LIMITS OF MALLOWS TREES

BENOÎT CORSINI

ABSTRACT. This article studies the limit of binary search trees drawn from Mallows permutations under various topologies. The main result, pertaining to the standard local topology for graphs, requires the introduction of a generalization of binary search trees to two-sided infinite sequences, referred to as *redwood trees*. We then show that the almost-sure local limit of finite Mallows trees is the redwood tree drawn from the two-sided infinite Mallows permutation, thus corresponding to swapping the local limit and the binary search tree structure.

Building off this result, we study various other natural topologies: the rooted topology of the local structure around the root, the Gromov-Hausdorff-Prokhorov topology of the tree seen as a metric space, and the subtree size topology of the ratio of nodes split between left and right subtrees. The limit of Mallows trees under these three topologies combined with the case of the local topology allow us to draw a global picture of what large Mallows trees look like from different perspective and further strengthen the relation between finite and infinite Mallows permutations.

CONTENTS

1. INTRODUCTION

Mallows permutations [Mal57] have recently started to attract more attention as a natural model to generalize uniform permutations. In particular, the last decade has seen an increasing number of articles studying structural properties of Mallows permutations, such as the length of the longest monotone subsequence [BB17, BP15, MS13], the number of cycles [GP18, HMV22, Muk16], or the occurrence of patterns [CDE18, He22, Pin21].

A notable property of Mallows permutations is their strong regenerative structure [PT19] which allowed their definition to be extended to more complex sets than $[n] := \{1, \ldots, n\}$, such as the one-sided infinite Mallows permutation on $\mathbb{N} = \{1, 2, \ldots\}$ [GO10], or the two-sided infinite Mallows permutation on $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$ [GO12]. In this work, we further highlight the relation between Mallows permutations on $[n]$, $\mathbb{N}$, and $\mathbb{Z}$ by studying the limit under various topologies of *Mallows trees*, defined to be binary search trees drawn from Mallows permutations.

The bulk of this work focuses on the topology of the local convergence of graphs, as introduced by Benjamini and Schramm [BS01]. In that case, we show that the limit of a finite Mallows tree is almost-surely an adequately defined binary search tree on a two-sided infinite Mallows permutation (Theorem 1.1). This result is interesting to compare with the rooted local topology, where the limit of a finite Mallows tree is the binary search tree drawn from the infinite one-sided Mallows distribution

---

(Theorem 1.3). Finally, two other natural topologies on trees are considered: the Gromov-Hausforff-Prokhorov topology, where trees are seen as metric spaces, for which the limit is a simple metric on $[0, 1]$ (Theorem 1.4), and the recently introduced subtree size topology [Grü23], for which the limit is the infinite rightward path (Theorem 1.5).

The main result of this work (Theorem 1.1) provides an interesting limit swap between binary search tree structures and local limits of discrete objects. Thus it should not fully come as a surprise: the local structure of a finite Mallows tree seen from a uniformly chosen node only depends on the ordering of the surrounding values of a uniformly chosen entry in the permutation and behaves similarly to the tree corresponding to a two-sided infinite Mallows permutation seen from any entry. Although swapping the two limits seems natural here, it is worth noting that a key ingredient in the proof relies on the relation between displacements in Mallows permutations and geometric random variables, and the fact that the sum of two geometric random variables is similar to the size-biased geometric distribution (see Section 3.3 and more precisely (14) for more details on this regard).

Apart from the main results on the convergence of finite Mallows trees under various topologies, this article provides two noteworthy novelties. First, it extends the notion of binary search tree from one-sided infinite sequences to two-sided infinite sequences, under some mild assumptions. This extension of binary search trees is referred to as *redwood trees* (see Section 2.2.1 for a discussion on the name and notations for redwood trees) and is defined and studied in Section 2.2. Second, the main result of this article proves the almost-sure local convergence of a sequence of trees, joining the rather small family of almost-sure local convergence results [DM10]. Moreover, while proving almost-sure local convergence often relies on applying the Borel-Cantelli lemma and thus does not directly require a properly defined common space, Mallows trees are, to the best of our knowledge, the first sequence of trees with a recursive construction for which such result is known.

1.1. **Main results.** For any $n \in \mathbb{N}$ and $q \in [0, \infty)$, the Mallows distribution [Mal57] is the distribution $\pi_{n,q}$ on the symmetric group $\mathcal{S}_n$ defined by

$$\pi_{n,q}(\sigma) := \frac{q^{\text{Inv}(\sigma)}}{Z_{n,q}},$$

where $\text{Inv}(\sigma) := |\{i < j : \sigma(i) > \sigma(j)\}|$ is the number of inversions of σ and $Z_{n,q} := \sum_{\sigma \in \mathcal{S}_n} q^{\text{Inv}(\sigma)}$ is a normalizing constant. If $\sigma = (\sigma(1), \ldots, \sigma(n)) \in \mathcal{S}_n$ is a random permutation distributed accorded to $\pi_{n,q}$, then the permutation $(\sigma(n), \ldots, \sigma(1)) = \sigma \cdot (n, n-1, \ldots, 1)$ is distributed according to $\pi_{n,1/q}$. This provides a natural symmetry between Mallows permutations with parameters q and $1/q$ and allows us to reduce the study of such permutations to $q \in [0, 1]$. Furthermore, the case $q = 1$ corresponds to uniformly sampled permutations and thus tends to have very different properties from the case $q \neq 1$. For these two reasons, we limit our study to the case $q \in [0, 1)$ for the rest of the article (except in Section 5.1 where we briefly describe how to adapt our results to the case $q > 1$).

The Mallows distribution happens to be the unique distribution on $\mathcal{S}_n$ which is q-invariant, that is such that

$$\pi_{n,q}\big(\sigma \cdot (i\ i+1)\big) = q^{\text{sgn}[\sigma(i+1) - \sigma(i)]} \cdot \pi_{n,q}(\sigma),$$

where $(i\ i+1)$ is the transposition sending i to $i+1$ and vice-versa, and $\text{sgn}[x]$ is the sign of x. In fact, the q-invariant property happens to be a natural way to generalize Mallows permutations to one-sided infinite sets such as $\mathbb{N}$ [GO10] and two-sided infinite sets such as $\mathbb{Z}$ [GO12]. We further describe the construction of one-sided and two-sided infinite Mallows permutations in Section 2.1.

Let $\mathbb{T} = \{\varnothing\} \cup \bigcup_{k \geq 1} \{\mathbf{0}, \mathbf{1}\}^k$ be the infinite binary tree seen as words on the alphabet $\{\mathbf{0}, \mathbf{1}\}$, where $\varnothing$ corresponds equivalently to the root of the tree and the empty word. We further identify binary trees with subsets of $\mathbb{T}$ rooted at $\varnothing$ and connected in the graph sense, that is subsets $T \subseteq \mathbb{T}$ stable by considering any prefix of a word.

For any finite sequence of distinct values $x = (x_1, \ldots, x_n)$, define $T\langle x \rangle$ the binary search tree of x as follows. If $n = 0$ and x is empty, then $T\langle x \rangle = \emptyset$. Otherwise, if $n \geq 1$, write $x_- = (x_i : x_i < x_1)$

(respectively $x_+ = (x_i : x_i > x_1)$) for the sequence of numbers smaller (respectively larger) than x_1 in the same order as in x and let

$$T\langle x\rangle := \{\varnothing\} \cup \mathbf{0}T\langle x_-\rangle \cup \mathbf{1}T\langle x_+\rangle \; ; \tag{1}$$

see Figure 1 for an example of the construction of $T\langle x\rangle$. This definition straightforwardly implies that $|T\langle x\rangle| = n$. It is also worth noting that, if $x = (x_1, \ldots, x_n)$, then for any sequence $\bar{x} = (x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+m})$ extending x, we have $T\langle x\rangle \subseteq T\langle\bar{x}\rangle$. Using this remark, we naturally extend the previous notations to infinite sequences $x = (x_1, x_2, \ldots)$ by letting

$$T\langle x\rangle := \bigcup_{n=1}^{\infty} T\langle(x_1, \ldots, x_n)\rangle \; .$$

Since $T\langle(x_1, \ldots, x_n)\rangle \subseteq T\langle(x_1, \ldots, x_n, x_{n+1})\rangle$, the previous definition can actually be understood as $T\langle x\rangle$ being the limit of $(T\langle(x_1, \ldots, x_n)\rangle)_{n\in\mathbb{N}}$.
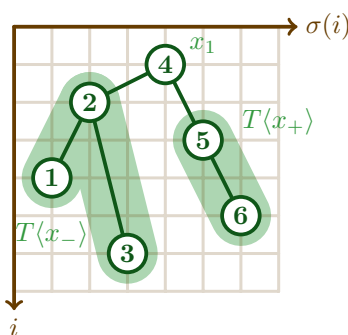

FIGURE 1. A representation of the binary search tree $T\langle x\rangle$ for the sequence $x = (4, 2, 5, 1, 6, 3)$. The tree is represented in dark green and the numbers on the node are the values of the corresponding x_i. In light green we show the construction of the binary search tree as defined in (1): x_1 (here $x_1 = 4$) is placed at the root of the tree and the sequence is split into x_- composed of values lower than x_1 (here $x_- = (2, 1, 3)$) at the left and x_+ composed of values larger higher than x_1 (here $x_+ = (5, 6)$) at the right. The root and both left and right binary search trees are then combined as in (1) to obtain the full binary search tree.

Consider now trying to define binary search trees for a two-sided infinite sequence $x = (\ldots, x_{-1}, x_0, x_1, \ldots)$. A simple first case to consider would be x such that $x_i = i$ for all $i \in \mathbb{Z}$. Since the binary search tree of the sequence $(1, 2, \ldots)$ is the infinite right-downward path, we would naturally like the binary search tree of this x to be the combination of the infinite right-downward and infinite left-upward paths. We thus start by defining the notion of *left-upward* in a binary tree.

Since right-downward paths are encoded by $\mathbf{1}^k$ for $k \geq 0$, we want to encode left-upward paths with $\mathbf{1}^k$ for $k < 0$. Naturally, negative powers of words are not properly defined, but we can add a new symbol Υ (*up*silon) to our alphabet $\{\mathbf{0}, \mathbf{1}\}$ to describe $\mathbf{1}^{-1}$. From now on, we thus use the convention that

$$\mathbf{1}^k = \begin{cases} \mathbf{1}^k & \text{if } k > 0 \\ \varnothing & \text{if } k = 0 \\ \Upsilon^{|k|} & \text{if } k < 0 \, , \end{cases} \tag{2}$$

and we let the extension of the infinite binary tree be

$$\mathbb{Y} := \bigcup_{k\in\mathbb{Z}} \{\mathbf{1}^k\} \cup \mathbf{1}^k\mathbf{0}\mathbb{T} = \mathbb{T} \cup \left(\bigcup_{k\geq 1} \{\Upsilon^k\} \cup \Upsilon^k\mathbf{0}\mathbb{T}\right) \; .$$

We further let $\mathbb{T_0} := \{\varnothing\} \cup \mathbf{0}\mathbb{T}$ be the binary tree obtained by removing the first right edge of the complete binary tree; this tree will play a key role in our study, in particular since its shifted sequence $(\mathbf{1}^k\mathbb{T_0})_{k\in\mathbb{Z}}$ is a partition of $\mathbb{Y}$.

