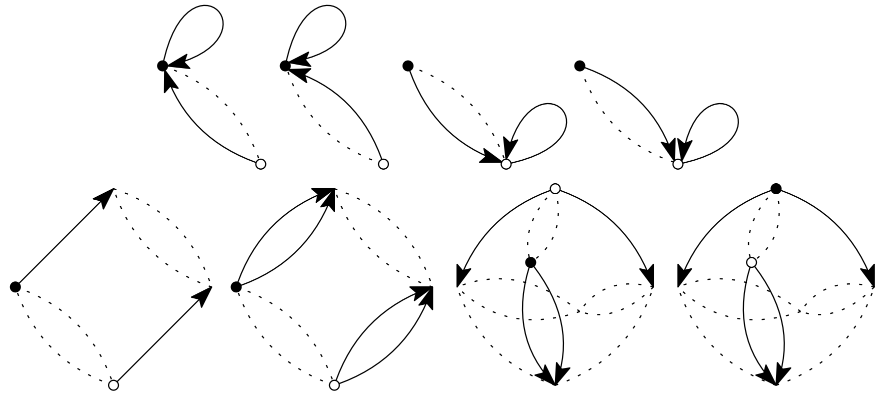

Figure 4.30: *Case 2:* A contractive coupling, when the starting positions (black and white vertices) are joined by a double edge.

To couple $\mathcal{T}_1', \mathcal{T}_1''$, we let $\{\mathcal{T}_0', \mathcal{T}_0' * \mathcal{G}'\}$ be a uniformly random edge in $\mathcal{E}(\mathcal{T}_0')$ and r_0 a Bernoulli(1/2). If $\mathcal{G}' = \mathcal{G}$, we put $\mathcal{T}_1' = \mathcal{T}_1'' = \mathcal{T}_0'$ if $r_0 = 0$ and $\mathcal{T}_1' = \mathcal{T}_1'' = \mathcal{T}_0''$ if

$r_0 = 1$. Otherwise, if $\mathcal{G}' \neq \mathcal{G}$, we put $\mathcal{T}_1' = \mathcal{T}_0'$ and $\mathcal{T}_1'' = \mathcal{T}_0''$ if $r_0 = 0$ and $\mathcal{T}_1' = \mathcal{T}_0' * \mathcal{G}'$ and $\mathcal{T}_1'' = \mathcal{T}_0'' * \mathcal{G}''$ if $r_0 = 1$, where $\mathcal{G}''$ is given by the bijection ψ.

In this case, $\mathcal{T}_1' = \mathcal{T}_1''$ with probability $2/d$. With probability γ'/d, the pair $\mathcal{T}_1', \mathcal{T}_1''$ move within one of the γ' crystals containing $\{\mathcal{T}_0', \mathcal{T}_0''\}$, and are then joined by a single edge. Otherwise, with probability $1 - (2 + \gamma')/d$, $\mathcal{T}_1', \mathcal{T}_1''$ remain joined by a double edge. Therefore,

$$\begin{aligned}
\mathbf{E}\left[w(\mathcal{T}_1', \mathcal{T}_1'')\right] &= \frac{\gamma'}{d}(1 + 1/\gamma) + 1 - \frac{2 + \gamma'}{d} \\
&= 1 - \frac{2\gamma - \gamma'}{d\gamma} \\
&\leq (1 - 1/d)w(\mathcal{T}_0', \mathcal{T}_0''),
\end{aligned} \tag{4.6.6}$$

since $w(\mathcal{T}_0', \mathcal{T}_0'') = 1$.