Let $x = (\ldots, x_{-1}, x_0, x_1, \ldots)$ be a two-sided infinite sequence with distinct values. Since the binary search tree of an infinite one-sided sequence is built as the limit of the tree obtained from the first n indices, we want to define the binary search tree of x as the limit of $T\langle x(n)\rangle$ where $x(n) = \{x_{-n}, x_{-n+1}, \ldots, x_n\}$. Unfortunately, this approach does not directly work, in particular since x_{-n} plays too big a role in the structure of $T\langle x(n)\rangle$. However, by re-rooting $T\langle x(n)\rangle$ according to some shift s_n so that the tree is approximately as high as it is low, we obtain a sequence of trees $(\mathbf{1}^{-s_n}T\langle x(n)\rangle)_{n\in\mathbb{N}}$ which are not directly subtrees of each other, but have a growth-like behaviour. Indeed, under mild assumptions on x, this sequence of trees eventually has a fixed structure on any ball of finite radius around the root $\varnothing$. This allows us to define $Y\langle x\rangle \subseteq \mathbb{Y}$ as the limit of $(\mathbf{1}^{-s_n}T\langle x(n)\rangle)_{n\in\mathbb{N}}$ on any finite ball, corresponding to the generalization of binary search trees to two-sided infinite sequences. We call $Y\langle x\rangle$ the *redwood tree* of x and refer to Section 2.2 for a more precise definition.

Fix $q \in [0, 1)$ and let $\mathcal{X}_q$ be a two-sided infinite Mallows permutation on $\mathbb{Z}$. Write $Y_q = Y\langle\mathcal{X}_q\rangle$ for the redwood tree drawn from $\mathcal{X}_q$ and let $\mathcal{Y}_q = (o, Y_q)$ be the rooted tree where o is uniformly sampled from the set

$$Y_q \cap \mathbb{T_0} = \{\varnothing\} \cup \left(Y_q \cap \mathbf{0}\mathbb{T}\right).$$

In other words, o is a uniform element chosen from the set composed of the root of the tree and the left subtree of the root. We now state our main theorem.

Theorem 1.1 (Local limit of Mallows trees). *Fix $q \in [0, 1)$. For all $n \in \mathbb{N}$, let $X_{n,q}$ be distributed as $\pi_{n,q}$. Then we have*

$$T\langle X_{n,q}\rangle \xrightarrow{\ell.\text{a.s.}} \mathcal{Y}_q,$$

where the convergence occurs according to the almost-sure local topology, defined in Section 2.3.

The proof of Theorem 1.1 can be found in Section 3.4. This result states an interesting limit swap between binary search tree structures and random permutation models; indeed, it can be interpreted as "the limit of the binary search tree of a finite Mallows permutation is the binary search tree of the infinite Mallows permutation". Another important fact related to this theorem is that, to the best of our knowledge, this is the first occurrence of a non-trivial almost-sure local convergence for naturally growing sequences of graphs. This strong type of convergence arises from the the regenerative property of Mallows permutations and heavily relies on the strong law of large numbers. To complement Theorem 1.1, we now provide a clearer characterization of the distribution of $\mathcal{Y}_q$. For the rest of this work, we write $\text{GEOMETRIC}^+(p)$ for the geometric distribution with parameter p and support $\mathbb{N} = \{1, 2, \ldots\}$, whereas $\text{GEOMETRIC}(p)$ refers to the geometric distribution with 0 in its support. We further let $\text{GEOMETRIC}^\star(p)$ be the sized biased version of $\text{GEOMETRIC}(p)$ subtracted by 1. All these distributions are summarized by

$$(3) \qquad \mathbb{P}(G = k) = \begin{cases} (1-p)^k p & \text{for } k \geq 0, \text{ if } G \sim \text{GEOMETRIC}(p) \\ (1-p)^{k-1} p & \text{for } k \geq 1, \text{ if } G \sim \text{GEOMETRIC}^+(p) \\ (k+1)(1-p)^k p^2 & \text{for } k \geq 0, \text{ if } G \sim \text{GEOMETRIC}^\star(p) \end{cases}$$

Theorem 1.2 (Distribution of the redwood Mallows tree). *Fix $q \in [0, 1)$ and let Y_q be the redwood tree drawn from a two-sided infinite Mallows permutation. For any $k \in \mathbb{Z}$, let*

$$Y_q(k) = Y_q \cap \mathbf{1}^k\mathbf{0}\mathbb{T} = \left(Y_q \cap \mathbf{1}^k\mathbb{T_0}\right) \setminus \{\mathbf{1}^k\}$$

be the left subtree of the node $\mathbf{1}^k$. Then $(Y_q(k))_{k\in\mathbb{Z}}$ is a two-sided sequence of independent Mallows trees with parameter q and size $\text{GEOMETRIC}(1-q)$, except for $Y_q(0)$ having size $\text{GEOMETRIC}^\star(1-q)$. Moreover, the structure of the tree is independent given the size, meaning that for $n \geq 0$ and t of size

n, we have

$$\mathbb{P}\Big(Y_q(k) = \mathbf{1}^k \mathbf{0} t \;\Big|\; \big|Y_q(k)\big| = n\Big) = \mathbb{P}\Big(T\langle X_{n,q}\rangle = t\Big),$$

where $X_{n,q}$ is $\pi_{n,q}$-distributed. Conversely, these properties fully characterize the distribution of Y_q.

The proof of Theorem 1.2 can be found in Section 3.3. The distribution of the left subtrees in Y_q follows from the global structure of Mallows trees as studied in [EGW12, Section 7], in particular the fact that the subtree of a Mallows tree is, conditionally given its size, a Mallows tree itself. Similarly, the fact that their sizes are geometrically distributed follows from the distribution of the distance between two consecutive records in a Mallows permutations and was already observed in the case of the one-sided infinite Mallows distribution in [ABC21, Lemma 1.10]. What might come as a surprise however, is the size-biased distribution of the left subtree of the root. This can be understood by seeing Y_q as the limiting object of $T\langle X_{n,q}\rangle$. Indeed, since the sizes of the left subtrees in a finite Mallows tree converge to a sequence of independent geometric random variables, the size of the left subtree containing a uniformly chosen node would have the size-biased version of the geometric distribution. The fact that Y_q, defined as the redwood tree drawn from the two-sided infinite Mallows distribution, also has this property arises from the fact that the size-biased geometric distribution subtracted by 1 and the sum of two independent geometric random variables have the same distribution. More details regarding this property and its relation to the proof of Theorem 1.2 can be found in Section 3.3 and is in particular shown in (14).

We conclude this section with three other limits of Mallows trees, depending on the topology considered. The first one is the *local rooted convergence*, or simply *rooted convergence*: if instead of considering the local behaviour of a uniformly sampled node in the tree, we fix one (the root), and consider the structure of the tree as seen from this node. This yields the classical binary search tree drawn from a Mallows permutation on $\mathbb{N}$, as stated below.

Theorem 1.3 (Rooted limit of Mallows trees). *Fix $q \in [0,1)$. For any $n \in \mathbb{N}$, let $X_{n,q}$ be $\pi_{n,q}$-distributed. Then we have*

$$T\langle X_{n,q}\rangle \xrightarrow{\varnothing} T\langle X_{\infty,q}\rangle,$$

where $X_{\infty,q}$ is a Mallows permutation on $\mathbb{N}$ with parameter q and the convergence occurs according to the rooted topology, defined in Section 2.3.1.

The second one is the strong Gromov–Hausdorff–Prokhorov (GHP) topology and corresponds to the distributional limit of the tree seen as a metric space. After proper scaling, arising from the height of the tree as shown in [ABC21], this yields a rather simple metric space with uniform measure on $[0,1]$.

Theorem 1.4 (Strong GHP limit of Mallows trees). *Fix $q \in [0,1)$. For any $n \in \mathbb{N}$, let $X_{n,q}$ be $\pi_{n,q}$-distributed. Let d_n be the graph distance in $T\langle X_{n,q}\rangle$ and μ_n be the uniform distribution on the nodes of $T\langle X_{n,q}\rangle$. Then we have*

$$\Big(T\langle X_{n,q}\rangle, \big((1-q)n\big)^{-1} \cdot d_n, \mu_n\Big) \xrightarrow{\mathrm{GHP}} \mathcal{E}_u,$$

where the convergence occurs according to the strong GHP topology, defined in Section 4.2, and

$$\mathcal{E}_u = \Big([0,1], d, \mu_{[0,1]}\Big)$$

with $d(x,y) = |x-y|$ and $\mu_{[0,1]}$ being the uniform measure on $[0,1]$.

The third and last one is according to the newly introduced *subtree size topology* [Grü23] and corresponds to the convergence of the ratio of nodes in each subtree. Since left hanging subtrees have geometric sizes in a Mallows tree, this yields the infinite rightward path.

Theorem 1.5 (Subtree size limit of Mallows trees). *Fix $q \in [0,1)$. For any $n \in \mathbb{N}$, let $X_{n,q}$ be $\pi_{n,q}$-distributed. Then we have*

$$T\langle X_{n,q} \rangle \xrightarrow{\mathrm{ssc}} \psi_R \,,$$

where

$$(4) \qquad \qquad \psi_R(v) = \begin{cases} 1 & \text{if } v \in \{\mathbf{1}^k : k \geq 0\} \\ 0 & \text{otherwise} \end{cases}$$

and the convergence occurs according to the subtree size topology, defined in Section 4.3.

The proofs of Theorem 1.3, Theorem 1.4, and Theorem 1.5 can respectively be found in Section 4.1, Section 4.2, and Section 4.3.

2. Notations and background work

In this section, we precise the definitions of the objects introduced in Section 1.1 and introduce further notations and results useful for the proofs of the different theorems.

2.1. **Mallows permutations.** In this section, we describe useful constructions of finite, one-sided, and two-sided Mallows permutations. These constructions will prove to be useful when studying the properties of the related binary search trees.

2.1.1. *Finite and one-sided Mallows permutations.* There exists various methods to sample Mallows permutations and, since we are interested in the limit of objects related to Mallows permutations with increasing size n, it is worth considering an adapted coupling. For this reason, we consider the *infinite Bernoulli model* as developed in [ABC21, Section 1.4] which we explain below.

Consider $B = (B_{i,j})_{i,j \geq 1}$ an infinite matrix with independent Bernoulli$(1-q)$ entries. Let $I_1, I_2, \dots$ be the first non-zero entry of each row not already chosen, that is

$$I_i = \min \left\{ j \notin \{I_1, \dots, I_{i-1}\} : B_{i,j} = 1 \right\},$$

and let $X_{n,q}$ be the permutation corresponding to the ordering of $(I_1, \dots, I_n)$. Then, thanks to [ABC21, Proposition 1.8], $X_{n,q}$ is $\pi_{n,q}$-distributed. Moreover, the sequence $X_{\infty,q} = (I_1, I_2, \dots)$ actually corresponds to a permutation, since every column of B has at least one non-zero entry, and is distributed as the Mallows distribution on $\mathbb{N}$ [GO10, Definition 4.4]. From now on, we assume that $X_{n,q}$, for $n \in \mathbb{N} \cup \{\infty\}$, refers to the random variable obtained using the coupling from the infinite Bernoulli model.

We now state some useful notations and results from [ABC21]. First, let

$$(5) \qquad \qquad R_{n,q} := \left| \left\{ j : \forall i < j, X_{n,q}(j) > X_{n,q}(i) \right\} \right| = \left| \left\{ j : \forall i < j, I_j > I_i \right\} \right|$$

be the number of records of $X_{n,q}$ and

$$F_{n,q} := \max \left\{ I_1, \dots, I_n \right\} - n$$

be the number of *free spaces* in the tree $T\langle X_{n,q} \rangle$. We refer to $F_{n,q}$ as the number of free spaces since it corresponds to the number of elements added to the left of $\mathbf{1}^{R_{n,q}}$ between $T\langle X_{n,q} \rangle$ and $T\langle X_{\infty,q} \rangle$; indeed, a better definition of $F_{n,q}$ is

$$(6) \qquad \qquad F_{n,q} = \left| \left(T\langle X_{\infty,q} \rangle \cap \bigcup_{0 \leq k \leq R_{n,q}} \mathbf{1}^k \mathbb{T}_{\mathbf{0}} \right) \setminus T\langle X_{n,q} \rangle \right| \,.$$

We refer to Figure 2 for a visual representation of this property. Interestingly, the previous equation implies that the number of left hanging subtrees in $T\langle X_{n,q} \rangle$ that are not full yet can be bounded by the number of free spaces

$$(7) \qquad \qquad \left| \left\{ 0 \leq k \leq R_{n,q} : T\langle X_{\infty,q} \rangle \cap \mathbf{1}^k \mathbb{T}_{\mathbf{0}} \neq T\langle X_{n,q} \rangle \cap \mathbf{1}^k \mathbb{T}_{\mathbf{0}} \right\} \right| \leq F_{n,q} \,,$$

bound which will be useful in proving Theorem 1.1.

$(I_j)_{1 \leq j \leq 1} = (4)$
$X_{1,q} = (1)$
$R_{1,q} = 1$
$F_{1,q} = 3$
$E_{1,q} = 1$

$(I_j)_{1 \leq j \leq 2} = (4, 2)$
$X_{2,q} = (2, 1)$
$R_{2,q} = 1$
$F_{2,q} = 2$
$E_{2,q} = 1$

$(I_j)_{1 \leq j \leq 3} = (4, 2, 5)$
$X_{3,q} = (2, 1, 3)$
$R_{3,q} = 2$
$F_{3,q} = 2$
$E_{3,q} = 1$

$(I_j)_{1 \leq j \leq 4} = (4, 2, 5, 1)$
$X_{4,q} = (3, 2, 4, 1)$
$R_{4,q} = 2$
$F_{4,q} = 1$
$E_{4,q} = 1$

$(I_j)_{1 \leq j \leq 5} = (4, 2, 5, 1, 6)$
$X_{5,q} = (3, 2, 4, 1, 5)$
$R_{5,q} = 3$
$F_{5,q} = 1$
$E_{5,q} = 1$

$(I_j)_{1 \leq j \leq 6} = (4, 2, 5, 1, 6, 3)$
$X_{6,q} = (4, 2, 5, 1, 6, 3)$
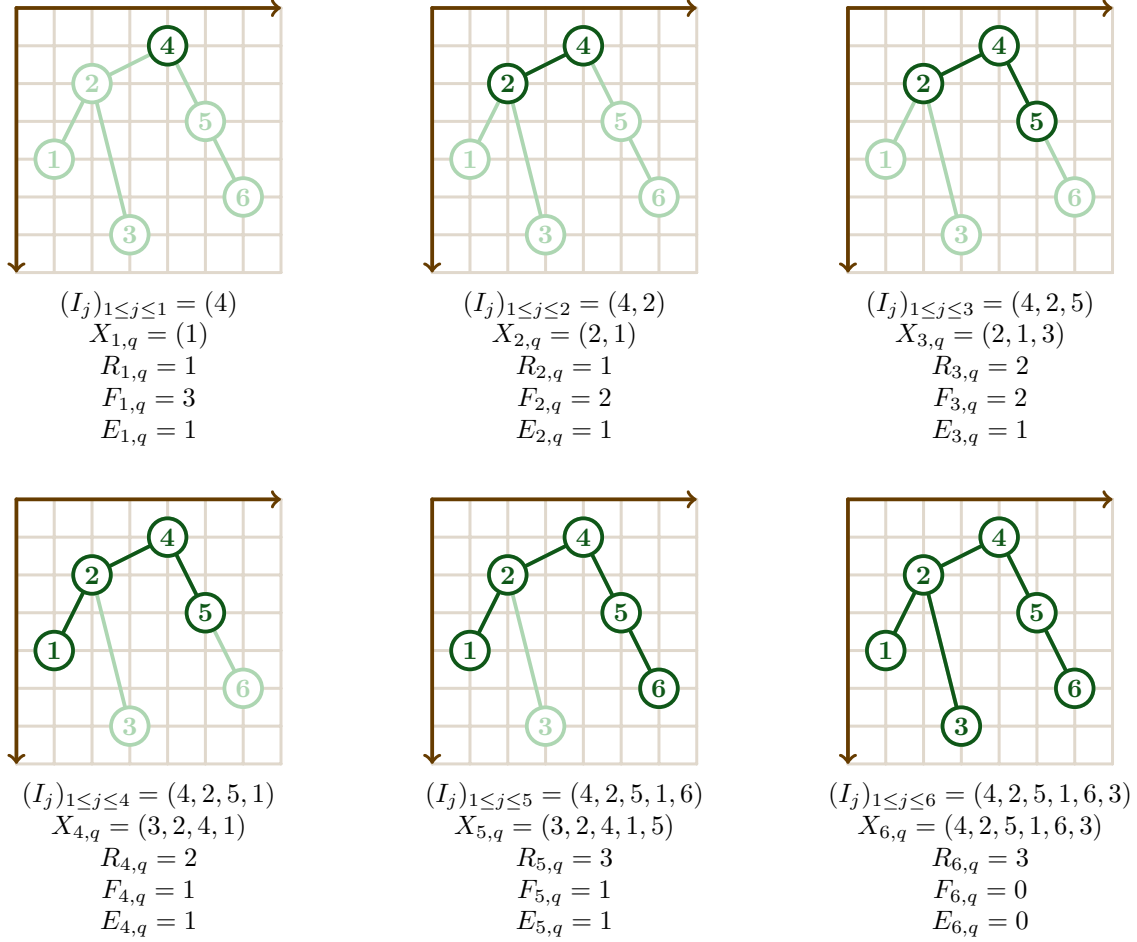$R_{6,q} = 3$
$F_{6,q} = 0$
$E_{6,q} = 0$

FIGURE 2. The representation of the first few number of records and free spaces for the sequence $(I_j)_{j \geq 1} = (4, 2, 5, 1, 6, 3, \ldots)$ along with the corresponding permutations and binary search trees. We further provide the value of $E_{n,q}$, the number of left subtrees to be filled, as defined on the left hand side of (7). From this representation, we verify the following properties. **(1)** $X_{n,q}$ is the ordering of the sequence $(I_1, \ldots, I_n)$. **(2)** $R_{n,q}$ is both the number of records of $X_{n,q}$ and the number of nodes on the rightmost path of the tree. **(3)** $F_{n,q}$ is both the difference between the maximal value of $(I_1, \ldots, I_n)$ and n and the number of nodes to be filled on the left part of the tree (i.e. the number of shaded green nodes to the left of the rightmost dark green node). **(4)** $E_{n,q} \leq F_{n,q}$ as claimed in (7) since the number of incomplete left subtrees is less than the number of left nodes to be filled.

We now provide two important results on $R_{n,q}$ and $F_{n,q}$. The first one states the asymptotic behaviour of $R_{n,q}$ and has strong enough tail bounds to provide an almost-sure convergence after re-scaling.

Lemma 2.1. *Fix $q \in [0, 1)$. Let $R_{n,q}$ be the number of records of $X_{n,q}$ coupled according to the infinite Bernoulli model. Then we have*

$$\frac{R_{n,q} - n(1-q)}{\sqrt{n} \log n} \xrightarrow{\text{a.s.}} 0.$$

In particular, this implies that $R_{n,q}/n \xrightarrow{\text{a.s.}} (1-q)$

Proof. Using [ABC21, Lemma 2.4], we know that, for any $\epsilon > 0$ and $n \in \mathbb{N}$, we have

$$\mathbb{P}\left(\left|\frac{R_{n,q}}{\mu(n,q)} - 1\right| > \epsilon\right) \leq \exp\left(-\left[1 - (1+\epsilon)\log\left(\frac{e}{1+\epsilon}\right)\right]\mu(n,q)\right)$$
$$+ \exp\left(-\left[1 - \frac{1}{1+\epsilon}\log\left((1+\epsilon)e\right)\right]\mu(n,q)\right),$$

where

$$\mu(n,q) = \sum_{k=2}^{n}\frac{1-q}{1-q^k}.$$

Now, let $\lambda(\epsilon)$ be the exponential power in the previous upper bound, that is

$$\lambda(\epsilon) = \min\left\{1 - (1+\epsilon)\log\left(\frac{e}{1+\epsilon}\right), 1 - \frac{1}{1+\epsilon}\log\left((1+\epsilon)e\right)\right\},$$

and observe that $\lambda(\epsilon) \sim \epsilon^2/2$ as $\epsilon \to 0$. Combining the previous results, we have that

$$\mathbb{P}\left(\left|\frac{R_{n,q} - n(1-q)}{\sqrt{n}\log n}\right| > \delta\right) = \mathbb{P}\left(\left|\frac{R_{n,q}}{\mu(n,q)} - \frac{n(1-q)}{\mu(n,q)}\right| > \delta \cdot \frac{\sqrt{n}\log n}{\mu(n,q)}\right)$$
$$\leq \mathbb{P}\left(\left|\frac{R_{n,q}}{\mu(n,q)} - 1\right| > \delta \cdot \frac{\sqrt{n}\log n}{\mu(n,q)} - \left|1 - \frac{n(1-q)}{\mu(n,q)}\right|\right)$$
$$\leq 2\exp\left(-\lambda\left(\delta \cdot \frac{\sqrt{n}\log n}{\mu(n,q)} - \left|1 - \frac{n(1-q)}{\mu(n,q)}\right|\right) \cdot \mu(n,q)\right).$$

But now, thanks to [ABC21, Proposition 2.5], we know that $\mu(n,q) = n(1-q) + O(\log n)$, which implies that

$$\delta \cdot \frac{\sqrt{n}\log n}{\mu(n,q)} - \left|1 - \frac{n(1-q)}{\mu(n,q)}\right| = \frac{\delta}{\sqrt{1-q}} \cdot \frac{\log n}{\sqrt{\mu(n,q)}} + o\left(\frac{\log n}{\sqrt{\mu(n,q)}}\right).$$

To conclude the proof of the lemma, recall that $\lambda(\epsilon) \sim \epsilon^2/2$ when $\epsilon \to 0$ to obtain

$$\mathbb{P}\left(\left|\frac{R_{n,q} - n(1-q)}{\sqrt{n}\log n}\right| > \delta\right) \leq \exp\left(-\frac{\delta^2\log^2 n}{2(1-q)} + o(\log^2 n)\right)$$

and the almost-sure convergences follows from applying the Borel-Cantelli lemma.   □

The second important result is related to $F_{n,q}$ and proves an other almost-sure convergence result.