By (4.6.4)–(4.6.6), we may apply Theorem 4.2.3 with $\alpha = O(1/\gamma d)$. Note that, by Theorem 4.4.1, it follows that $D_w = O((1 + 1/\gamma)D) = O(n^2)$. We conclude that the mixing time is bounded by $O(\gamma d \log n)$, as claimed. □

Bibliography

[1] Louigi Addario-Berry and Marie Albenque. The scaling limit of random simple triangulations and random simple quadrangulations. *The Annals of Probability*, 45(5):2767–2825, 2017. 14, 15, 47

[2] Louigi Addario-Berry and Marie Albenque. Convergence of non-bipartite maps via symmetrization of labeled trees. *Annales Henri Lebesgue*, 4:653–683, 2021. 14, 47, 48

[3] Louigi Addario-Berry, Nicolas Broutin, and Christina Goldschmidt. The continuum limit of critical random graphs. *Probability Theory and Related Fields*, 152(3):367–406, 2012. 2

[4] Louigi Addario-Berry, Serte Donderwinkel, Christina Goldschmidt, and Rivka Mitchell. Discrete snakes with globally centered displacements. *arXiv preprint arXiv:2505.21823*, 2025. 40

[5] Louigi Addario-Berry, Serte Donderwinkel, and Igor Kortchemski. Critical trees are neither too short nor too fat. *Annales Henri Lebesgue*, 8:113–149, 2025. 120

[6] David Aldous. The Random Walk Construction of Uniform Spanning Trees and Uniform Labelled Trees. *SIAM Journal on Discrete Mathematics*, 3(4):450–465, 1990. 8, 9

[7] David Aldous. The Continuum Random Tree. I. *The Annals of Probability*, 19(1):1–28, 1991. 2, 7

[8] David. Aldous. *The Continuum Random Tree II: An Overview*, pages 23–70. London Mathematical Society Lecture Note Series. Cambridge University Press, 1991.

[9] David Aldous. The Continuum Random Tree. III. *The Annals of Probability*, 21(1):248–289, 1993. 2, 7, 8, 54, 55, 67

[10] David Aldous and James Allen Fill. Reversible Markov Chains and Random Walks on Graphs, 2002. Unfinished monograph, recompiled 2014, available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`. 28, 167

[11] David Aldous and Jim Pitman. Invariance Principles for Non-Uniform Random Mappings and Trees. In *Asymptotic Combinatorics with Application to Mathematical Physics*, pages 113–147. Springer, Dordrecht, 2002. 55

[12] Omer Angel, Asaf Ferber, Benny Sudakov, and Vincent Tassion. Long Monotone Trails in Random Edge-Labellings of Random Graphs. *Combinatorics, Probability and Computing*, 29(1):22–30, 2020. 22, 138

[13] Federico Ardila, Federico Castillo, Christopher Eur, and Alexander Postnikov. Coxeter submodular functions and deformations of Coxeter permutahedra. *Advances in Mathematics*, 365:107039, 36, 2020. 157, 160, 163, 164

[14] Caelan Atamanchuk, Luc Devroye, and Gábor Lugosi. On the size of temporal cliques in subcritical random temporal graphs. *Combinatorics, Probability and Computing*, page 1–9, 2025. 21, 22, 139

[15] Spencer Backman. Riemann–Roch theory for graph orientations. *Advances in Mathematics*, 309:655–691, 2017. 37

[16] Spencer Backman. Partial graph orientations and the Tutte polynomial. *Advances in Applied Mathematics*, 94:103–119, 2018. Special issue on the Tutte polynomial. 37

[17] Ruben Becker, Arnaud Casteigts, Pierluigi Crescenzi, Bojana Kodric, Malte Renken, Michael Raskin, and Viktor Zamaraev. Giant Components in Random Temporal Graphs. In Nicole Megow and Adam Smith, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2023)*, volume 275 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 29:1–29:17, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 21, 24, 138, 139, 140

[18] Patrick Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, second edition, 1999. 54, 113

[19] Magnus Bordewich and Martin Dyer. Path coupling without contraction. *Journal of Discrete Algorithms*, 5(2):280–292, 2007. 188

[20] Anna Brandenberger, Serte Donderwinkel, Céline Kerriou, Gábor Lugosi, and Rivka Mitchell. Temporal connectivity of Random Geometric Graphs. *arXiv preprint arXiv:2502.15274*, 2025. 137

[21] Andrei Z Broder. Generating random spanning trees. In *FOCS*, volume 89, pages 442–447, 1989. 8