Lemma 2.2. *Fix $q \in [0,1)$. Let $F_{n,q}$ be the number of free spaces as defined in* (6)*. Then, for any* $(a_n)_{n\geq 1}$ *such that $a_n/\log n \to \infty$, we have*

$$\frac{F_{n,q}}{a_n} \xrightarrow{\text{a.s.}} 0.$$

In particular, this implies that

$$\frac{F_{n,q}}{n} \xrightarrow{\text{a.s.}} 0.$$

Proof. Thanks to [ABC21, Proposition 1.11], we know that the moment generating function of $F_{n,q}$ is

$$\mathbb{E}\left[x^{F_{n,q}}\right] = \prod_{k=1}^{n}\frac{1-q^k}{1-q^k x}.$$

Moreover, for any $k \geq 1$ and $1 < x < 1/q$, we have

$$\frac{1-q^k}{1-q^k x} = 1 + (x-1)\frac{q^k}{1-q^k x} \leq 1 + (x-1)\frac{q^k}{1-qx}.$$

Fix now $\epsilon > 0$ and let $x = \sqrt{1/q} < 1/q$. Then, using Chernoff's bounds along with the previous inequality, we have

$$\mathbb{P}\left(\frac{F_{n,q}}{a_n} > \epsilon\right) \leq x^{-a_n\epsilon}\mathbb{E}\left[x^{F_{n,q}}\right] \leq q^{\frac{a_n\epsilon}{2}}\prod_{k=1}^{n}\left[1 + q^{k-1/2}\right],$$

where we used that $(x-1)/(1-qx) = (1/\sqrt{q} - 1)/(1 - \sqrt{q}) = 1/\sqrt{q}$. Finally, using the convexity of exponential, we obtain

$$\mathbb{P}\left(\frac{F_{n,q}}{a_n} > \epsilon\right) \leq \exp\left(-a_n \cdot \frac{\epsilon|\log q|}{2} + \sum_{k=1}^{n}q^{k-1/2}\right) \leq \exp\left(-a_n \cdot \frac{\epsilon|\log q|}{2} + \frac{\sqrt{q}}{1-q}\right),$$

which, combined with the Borel-Cantelli lemma, proves the desired almost-sure convergence. □

2.1.2. *Two-sided infinite Mallows permutations.* Using the results from [GO12], we provide a construction of Mallows permutations on $\mathbb{Z}$. A permutation on $\mathbb{Z}$ is said to be *admissible* if

$$\left|\{i \in \mathbb{N} : \sigma(i) \notin \mathbb{N}\}\right| + \left|\{i \in \mathbb{Z}\setminus\mathbb{N} : \sigma(i) \in \mathbb{N}\}\right| < \infty. \tag{8}$$

This property plays a key role here, since Mallows permutations on $\mathbb{Z}$ are almost-surely admissible, and since admissibility will be used as a condition for the construction of redwood trees in Section 2.2.

Let $X_{+,q}, X_{-,q}$ be two independent Mallows permutations on $\mathbb{N}$. Let $M = (M_1, M_2, \ldots)$ be a random sequence of independent non-negative integers such that, for all $i \geq 1$, M_i is distributed as a GEOMETRIC$(1 - q^i)$, that is, for all $m \in \mathbb{N}$

$$\mathbb{P}(M_i = m) = q^{im}(1 - q^i).$$

Note that this definition implies that $\sum_i M_i$ is almost-surely finite. Let now $\Lambda = \Lambda(M) = (\Lambda_1 \geq \Lambda_2 \geq \ldots)$ be an infinite decreasing sequence such that, for all $i \geq 1$, we have

$$M_i = \left|\{j \geq 1 : \Lambda_j = i\}\right|.$$

Using the notations of [Mac98, Chapter I], Λ is a *partition* and M is the corresponding sequence of *multiplicities*; in particular, Λ is a random partition such that, for any $\lambda = (\lambda_1, \lambda_2, \ldots)$, we have

$$\mathbb{P}(\Lambda = \lambda) = q^{|\lambda|}\cdot\prod_{k=1}^{\infty}(1 - q^k),$$

where $|\lambda| = \sum_{k=1}^{\infty}\lambda_k$, as provided in [GO12, (8)]. We now combine the triplet $(X_{+,q}, X_{-,q}, \Lambda)$ to create a permutation as follows.

Given Λ, let $\ell_i = i - \Lambda_i$ and note that $\ell_1 < \ell_2 < \ldots$. Further define $k_0 > k_{-1} > k_{-2} > \ldots$ such that $\{\ell_i : i \in \mathbb{N}\}$ and $\{k_i : i \in \mathbb{Z}\setminus\mathbb{N}\}$ form a partition of $\mathbb{Z}$. Now define $\mathcal{X}_q$ as a function of $(X_{+,q}, X_{-,q}, \Lambda)$ by

$$\mathcal{X}_q(x) = \begin{cases} X_{+,q}(i) & \text{if } x = \ell_i \\ 1 - X_{-,q}(1-i) & \text{if } x = k_i \end{cases} \tag{9}$$

and note that $\mathcal{X}_q$ is indeed a permutation of $\mathbb{Z}$. Thanks to [GO12, Section 3], $\mathcal{X}_q$ is distributed as a Mallows permutation on $\mathbb{Z}$, in particular by uniqueness of the q-invariant distribution.

We note here that the *balanced* property (also introduced in [GO12]) does not play a role when considering the corresponding binary search trees. So, although Mallows permutations on $\mathbb{Z}$ are also almost-surely balanced, this property will not be mentioned in the remainder of this work; the focus here is placed on admissible permutations, for which we can define the corresponding redwood tree.

2.2. **Redwood trees.** In this section, we properly define redwood trees to complete their introduction from Section 1.1. We recall the convention on the alphabet $\{\mathbf{0}, \mathbf{1}, \Upsilon\}$ as introduced in (2) and that

$$\mathbb{Y} = \bigcup_{k \in \mathbb{Z}} \mathbf{1}^k \mathbb{T}_\mathbf{0} \,,$$

where $\mathbb{T}_\mathbf{0} = \{\varnothing\} \cup \mathbf{0}\mathbb{T}$. We also recall that $\sigma : \mathbb{Z} \to \mathbb{Z}$ is said to be admissible if it satisfies (8). We start by providing an important property of admissible functions. A record index of σ is a an index r such that, for all $k < r$, we have $\sigma(r) > \sigma(k)$; given a record index r, $\sigma(r)$ is the corresponding record.

Lemma 2.3. *Let $\sigma : \mathbb{Z} \to \mathbb{Z}$ be an admissible injective function. Then, for any $k \in \mathbb{Z}$, there exists two record indices r_- and r_+ such that $r_- < k < r_+$.*

Proof. Let $\sigma : \mathbb{Z} \to \mathbb{Z}$ be injective and satisfying (8). Fix $k \in \mathbb{Z}$. Using that σ is admissible, we first have that

$$M_k = \max\big\{\sigma(\ell) : \ell \leq k\big\} < \infty \,.$$

But then, again thanks to σ being admissible, we know that

$$r_+ = \min\big\{\ell > k : \sigma(\ell) > M_k\big\}$$

is well-defined and is a record index on the right of k. Similarly, the admissibility property of σ gives us that

$$m_k = \max\big\{\sigma(\ell) : \ell < k\big\} < \infty \,,$$

which then, combined with the fact that σ is injective and thus admits an inverse $\sigma^{-1} : \sigma(\mathbb{Z}) \to \mathbb{Z}$, gives us that

$$r_- = \sigma^{-1}(m_k)$$

is well-defined and is a record index on the left of k. □

Let σ be an admissible permutation of $\mathbb{Z}$. Using Lemma 2.3, the set of record indices of σ is a two-sided infinite sequence. Thus, we say that the sequence $r = (\ldots, r_{-1}, r_0, r_1, \ldots)$ is the *standard record representation* of σ if and only if

- $\{r_k : k \in \mathbb{Z}\}$ is the complete set of record indices of σ;
- the sequence is in strictly increasing order; and
- it is (uniquely) indexed so that $\sigma(r_0) \leq 0 < \sigma(r_1)$.

When the index 0 of the sequence is not completely clear, we will underline it: $(\ldots, r_{-1}, \underline{r_0}, r_1, \ldots)$.

Given $r = (\ldots, r_{-1}, r_0, r_1, \ldots)$ the standard record representation of σ, we let $\sigma[k] = (\sigma(i) : \sigma(r_k) < \sigma(i) < \sigma(r_{k+1}))$ be the sequence of images of the permutation between two consecutive records, in the same order as in σ; it is worth noting that this is well-defined since, for two record indices $r < r'$, we also have $\sigma(r) < \sigma(r')$. Finally, we define the redwood tree of σ to be

$$(10) \qquad\qquad Y\langle\sigma\rangle := \bigcup_{k \in \mathbb{Z}} \{\mathbf{1}^k\} \cup \mathbf{1}^k \mathbf{0} T\langle\sigma[k]\rangle \,.$$

From this definition, it is clear that $Y\langle\sigma\rangle \subseteq \mathbb{Y}$ and we now explore further properties of this object.

Lemma 2.4. *Let σ be a permutation on $\mathbb{N}$. Let $\tilde{\sigma}$ be the extension of σ to $\mathbb{Z}$ by the identity, that is*

$$\tilde{\sigma}(k) = \begin{cases} \sigma(k) & \text{if } k \in \mathbb{N} \\ k & \text{otherwise}. \end{cases}$$

Then, we have

$$Y\langle\tilde{\sigma}\rangle = T\langle\sigma\rangle \cup \bigcup_{k<0}\{\mathbf{1}^k\} = T\langle\sigma\rangle \cup \bigcup_{k \in \mathbb{N}}\{\Upsilon^k\} \,,$$

which implies that $Y\langle\tilde{\sigma}\rangle \cap \mathbb{T} = T\langle\sigma\rangle$.

Proof. First, note that, if $r_1 < r_2 < \ldots$ is the ordered sequence of record indices of σ, then it can be used to construct the binary search tree of σ as follows

$$T\langle\sigma\rangle = \bigcup_{k\geq 0}\{\mathbf{1}^k\} \cup \mathbf{1}^k\mathbf{0}T\langle\sigma[k]\rangle \,,$$

where $\sigma[k]$ is defined as the sequence of images between the two consecutive records $\sigma(r_k)$ and $\sigma(r_{k+1})$ (using the convention that $r_0 = 0$ and $\sigma(r_0) = 0$). Now, it suffices to see that the standard record representation of $\tilde{\sigma}$ is $(\ldots, -2, -1, \underline{0}, r_1, r_2, \ldots)$ to obtain that $\tilde{\sigma}[k]$ is equal to $\sigma[k]$ when $k \geq 0$ and is the empty permutation otherwise. The result then follows from the definition of the redwood tree and the previous identity on the binary search tree of σ. $\square$

The previous lemma provides a first reason as to why the redwood tree can be considered as the extension of the binary search tree to two-sided admissible permutations. However, another characteristic relation between redwood trees and binary search trees is given in the following proposition. For $a, b \in \mathbb{Z}$, we let $[\![a, b]\!] = \{a, a+1, \ldots, b-1, b\}$ be the set of integers between a and b.

Proposition 2.5. *Let σ be an admissible permutation on $\mathbb{Z}$ and $r = (\ldots, r_{-1}, r_0, r_1, \ldots)$ be its standard record representation. For any $n \in \mathbb{N}$, let $\sigma_n = (\sigma(-n), \ldots, \sigma(n-1), \sigma(n))$ be the restriction of σ to $[\![-n, n]\!]$ and let s_n be the number of records of σ_n smaller or equal than 0. Then, for any finite $I \subseteq \mathbb{Z}$, there exists n_0 such that, for all $n \geq n_0$ we have*

$$Y\langle\sigma\rangle \cap \left(\bigcup_{k\in I}\mathbf{1}^k\mathbb{T}_\mathbf{0}\right) = \mathbf{1}^{-s_n}T\langle\sigma_n\rangle \cap \left(\bigcup_{k\in I}\mathbf{1}^k\mathbb{T}_\mathbf{0}\right).$$

In words, the redwood tree $Y\langle\sigma\rangle$ and the shifted binary search tree $\mathbf{1}^{-s_n}T\langle\sigma_n\rangle$ are locally equal for n large enough.

Proof. This proof being based on a very visual idea, we invite the reader to first take a look at Figure 3 before continuing below. Without loss of generality, we assume that $I = [\![-K, K]\!]$ for some $K \geq 1$. For any $n \in \mathbb{N}$, let $r(n) = (r_{1-s_n}(n), \ldots, r_{-1}(n), r_0(n), \ldots, r_{m_n}(n))$ be the ordered set of record indices of σ_n with the convention that $\sigma_n(r_0(n)) \leq 0 < \sigma_n(r_1(n))$ (so that s_n is indeed the number of records of σ_n smaller or equal than 0). Finally, since σ is a bijection of $\mathbb{Z}$, define

$$n_0 = \max\left\{|k| : k \in \sigma^{-1}\Big([\![\sigma(r_{-K}), \sigma(r_{K+1})]\!]\Big)\right\}$$

so that

$$[\![\sigma(r_{-K}), \sigma(r_{K+1})]\!] \subseteq \sigma\big([\![-n_0, n_0]\!]\big).$$

In particular, this implies that $n_0 \geq \max\{r_{K+1}, -r_{-K}\}$. Note that $\sigma([\![-n_0, n_0]\!])$, or in general $\sigma([\![-n, n]\!])$, is not necessarily a segment of $\mathbb{Z}$ itself; thus the previous inclusion simply means that it contains the segment defined by $\sigma(r_{-K})$ and $\sigma(r_{K+1})$ but may very well contain other isolated segments or singletons. We now prove that n_0 satisfies the desired condition.

Let $n \geq n_0$. First, for any $-K \leq k \leq K+1$, thanks to the fact that r_k is a record index of σ, it is also a record index of σ_n. Then, by the definition of both r and $r(n)$, this implies that $r_k = r_k(n)$. Now, for any $-K \leq k \leq K$, using that $\sigma(r_k) < \sigma(r_{k+1})$, we have that

$$[\![\sigma(r_k), \sigma(r_{k+1})]\!] \subseteq [\![\sigma(r_{-K}), \sigma(r_{K+1})]\!] \subseteq \sigma\big([\![-n_0, n_0]\!]\big) \subseteq \sigma\big([\![-n, n]\!]\big).$$

Since σ_n is the restriction of σ to $[\![-n, n]\!]$, this also implies that

$$[\![\sigma_n(r_k(n)), \sigma_n(r_{k+1}(n))]\!] = [\![\sigma_n(r_k), \sigma_n(r_{k+1})]\!] \subseteq \sigma_n\big([\![-n, n]\!]\big).$$

Finally, recall the definition of $\sigma[k]$ as the sequence of images of σ between $\sigma(r_k)$ and $\sigma(r_{k+1})$ and similarly define $\sigma_n[k]$ to obtain that $\sigma[k] = \sigma_n[k]$ for any $-K \leq k \leq K$.

Decompose $T\langle\sigma_n\rangle$ using its records in the same way as in the proof of Lemma 2.4 to obtain

$$T\langle\sigma_n\rangle = \bigcup_{-s_n \leq k \leq m_n+1} \{\mathbf{1}^{k+s_n}\} \cup \mathbf{1}^{k+s_n}\mathbf{0}T\langle\sigma_n[k]\rangle$$

$$= \mathbf{1}^{s_n} \bigcup_{-s_n \leq k \leq m_n+1} \{\mathbf{1}^k\} \cup \mathbf{1}^k\mathbf{0}T\langle\sigma_n[k]\rangle \ .$$

Using the previously shown result stating that $\sigma_n[k] = \sigma[k]$ for $-K \leq k \leq K$, we have that

$$\mathbf{1}^{-s_n}T\langle\sigma_n\rangle \cap \left(\bigcup_{k \in I}\mathbf{1}^k\mathbb{T}_\mathbf{0}\right) = \bigcup_{|k| \leq K} \{\mathbf{1}^k\} \cup \mathbf{1}^k\mathbf{0}T\langle\sigma[k]\rangle$$

$$= Y\langle\sigma\rangle \cap \left(\bigcup_{k \in I}\mathbf{1}^k\mathbb{T}_\mathbf{0}\right) ,$$

where the last inequality directly follows from the definition of $Y\langle\sigma\rangle$. This concludes the proof of the proposition by recalling that $n \geq n_0$ was arbitrarily chosen. $\qquad\square$

2.2.1. *Disclaimer on the name and notations for redwood trees.* One might wonder why choosing to name the extension of binary search trees as *redwood trees*. A first version of the article actually referred to them as *right infinite binary search trees*, since they are two-sided infinite extensions of binary search tree growing in a right-downward direction[1]. This name also implied the existence of *left* infinite binary search tree, which are briefly discussed in Section 5.1, and relate here to the limit of finite Mallows trees when $q > 1$.

The name right infinite binary search tree, although easily abbreviated to RIBST, was a bit too long (the name *binary search tree* is already rather long on its own) and felt like a precise description of an object rather than its actual name. For this reason, and since these trees are both infinitely high and infinitely deep, we decided to name them based on the highest species of trees in the world: the *sequoia sempervirens*, also called *California redwoods*[2].

Following this logic, redwood trees could easily refer to both right and left infinite binary search trees. However, the fact that both *right infinite binary search trees* and *redwood trees* start with an 'r' convinced us of this choice; thus, we invite anyone who would rather use left infinite binary search trees than right ones to refer to them as *ledwood trees*.

Finally, and as far as it made sense in the context of this article, we always tried to use the letter T when considering one-sided trees, that is trees with a natural node at the top, and the letter Y for two-sided trees. The reason for choosing Y is that it can be interpreted as the letter T where we extended its two top branches upwards; with that logic, Y can be used for both redwood and ledwood trees.

2.3. **Local convergence of graphs.** In this section, we provide some background on the topology of the local convergence of graphs. Most of what is presented here can be found in more details in [Hof23, Section 2].

A graph $G = (V, E)$ is defined as a pair of vertices V and edges E. We write $|G| := |V|$ for the size of G and use the notation $v \in G$ to say that v is a vertex from V. Given a graph G and one of its vertex $v \in G$, write $B_r(v, G)$ for the rooted ball of radius r around v in G; for example $B_0(v, G) = (v, (\{v\}, \emptyset))$. Given two rooted graphs (o, G) and (o', G'), where $G = (V, E)$ and $G' = (V', E')$, say that (o, G) and (o', G') are *isomorphic* and write it $(o, G) \equiv (o', G')$ if there exists a bijection $\varphi : V \to V'$ such that $\varphi(o) = o'$ and

$$(u, v) \in E \iff (\varphi(u), \varphi(v)) \in E' \ .$$

---

[1]This is naturally equivalent to growing in a left-upward direction in the case of a two-sided infinite tree, but since the idea was to extend right-downward paths in classical binary search trees, we used the downward direction as the standard one.