[22] Nicolas Broutin, Nina Kamčev, and Gabor Lugosi. Increasing paths in random temporal graphs. *The Annals of Applied Probability*, 34:5498–5521, 2024. 21, 22, 139

[23] Nicolas Broutin and Jean-François Marckert. Asymptotics of trees with a prescribed degree sequence and applications. *Random Structures & Algorithms*, 44(3):290–316, 2014. 2, 59, 92

[24] Richard A Brualdi and Qiao Li. The interchange graph of tournaments with the same score vector. In *Progress in graph theory (Waterloo, Ont., 1982)*, pages 129–151. Academic Press, Toronto, ON, 1984. 28, 34, 37, 158, 160

[25] Russ Bubley and Martin Dyer. Path coupling: A technique for proving rapid mixing in Markov chains. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 223–231. IEEE, 1990. 25, 28, 160, 167, 168

[26] Matija Bucić, Matthew Kwan, Alexey Pokrovskiy, Benny Sudakov, Tuan Tran, and Adam Zsolt Wagner. Nearly-linear monotone paths in edge-ordered graphs. *Israel Journal of Mathematics*, 238(2):663–685, 2020. 138

[27] Matthew Buckland, Brett Kolesnik, Rivka Mitchell, and Tomasz Przybyłowski. Random walks on Coxeter interchange graphs. *Electronic Journal of Probability*, 30(none):1 – 31, 2025. 36, 157

[28] A Robert Calderbank, Fan RK Chung, and Dean G Sturtevant. Increasing sequences with nonzero block sums and increasing paths in edge-ordered graphs. *Discrete mathematics*, 50:15–28, 1984. 138

[29] Élie Cartan. Sur la Réduction à sa Forme Canonique de la Structure d'un Groupe de Transformations Fini et Continu. *American Journal of Mathematics*, 18(1):1–61, 1896. 160

[30] Arnaud Casteigts, Michael Raskin, Malte Renken, and Viktor Zamaraev. Sharp thresholds in random simple temporal graphs. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 319–326, 2022. 21, 138

[31] Arnaud Casteigts, Michael Raskin, Malte Renken, and Viktor Zamaraev. Sharp thresholds in random simple temporal graphs. *SIAM Journal on Computing*, 53(2):346–388, 2024. 22

[32] Philippe Chassaing and Gilles Schaeffer. Random planar lattices and integrated superBrownian excursion. *Probability Theory and Related Fields*, 128(2):161–212, 2004. 14, 47

[33] An-Hang Chen, Jou-Ming. Chang, and Yue-Li Wang. The interchange graphs of tournaments with minimum score vectors are exactly hypercubes. *Graphs and Combinatorics*, 25(1):27–34, 2009. 36, 38, 158

[34] Václav Chvátal and J Komlós. Some combinatorial theorems on monotonicity. *Canadian Mathematical Bulletin*, 14(2):151–157, 1971. 138

[35] Guillaume Conchon–Kerjan and Christina Goldschmidt. The stable graph: The metric space scaling limit of a critical random graph with i.i.d. power-law degrees. *The Annals of Probability*, 51(1):1–69, 2023. 2

[36] Nicolas Curien and Igor Kortchemski. Random stable looptrees. *Electronic Journal of Probability*, 19(108):1–35, 2014. 51

[37] Serte Donderwinkel and Zheneng Xie. Universality for the directed configuration model: Metric space convergence of the strongly connected components at criticality. *Electronic Journal of Probability*, 29:1–85, 2024. 2

[38] Thomas Duquense and Fael Rebei. Scaling limits of critical branching random walks via their snakes. In preparation, 2025. 20, 49

[39] Thomas Duquesne. A limit theorem for the contour process of conditioned Galton–Watson trees. *The Annals of Probability*, 31(2):996–1027, 2003. 2, 7, 9, 10, 49

[40] Thomas Duquesne and Jean-François Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, 281:vi+147, 2002. 10, 43