[2]https://en.wikipedia.org/wiki/Sequoia_sempervirens

$$\sigma_0 = (3)$$
$$r(0) = (1)$$
$$\sigma_0(r(0)) = (3)$$
$$s_0 = 0$$

$$\sigma_1 = (-1, 3, 1)$$
$$r(1) = (1, 2)$$
$$\sigma_1(r(1)) = (-1, 3)$$
$$s_1 = 1$$

$$\sigma_2 = (-3, -1, 3, 1, 4)$$
$$r(2) = (1, 2, 3, 5)$$
$$\sigma_2(r(2)) = (-3, -1, 3, 4)$$
$$s_2 = 2$$

$$\sigma_3 = (0, -3, -1, 3, 1, 4, -5)$$
$$r(3) = (1, 4, 6)$$
$$\sigma_3(r(3)) = (0, 3, 4)$$
$$s_3 = 1$$

$$\sigma_4 = (-2, 0, -3, -1, 3, 1, 4, -5, 7)$$
$$r(4) = (1, 2, 5, 7, 9)$$
$$\sigma_4(r(4)) = (-2, 0, 3, 4, 7)$$
$$s_4 = 2$$

$$\sigma_5 = (-7, -2, 0, -3, -1, 3, 1, 4, -5, 7, 2)$$
$$r(5) = (1, 2, 3, 6, 8, 10)$$
$$\sigma_5(r(5)) = (-7, -2, 0, 3, 4, 7)$$
$$s_5 = 3$$

$$\sigma_6 = (-6, -7, -2, 0, -3, -1, 3, 1, 4, -5, 7, 2, 6)$$
$$r(6) = (1, 3, 4, 7, 9, 11)$$
$$\sigma_6(r(6)) = (-6, -2, 0, 3, 4, 7)$$
$$s_6 = 3$$

$$\sigma_7 = (-4, -6, -7, -2, 0, -3, -1, 3, 1, 4, -5, 7, 2, 6, 5)$$
$$r(7) = (1, 4, 5, 8, 10, 12)$$
$$\sigma_7(r(7)) = (-4, -2, 0, 3, 4, 7)$$
$$s_7 = 3$$

$$\sigma = (\dots, -4, -6, -7, -2, 0, -3, -1, \underline{3}, 1, 4, -5, 7, 2, 6, 5, \dots)$$
$$r = (\dots, -7, -4, \underline{-3}, 0, 2, 4, \dots)$$
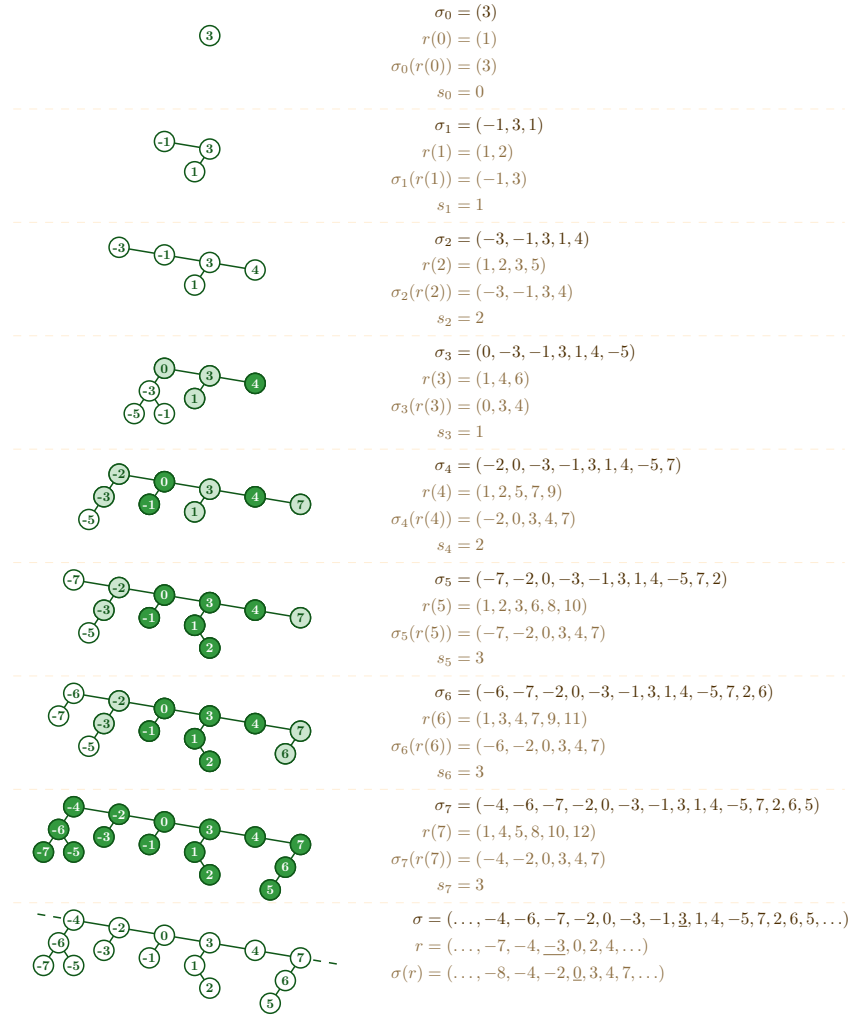$$\sigma(r) = (\dots, -8, -4, -2, \underline{0}, 3, 4, 7, \dots)$$

FIGURE 3.   A visual representation of the statement of Proposition 2.5 applied to $\sigma = (\dots, -4, -6, -7, -2, 0, -3, -1, \underline{3}, 1, 4, -5, 7, 2, 6, 5, \dots)$ (assuming that $\sigma(k) > 7$ for $k > 7$ and $\sigma(k) < -7$ for $k < -7$). On the right is σ_n corresponding to $(\sigma(-n), \dots, \sigma(n))$, $r(n)$ the record indices of σ_n, $\sigma_n(r(n))$ the records of σ_n, and s_n the number of records below 0. In the case of σ we also provide r, the standard record representation of σ, and $\sigma(r)$, the corresponding records; we note that, while we do not know the index of the next record on the left, we know this record equals -8. On the left are the binary search trees $(\mathbf{1}^{-s_n} T\langle \sigma_n \rangle)_{0 \le n \le 7}$ and the redwood tree $Y\langle \sigma \rangle$ with all the nodes $\varnothing$ vertically aligned. The nodes in light green are the ones whose label and position is the same as in the final redwood tree and the nodes in dark green correspond to subtrees in the final redwood tree (with the exact shape and labels). If writing $n_0(I)$ for the minimal value satisfying the statement of Proposition 2.5, we observe that $n_0(\emptyset) = 0$, $n_0(\{1\}) = 3$, $n_0(\{-1\}) = 4$, $n_0(\{0\}) = 5$, and $n_0(\{-3, -2, 2\}) = 7$. We also observe that $\sigma(\llbracket -7, 7 \rrbracket) = \llbracket -7, 7 \rrbracket$, explaining why all the nodes are dark green in the last binary search tree.

We now use these notations to define almost-sure local convergence.

Let $(G_n)_{n\geq 1}$ be a sequence of finite random graphs and let (o, G) be a (usually infinite) random rooted graph. We say that $(G_n)_{n\geq 1}$ converges locally almost-surely to (o, G) and we write it $G_n \xrightarrow{\ell.a.s.} (o, G)$ if, for any $r \geq 1$ and any rooted graph H, we have

$$(11) \qquad \frac{1}{|G_n|} \sum_{v \in G_n} \mathbb{1}_{\{B_r(v, G_n) \equiv H\}} \xrightarrow{a.s.} \mathbb{P}\big(B_r(o, G) \equiv H\big),$$

where the convergence occurs according to the standard almost-sure convergence of random variables. We note that the general definition of the local almost-sure convergence allows for the limit to be a random variable (ie a random measure on the space of rooted graphs) and we decided not to explicit this dependency for clarity and since all almost-sure local limits will be deterministic here; we refer to [Hof23, Section 2] for a more detailed discussion on the properties of local limits of random graphs.

2.3.1. *Other types of local convergence.* For the sake of simplicity and since it suits our purpose, we directly stated the right type of convergence for the main result without explaining how this definition came to existence. However, it is worth mentioning that there exists weaker and more common types of local convergence for graphs. For example, one could replace the convergence in (11) by any other type of random variable convergence, for example in probability. In particular, the most common type of local convergence is the *local weak convergence*, introduced in [AS04, BS01], obtained by simply taking the expectation on both sides of (11):

$$\frac{1}{|G_n|} \sum_{v \in G_n} \mathbb{P}\big(B_r(v, G_n) \equiv H\big) \longrightarrow \mathbb{P}\big(B_r(o, G) \equiv H\big);$$

note that the convergence here is simply that of real numbers.

A natural way to further generalize the notion of local convergence is to restrict the distribution of v. Indeed, considering the local weak convergence, the previous condition can be re-written as

$$\mathbb{P}\big(B_r(v, G_n) \equiv H\big) \longrightarrow \mathbb{P}\big(B_r(o, G) \equiv H\big),$$

where v is a uniformly sampled vertex of G_n. One could however allow any type of distribution for v. In particular, in the case of binary search tree, there exists a node with a special role: the root $\varnothing$. Therefore, we define the *rooted local convergence* for binary search trees as follows.

Let $(T_n)_{n\geq 1}$ be a sequence of finite binary search trees and let T be a binary search tree. We say that T is the rooted (local) limit of $(T_n)_{n\geq 1}$ and we write it $T_n \xrightarrow{\varnothing} T$ if, for any $r \geq 1$ and any rooted t, we have

$$\mathbb{P}\big(B_r(\varnothing, T_n) \equiv t\big) \longrightarrow \mathbb{P}\big(B_r(\varnothing, T) \equiv t\big).$$

It is worth mentioning that this could also be called *rooted local weak convergence* as it based on the convergence condition of the local weak convergence. However, this would imply that both *rooted local convergence in probability* or *rooted local almost-sure convergence* would also make sense; and while these two concepts could be defined using the same logic as before, these topologies actually simply state that $B_r(\varnothing, T_n) = B_r(\varnothing, T)$ for n large enough, which slightly defeats the purpose of local limits (all other types of local convergence do not require any specific coupling of the random variables involved in the definitions).

3. THE ALMOST-SURE LOCAL CONVERGENCE

Since it is easier to prove the convergence result from Theorem 1.1 on the tree characterized in Theorem 1.2 than directly on the redwood tree drawn from the two-sided Mallows permutation, we start by proving the result of Theorem 1.2.

3.1. **Records of the two-sided infinite Mallows tree.** This section strongly relies on the notations introduced in Section 2.1.2 and, in particular, uses the construction of the two-sided infinite Mallows permutation $\mathcal{X}_q$ from the triplet $(X_{+,q}, X_{-,q}, \Lambda)$. We recall that $\{\ell_1 < \ell_2 < \ldots\}$ and $\{\ldots < k_{-2} < k_{-1} < k_0\}$ form a partition of $\mathbb{Z}$, are defined by $\ell_i = 1 - \Lambda_i$, and are at the essence of the definition of $\mathcal{X}_q$ as given in (9).