[41] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publication of the Mathematical Institution of Hungarian Academy of Sciences*, 5:17–61, 1960. 22

[42] Jean François Le Gall. Random trees and applications. *Probability Surveys*, 2:245–311, 2005. 4, 6, 12, 55

[43] Pavel Galashin, Sam Hopkins, Thomas McConville, and Alexander Postnikov. Root system chip-firing I: interval-firing. *Mathematische Zeitschrift*, 292(3):1337–1385, 2019. 157

[44] Pavel Galashin, Sam Hopkins, Thomas McConville, and Alexander Postnikov. Root system chip-firing II: Central-firing. *International Mathematics Research Notices*, 2021(13):10037–10072, 2021. 157

[45] Emeric Gioan. Enumerating degree sequences in digraphs and a cycle–cocycle reversing system. *European Journal of Combinatorics*, 28(4):1351–1366, 2007. 37

[46] Bernhard Gittenberger. A Note on "State Spaces of the Snake and Its Tour—Convergence of the Discrete Snake" by J.-F. Marckert and A. Mokkadem. *Journal of Theoretical Probability*, 16:1063–1067, 2003. 47

[47] Christina Goldschmidt. Scaling limits of random trees and random graphs. In *PIMS-CRM Summer School in Probability*, pages 1–33. Springer, 2017. 8

[48] Christina Goldschmidt and Bénédicte Haas. A line-breaking construction of stable trees. *Electronic Journal of Probability*, 20:1–24, 2015. 20

[49] Ronald L Graham and Daniel J Kleitman. Increasing paths in edge ordered graphs. *Periodica Mathematica Hungarica*, 3(1-2):141–148, 1973. 138

[50] Bénédicte Haas and Grégory Miermont. Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees. *The Annals of Probability*, 40(6):2589–2666, 2012. 55, 94

[51] Seifollah Louis Hakimi. On realizability of a set of integers as degrees of the vertices of a linear graph. I. *Journal of the Society for Industrial and Applied Mathematics.*, 10:496–506, 1962. 27

[52] Brian Hall. *Lie Groups, Lie Algebras, and Representations*, volume 222 of *Graduate Texts in Mathematics.* Springer, Cham, second edition, 2015. An Elementary Introduction. 164

[53] Theodore Edward Harris. *The theory of branching processes*, volume 1. Springer Berlin, 1963. 6

[54] Václav Havel. A remark on the existence of finite graphs. *Časopis pro pěstování matematiky*, 80:477–480, 1955. 27

[55] James Humphreys. *Reflection groups and Coxeter groups*, volume 29 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990. 160, 163, 164

[56] Mikhail Isaev, Tejas Iyer, and Brendan D McKay. Asymptotic enumeration of orientations of a graph as a function of the out-degree sequence. *Electronic Journal of Combinatorics*, 27(1):Paper No. 1.26, 30, 2020. 27, 37, 158

[57] Svante Janson. Left and right pathlengths in random binary trees. *Algorithmica. An International Journal in Computer Science*, 46(3-4):419–429, 2006. 51

[58] Svante Janson. Probability asymptotics: notes on notation. Preprint `arXiv:1108.3924`, 2011. 56

[59] Svante Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probability Surveys*, 9:103–252, 2012. 85, 88

[60] Svante Janson and Jean-François Marckert. Convergence of discrete snakes. *Journal of Theoretical Probability*, 18(3):615–647, 2005. 15, 16, 18, 46, 47, 48, 100

[61] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000. 149, 154

[62] WD Kaigh. A conditional local limit theorem for recurrent random walk. *The Annals of Probability*, 3(5):883–888, 1975. 6

[63] Olav Kallenberg. *Foundations of modern probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer, Cham, third edition, 2021. 73, 131

[64] Ravi Kannan, Prasad Tetali, and Santosh Vempala. Simple Markov-chain algorithms for generating bipartite graphs and tournaments. *Random Structures & Algorithms*, 14(4):293–308, 1999. 25, 27, 158, 160, 163, 169

[65] Robin Khanfir. Convergences of looptrees coded by excursions. Preprint `arXiv:2208.11528`, *Ann. Inst. Henri Poincaré Probab. Stat.* to appear., 2022+. 54

[66] Wilhelm Killing. Die Zusammensetzung der stetigen endlichen Transformationsgruppen. *Mathematische Annalen*, 33:1–48, 1888. 160

[67] B. Kolesnik and M. Sanchez. The geometry of random tournaments. *Discrete & Computational Geometry*, 71:1343–1351, 2024. 157, 158

[68] Brett Kolesnik, Rivka Mitchell, and Tomasz Przybyłowski. Tournaments on signed graphs. *Annals of Combinatorics*, 2025. 33, 34, 35, 37, 38, 158, 161, 163, 165, 166, 167

[69] Brett Kolesnik and Mario Sanchez. Coxeter tournaments. Preprint available at arXiv:2302.14002. 30, 157, 158, 161, 163, 165, 169

[70] Igor Kortchemski. A simple proof of Duquesne's theorem on contour processes of conditioned Galton–Watson trees. *Séminaire de probabilités xlv*, pages 537–558, 2013. 2

[71] Igor Kortchemski and Cyril Marzouk. Random Lévy Looptrees and Lévy maps. arXiv:2402.04098, February 2024. 54

[72] Igor Kortchemski and Loïc Richier. The boundary of random planar maps via looptrees. *Annales de la Faculté des Sciences de Toulouse. Mathématiques. Série 6*, 29(2):391–430, 2020. 54

[73] H. G. Landau. On dominance relations and the structure of animal societies. III. The condition for a score structure. *The Bulletin of Mathematical Biophysics*, 15:143–148, 1953. 25, 27, 33, 159, 165

[74] Mikhail Lavrov and Po-Shen Loh. Increasing Hamiltonian paths in random edge orderings. *Random Structures & Algorithms*, 48(3):588–611, 2016. 138, 153

[75] Jean-François Le Gall. A class of path-valued Markov processes and its applications to superprocesses. *Probability Theory and Related Fields*, 95(1):25–46, 1993. 12, 41

[76] Jean-François Le Gall. The Brownian snake and solutions of $\Delta u = u^2$ in a domain. *Probability Theory and Related Fields*, 102(3):393–432, 1995. 12, 41

[77] Jean-François Le Gall. Uniqueness and universality of the Brownian map. *The Annals of Probability*, 41(4):2880–2960, 2013. 8, 14, 47

[78] Tao Lei. Scaling limit of random forests with prescribed degree sequences. *Bernoulli. Official Journal of the Bernoulli Society for Mathematical Statistics and Probability*, 25(4A):2409–2438, 2019. 2, 92

[79] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and Mixing Times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson. 28, 167

[80] Jean-François Marckert. The rotation correspondence is asymptotically a dilatation. *Random Structures & Algorithms*, 24(2):118–132, 2004. 51

[81] Jean-François Marckert and Grégory Miermont. Invariance principles for random bipartite planar maps. *The Annals of Probability*, 35(5):1642–1705, 2007. 14, 47, 48

[82] Jean-François Marckert and Abdelkader Mokkadem. The depth first processes of Galton-Watson trees converge to the same Brownian excursion. *The Annals of Probability*, 31(3):1655–1678, 2003. 7, 43, 49, 50

[83] Jean-François Marckert and Abdelkader Mokkadem. States Spaces of the Snake and Its Tour—Convergence of the Discrete Snake. *Journal of Theoretical Probability*, 16(4):1015–1046, 2003. 13, 14, 43, 47

[84] Jean-François Marckert. The lineage process in Galton Watson trees and globally centered discrete snakes. *The Annals of Applied Probability*, 18(1):209–244, 2008. 16, 17, 48, 49

[85] Anders Martinsson. Most edge-orderings of K_n have maximal altitude. *Random Structures & Algorithms*, 54(3):559–585, 2019. 138