We start by providing an exact characterization of the standard record representation of $\mathcal{X}_q$ given $X_{+,q}$, $X_{-,q}$, and Λ. Write $(r_1, r_2, \ldots)$ for the the ordered sequence of record indices of $X_{+,q}$ and define $(\rho_1, \rho_2, \ldots)$ inductively as follows. First, $\rho_1 = X_{-,q}^{-1}(1)$ and then, given ρ_i, let

$$\rho_{i+1} = X_{-,q}^{-1}\Big(\min \big\{X_{-,q}(k) : k > \rho_i\big\}\Big).$$

Proposition 3.1. *Using the previous notations, let $m \in \mathbb{N}$ be minimal so that $k_{1-\rho_m} < \ell_1$. Then, the standard record representation of $\mathcal{X}_q$ is given by*

$$(\ldots, k_{1-\rho_{m+2}}, k_{1-\rho_{m+1}}, \underline{k_{1-\rho_m}}, \ell_{r_1}, \ell_{r_2}, \ldots).$$

Proof. We first prove that $\ell_{r_1}, \ell_{r_2}, \ldots$ is a sequence of consecutive records for $\mathcal{X}_q$. Let $j \in \mathbb{N}$ and consider $k < \ell_{r_j}$. Then, either we have $k = \ell_i$ with $i < r_1$ and then

$$\mathcal{X}_q(k) = X_{+,q}(i) < X_{+,q}(r_j) = \mathcal{X}_q(\ell_{r_j}),$$

where the inequality uses that r_j is a record index of $X_{+,q}$; or we have $k = k_i$ for some $i \leq 0$ and then

$$\mathcal{X}_q(k) = 1 - X_{-,q}(1 - i) \leq 0 < X_{+,q}(r_j) = \mathcal{X}_q(\ell_{r_j}).$$

In both case, $\mathcal{X}_q(k) < \mathcal{X}_q(\ell_{r_j})$ so ℓ_{r_j} is a record index of $\mathcal{X}_q$. On the other hand, and using the same type of argument, if $\mathcal{X}_q$ has a record index between ℓ_{r_i} and $\ell_{r_{i+1}}$, then there has to exist a record between r_i and r_{i+1} in $X_{+,q}$, which contradicts the definition of $(r_1, r_2, \ldots)$. Before moving to the second part of the proof, we note that we always have $r_1 = 1$ and that $\mathcal{X}_q(\ell_{r_1}) > 0$.

We now prove that $\rho_1, \rho_2, \ldots$ are consecutive anti-record indices of $X_{-,q}$, that is indices ρ such that, for all $k > \rho$, $X_{-,q}(\rho) < X_{-,q}(k)$. The proof is divided into three steps: proving that this sequence is increasing, then proving that they are anti-record indices, and finally proving that there are not anti-record index between them.

By their definition, we note that

$$X_{-,q}(\rho_{i+1}) = \min \big\{X_{-,q}(k) : k > \rho_i\big\},$$

and since $X_{-,q}$ is bijective, this implies that $\rho_{i+1} > \rho_i$. Now, for any $i \in \mathbb{N}$ and any $k > \rho_i$, we also have that $k \geq \rho_{i-1}$ and so, by definition

$$X_{-,q}(\rho_i) \leq X_{-,q}(k).$$

Using that $X_{-,q}$ is a bijection along with the fact that $k \neq \rho_i$ leads to the desired strict inequality; this argument also works when $i = 1$ since, by definition, for any $k \neq X_{-,q}^{-1}(1) = \rho_1$, we have $X_{-,q}(k) > 1 = X_{-,q}(\rho_1)$. Finally, assume that there exists $\rho_i < k < \rho_{i+1}$ such that k is an anti-record index of $X_{-,q}$. Then, by definition of k being an anti-record, we have that

$$X_{-,q}(k) < X_{-,q}(\rho_{i+1}).$$

However, since $k > \rho_i$ and by definition of ρ_{i+1}, we also have that

$$X_{-,q}(k) \geq X_{-,q}(\rho_{i+1}),$$

which is a contradiction. This proves that $\rho_1, \rho_2, \ldots$ is the ordered sequence of anti-record indices of $X_{-,q}$

We conclude the proof by proving that the desired sequence is the standard record representation of $\mathcal{X}_q$. First of all, since $\mathcal{X}_q(\ell_{r_1}) > 0$ and any other positive record of $\mathcal{X}_q$ would also be a record of $X_{+,q}$, the record prior to ℓ_{r_1} necessarily corresponds to the index 0 in the standard record representation of

$\mathcal{X}_q$. Now, let r be a record index of $\mathcal{X}_q$ such that $\mathcal{X}_q(r) \leq 0$. Then $r = k_{1-i}$ for some $i \in \mathbb{N}$ and so, for any $k > i$, by the definition of r and since $k_{1-k} < k_{1-i} = r$, we have

$$X_{-,q}(k) = 1 - \mathcal{X}_q(k_{1-k}) > 1 - \mathcal{X}_q(k_{1-i}) = X_{-,q}(i).$$

This means that i is an anti-record index of $X_{-,q}$ and so any negative record of $\mathcal{X}_q$ can be mapped to an anti-record of $X_{-,q}$. For the same reason, the sequence of consecutive negative record of $\mathcal{X}_q$ corresponds to the sequence of consecutive anti-records of $X_{-,q}$. To conclude the proof, we just need to show that $k_{1-\rho_i}$ is not a record of $\mathcal{X}_q$ for $i < m$. This holds since, by definition of m, for any $i < m$, we have $k_{1-\rho_i} > \ell_1 = \ell_{r_1}$ (they cannot be equal) but also $\mathcal{X}_q(k_{1-\rho_i}) \leq 0 < \mathcal{X}_q(\ell_{r_1})$. $\qquad \square$

Proposition 3.1 provides a first understanding of the records of $\mathcal{X}_q$ given $(X_{+,q}, X_{-,q}, \Lambda)$. However, the definition of $k_{1-\rho_m}$ might not be fully clear and we provide here a simpler characterization.

Corollary 3.2. *Using the notations of Proposition 3.1, we have that*

$$\mathcal{X}_q(k_{1-\rho_m}) = 1 - \min\left\{X_{-,q}(k) : k > \Lambda_1\right\}.$$

Proof. First recall that m is defined so that $k_{1-\rho_m} < \ell_1 < k_{1-\rho_{m-1}}$. Using that the sequence $\Lambda_1 \geq \Lambda_2 \geq \ldots$ is eventually equal to 0, we observe that

$$\left|\left\{j \in \mathbb{Z} \setminus \mathbb{N} : k_j > \ell_1\right\}\right| = \Lambda_1.$$

In particular, this implies that $k_{1-\Lambda_1-1} < \ell_1 < k_{1-\Lambda_1}$. Since m is defined so that $k_{1-\rho_m} < \ell_1 < k_{1-\rho_{m-1}}$, it follows that $\rho_m > \Lambda_1 \geq \rho_{m-1}$.

Recall now the definition of ρ_m:

$$\rho_m = X_{-,q}^{-1}\left(\min\left\{X_{-,q}(k) : k > \rho_{m-1}\right\}\right),$$

which equivalently means that

$$X_{-,q}(\rho_m) = \min\left\{X_{-,q}(k) : k > \rho_{m-1}\right\}.$$

Knowing that $\rho_m > \Lambda_1$, the minimum is not reached for $\rho_{m-1} < k \leq \Lambda_1$, so that

$$X_{-,q}(\rho_m) = \min\left\{X_{-,q}(k) : k > \Lambda_1\right\}.$$

To conclude the proof, recall that

$$\mathcal{X}_q(k_{1-\rho_m}) = 1 - X_{-,q}(\rho_m),$$

giving us the desired result. $\qquad \square$

Note that the statements of Proposition 3.1 and Corollary 3.2, although applied to $\mathcal{X}_q$, are actually deterministic results: they only depend on the construction of $\mathcal{X}_q$ from $(X_{+,q}, X_{-,q}, \Lambda)$. We now want to combine these results to characterize the distribution of the distance between consecutive records. We first provide a characterization of the positive records of the two-sided infinite Mallows permutation.

Corollary 3.3. *Fix $q \in [0, 1)$. Let $\mathcal{X}_q$ be a two-sided infinite Mallows permutation and let $(\ldots, r_{-1}, r_0, r_1, \ldots)$ be its standard record representation. Then the distance between consecutive positive records $(\mathcal{X}_q(r_{i+1}) - \max\{\mathcal{X}_q(r_i), 0\})_{i \geq 0}$ is a sequence of independent* GEOMETRIC$^+(1 - q)$ *random variables.*

Proof. Thanks to Proposition 3.1, we know that the positive records of $\mathcal{X}_q$ are equal to the records of $X_{+,q}$. But now [ABC21, Proposition 1.11] states that the distances between records is distributed as a GEOMETRIC$^+(1 - q)$ and that the consecutive values and number of records is a Markov chain, thus proving the independence of the sequence. $\qquad \square$

While Corollary 3.3 characterizes the distribution of the jumps between positive records, proving a similar result for the negative proves to be much harder. Indeed, negative records are related to anti-records of $X_{-,q}$, and the sequence of anti-records of a Mallows permutation does not have independent geometric jumps. However, their jump distribution is also biased by the shift created by Λ and will thus end up being geometric, but it requires a different approach, which we now describe.

In order to fully characterize the distribution of the distance between consecutive jumps in a two-sided infinite Mallows permutation, we shift the permutation so that negative records become related to records of $X_{+,q}$. More precisely, we strongly rely on the following proposition.

Proposition 3.4. *Fix $q \in [0,1)$. Let $\mathcal{X}_q$ be the infinite Mallows permutation constructed from the triplet $(X_{+,q}, X_{-,q}, \Lambda)$, whose standard record representation is given by*

$$\left(\ldots, k_{1-\rho_{m+2}}, k_{1-\rho_{m+1}}, \underline{k_{1-\rho_m}}, \ell_{r_1}, \ell_{r_2}, \ldots\right),$$

as in Proposition 3.1. For any $n \in \mathbb{N}$, further write $\mathcal{X}_q^n$ for the permutation defined by $\mathcal{X}_q^n(k) = \mathcal{X}_q(n+k) - n$, for any $k \in \mathbb{Z}$. Then, $\mathcal{X}_q^n$ is distributed as a two-sided infinite Mallows permutation with parameter q. Moreover, by defining $s_n = \max\{i : X_{+,q}(r_i) \le n\}$, the standard record representation of $\mathcal{X}_q^n$ is given by

$$\left(\ldots, k_{1-\rho_{m+1}} - n, k_{1-\rho_m} - n, \ell_{r_1} - n, \ldots, \ell_{r_{s_n-1}} - n, \underline{\ell_{r_{s_n}} - n}, \ell_{r_{s_n+1}} - n, \ldots\right).$$