[86] Cyril Marzouk. Scaling limits of discrete snakes with stable branching. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 56(1):502–523, 2020. 15, 19, 20, 48

[87] Brendan D. McKay. The asymptotic numbers of regular tournaments, Eulerian digraphs and Eulerian oriented graphs. *Combinatorica. An International Journal on Combinatorics and the Theory of Computing*, 10(4):367–377, 1990. 27, 37, 158

[88] Brendan D. McKay and Xiaoji Wang. Asymptotic enumeration of tournaments with a given score sequence. *Journal of Combinatorial Theory. Series A*, 73(1):77–90, 1996. 27, 37, 158

[89] Lisa McShine. Random sampling of labeled tournaments. *Electronic Journal of Combinatorics*, 7:Research Paper 8, 9, 2000. 25, 26, 27, 28, 29, 30, 35, 158, 160, 161, 163, 169

[90] Grégory Miermont. The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Mathematica*, 210(2):319–401, 2013. 14, 47

[91] John W. Moon. An extension of Landau's theorem on tournaments. *Pacific Journal of Mathematics*, 13:1343–1345, 1963. 165

[92] John W. Moon. *Topics on tournaments*. Holt, Rinehart and Winston, New York-Montreal, Que.-London, 1968. 167

[93] Mathew Penrose. Connectivity of soft random geometric graphs. *The Annals of Applied Probability*, 26, 11 2013. 23, 24, 140

[94] Valentin V. Petrov. *Sums of independent random variables*, volume Band 82 of *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]*. Springer–Verlag, New York-Heidelberg, 1975. Translated from the Russian by A. A. Brown. 56, 119, 120

[95] Jim Pitman. Poisson-Kingman partitions. In *Statistics and science: a Festschrift for Terry Speed*, volume 40 of *IMS Lecture Notes Monogr. Ser.*, pages 1–34. Inst. Math. Statist., Beachwood, OH, 2003. 93

[96] Jim Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer–Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. 56

[97] Alexander Postnikov. Permutohedra, associahedra, and beyond. *International Mathematics Research Notices. IMRN*, 2009(6):1026–1106, 2009. 159

[98] Richard Rado. An inequality. *Journal of the London Mathematical Society. Second Series*, 27:1–6, 1952. 159

[99] Abydr Sherry Sarkar. Mixing times of tournaments. Undergraduate thesis, Georgia Institute of Technology, 2020. 36, 160

[100] Alistair Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1(4):351–370, 1992. 158

[101] J. H. Spencer. Random regular tournaments. *Periodica Mathematica Hungarica. Journal of the János Bolyai Mathematical Society*, 5:105–120, 1974. 27, 37, 158

[102] Richard P. Stanley. Decompositions of rational convex polytopes. *Annals of Discrete Mathematics*, 6:333–342, 1980. 157, 159

[103] Johan Frederik Steffensen. Deux problèmes du calcul des probabilités. In *Annales de l'Institut Henri Poincaré*, volume 3, pages 319–344, 1933. 6

[104] John Tang, Ilias Leontiadis, Salvatore Scellato, Vincenzo Nicosia, Cecilia Mascolo, Mirco Musolesi, and Vito Latora. Applications of temporal graph metrics to real-world networks. *Temporal Networks*, pages 135–159, 2013. 138

[105] David Bruce Wilson. Mixing times of Lozenge tiling and card shuffling Markov chains. *The Annals of Applied Probability*, 14(1):274–325, 2004. 169

[106] Thomas Zaslavsky. The geometry of root systems and signed graphs. *American Mathematical Monthly*, 88(2):88–105, 1981. 31, 161, 163

[107] Thomas Zaslavsky. Signed graphs. *Discrete Applied Mathematics. The Journal of Combinatorial Algorithms, Informatics and Computational Sciences*, 4(1):47–74, 1982.

[108] Thomas Zaslavsky. Orientation of signed graphs. *European Journal of Combinatorics*, 12(4):361–375, 1991. 31, 32, 157, 161, 163, 169

[109] Günter M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer–Verlag, New York, 1995. 37, 157, 159