Proof. This results mostly follows from results of [GO12]. Indeed, by letting $t^n(k) = k + n$ be the permutation translating indices by n, we have that $\mathcal{X}_q^n = t^{-n} \cdot \mathcal{X}_q \cdot t^n$. But then, thanks to [GO12, Proposition 2.3] (the additive character of the balance) and [GO12, Theorem 3.3] (the uniqueness of the two-sided Mallows distribution), $\mathcal{X}_q^n$ is also a Mallows permutation. Finally, using the definition of records, we have that

$$
\begin{aligned}
r \text{ is a record of } \mathcal{X}_q &\iff \forall k < r : \mathcal{X}_q(k) < \mathcal{X}_q(r) \\
&\iff \forall k < r - n : \mathcal{X}_q(n+k) - n < \mathcal{X}_q(n+(r-n)) - n \\
&\iff \forall k < r - n : \mathcal{X}_q^n(k) < \mathcal{X}_q^n(r-n) \\
&\iff r - n \text{ is a record of } \mathcal{X}_q^n,
\end{aligned}
$$

thus explaining the desired sequence of records. To conclude the proof, simply note that, by definition of s_n, we have $X_{+,q}(r_{s_n}) \le n < X_{+,q}(r_{s_n+1})$, leading to

$$\mathcal{X}_q^n(\ell_{r_{s_n}} - n) = \mathcal{X}_q(\ell_{r_{s_n}}) - n = X_{+,q}(r_{s_n}) - n \le 0$$

and

$$\mathcal{X}_q^n(\ell_{r_{s_n+1}} - n) = \mathcal{X}_q(\ell_{r_{s_n+1}}) - n = X_{+,q}(r_{s_n} + 1) - n > 0,$$

thus proving the desired standard record representation. $\qquad\square$

We now conclude this first part of the proof by providing a useful equality in distributions for the records of the two-sided infinite Mallows permutation. The equality in distribution relies on the following transformation. Let $g = (g_i)_{i \in \mathbb{Z}}$ be a sequence of positive integers indexed by $\mathbb{Z}$. For any $n \ge 0$, write $s_n = \min\{i \ge 0 : g_0 + \ldots + g_i > n\}$ and let

$$(12) \qquad \Phi_n(g) = \left(\ldots, g_{-2}, g_{-1} + g_0 - 1, g_1, \ldots, g_{s_n-1}, n + 1 - \sum_{i=0}^{s_n-1} g_i, \underline{\sum_{i=0}^{s_n} g_i - n}, g_{s_n+1}, \ldots\right).$$

More precisely, $\Phi_n(g) = (\phi_i)_{i \in \mathbb{Z}}$ is defined by

$$
\phi_i = \begin{cases}
g_{s_n+i} & \text{if } i \ge 1 \\
\sum_{i=0}^{s_n} g_i - n & \text{if } i = 0 \\
n + 1 - \sum_{i=0}^{s_n-1} g_i & \text{if } i = -1 \\
g_{s_n+i+1} & \text{if } -s_n \le i \le -2 \\
g_0 + g_{-1} - 1 & \text{if } i = -s_n - 1 \\
g_{s_n+i} & \text{if } i \le -s_n - 2,
\end{cases}
$$

where we remove the fourth condition when $s_n = 1$ and combine the third and fifth terms to obtain $n + g_{-1}$ when $s_n = 0$.

Corollary 3.5. *Use the notations of Proposition 3.4 and the definition of Φ_n from (12). Write*

$$(\ldots, T_{-2}, T_{-1}, \underline{T_0}, T_1, T_2, \ldots)$$
$$= (\ldots, \mathcal{X}_q(k_{1-\rho_{m+2}}) - 1, \mathcal{X}_q(k_{1-\rho_{m+1}}) - 1, \mathcal{X}_q(k_{1-\rho_m}) - 1, \underline{0}, \mathcal{X}_q(\ell_{r_1}), \mathcal{X}_q(\ell_{r_2}), \ldots)$$

for the sequence of records of $\mathcal{X}_q$ where we add 0 at position 0 and shift all negative records by -1. Then, for any $n \in \mathbb{N}$, we have the equality in distribution

$$\Phi_n\big((T_{i+1} - T_i)_{i \in \mathbb{Z}}\big) \overset{d}{=} (T_{i+1} - T_i)_{i \in \mathbb{Z}}.$$

Proof. This result directly follows from the equality in distribution stated in Proposition 3.4. Indeed, the proposition implies that

$$(13) \qquad T_i \overset{d}{=} \begin{cases} \mathcal{X}_q^n(\ell_{r_{s_n+i}} - n) \\ 0 \\ \mathcal{X}_q^n(\ell_{r_{s_n+i+1}} - n) - 1 \\ \mathcal{X}_q^n(k_{1-\rho_{m-i-s_n-1}} - n) - 1 \end{cases} = \begin{cases} T_{s_n+i} - n & \text{if } i \geq 1 \\ 0 & \text{if } i = 0 \\ T_{s_n+i+1} - n - 1 & \text{if } -s_n \leq i \leq -1 \\ T_{s_n+i} - n & \text{if } i \leq -s_n - 1, \end{cases}$$

where the equality in distribution applies to the whole sequence of $(T_i)_{i \in \mathbb{Z}}$. This then implies that

$$T_{i+1} - T_i = \begin{cases} T_{s_n+i+1} - T_{s_n+i} & \text{if } i \geq 1 \\ T_{s_n+1} - n & \text{if } i = 0 \\ n + 1 - T_{s_n} & \text{if } i = -1 \\ T_{s_n+i+2} - T_{s_n+i+1} & \text{if } -s_n \leq i \leq -2 \\ T_1 - T_{-1} - 1 & \text{if } i = -s_n - 1 \\ T_{s_n+i+1} - T_{s_n+i} & \text{if } i \leq -s_n - 2, \end{cases}$$

and the right hand side is the exact definition of Φ_n (also in the case $s_n \in \{0, 1\}$). $\qquad\square$

In order to use this result, we now need to better understand what this equality in distribution means. We explore properties of sequences satisfying the equality in distribution from Corollary 3.5 below.

3.2. **Geometric random variables and point processes.** The main goal of this section is to use the result of Corollary 3.5 to understand the global distribution of records of a two-sided infinite Mallows permutation. In particular, we intend to prove the following theorem.

Theorem 3.6. *Let $G = (G_i)_{i \in \mathbb{Z}}$ be a sequence of random variables taking values in $\mathbb{Z}$ and define Φ_n as in (12). Then, the two following properties are equivalent.*

(P1) *The full sequence $G = (G_i)_{i \in \mathbb{Z}}$ is a family of independent $\textsc{Geometric}^+(p)$ random variables.*

(P2) *The sub-sequence $(G_i)_{i \geq 0}$ is a family of independent $\textsc{Geometric}^+(p)$ random variables and the whole sequence $G = (G_i)_{i \in \mathbb{Z}}$ satisfies the equality in distribution $\Phi_n(G) \overset{d}{=} G$ for any $n \geq 0$.*

In order to prove Theorem 3.6, we use the natural link between geometric random variables and point processes. In particular, we start by defining a point process which provides a sequence $(G_i)_{i \in \mathbb{Z}}$ satisfying both (P1) and (P2). We then use the existence of such sequence to prove that the two properties are equivalent.

Proof of Theorem 3.6. Fix $p \in (0, 1]$. Let $P \subseteq \mathbb{Z}$ be a random set of points obtained by independently keeping each $k \in \mathbb{Z}$ with probability p. Define now the random process $(N_k)_{k \in \mathbb{Z}}$ by $N_0 = 0$ and

$$N_k = \begin{cases} \big|P \cap [\![0, k-1]\!]\big| & \text{if } k \geq 1 \\ -\big|P \cap [\![k, -1]\!]\big| & \text{if } k \leq -1. \end{cases}$$

Note that $(N_k)_{k \geq 0}$ and $(-N_{-k})_{k \geq 0}$ independent, identically distributed, and are both jumping processes with increments of size 1 defined by

$$\mathbb{P}(N_{k+1} = N_k) = \mathbb{P}(N_{-k-1} = N_{-k}) = 1 - p,$$

for any $k \geq 0$. This implies that the distance between jumping times are geometrically distributed and independent of each other. More precisely, by defining $T_0 = 0$ and, for any $i \geq 1$, the stopping times by

$$T_i = \min\{k \in \mathbb{N} : N_k = i\}$$
$$T_{-i} = -\min\{k \in \mathbb{N} : N_{-k} = -i\},$$

Then $(T_{i+1} - T_i)_{i \in \mathbb{Z}}$ is a sequence of independent $\text{GEOMETRIC}^+(p)$ random variables. Note that T_i for $i \geq 1$ is a standard stopping time for $(N_k)_{k \in \mathbb{Z}}$, whereas T_{-i} is not since it is defined by

$$T_{-i} = \max\{k : N_k = -i\}$$

and so $\{T_{-i} \leq t\}$ belongs to the σ-algebra generated by $\{N_k : k \leq t + 1\}$ but not that of $\{N_k : k \leq t\}$. We now use this point process and the related variables to prove the statement of Theorem 3.6.

First, let $G_i = T_{i+1} - T_i$ so that $(G_i)_{i \in \mathbb{Z}}$ satisfies (P1). We now prove that it further satisfies (P2). Note that $G_0 + \ldots + G_i = T_{i+1}$. Moreover using the definition of s_n and the fact that that $\{T_i > k\} = \{N_k < i\}$ for $i, k \geq 0$, we obtain

$$s_n = \min\left\{i \in \mathbb{N} : T_{i+1} > n\right\} = \min\left\{i \in \mathbb{N} : N_n < i + 1\right\} = N_n.$$

Consider now the new process defined by $N^n = (N_k^n)_{k \in \mathbb{Z}} = (N_{n+k} - N_n)_{k \in \mathbb{Z}}$. First of all, note that $N_0^n = 0$. Moreover, for any $k \geq 1$, we have

$$N_k^n = \left|P \cap [\![0, n+k-1]\!]\right| - \left|P \cap [\![0, n-1]\!]\right| = \left|(n+P) \cap [\![0, k-1]\!]\right|.$$

Similarly, for any $-n \leq k \leq -1$, we have

$$N_k^n = \left|P \cap [\![0, n-|k|-1]\!]\right| - \left|P \cap [\![0, n-1]\!]\right| = -\left|(n+P) \cap [\![-|k|, -1]\!]\right| = -\left|(n+P) \cap [\![k, -1]\!]\right|.$$

Finally, for $k \leq -n-1$, we have

$$N_k^n = -\left|P \cap [\![n+k, -1]\!]\right| - \left|P \cap [\![0, n-1]\!]\right| = -\left|(n+P) \cap [\![k, -1]\!]\right|.$$

This means that N^n is defined in the same way as N but with $P^n = n+P$ instead of P. By translation invariance of the distribution of P, it follows that N^n is distributed as N.

To conclude the proof that $(G_i)_{i \in \mathbb{Z}}$ satisfies (P2), define $(T_i^n)_{i \in \mathbb{Z}}$ for N^n in the same way that $(T_i)_{i \in \mathbb{Z}}$ is defined with regards to N. Then, we verify that

$$T_i^n = \begin{cases} T_{i+N_n} - n & \text{if } i \geq 1 \\ 0 & \text{if } i = 0 \\ T_{i+N_n+1} - n - 1 & \text{if } -N_n \leq i \leq -1 \\ T_{i+N_n} - n & \text{if } i \leq -N_n - 1. \end{cases}$$

But this set of equality is exactly the same as in (13) from the proof of Corollary 3.5. This proves that $(G_i)_{i \in \mathbb{Z}}$ satisfies (P2).

To conclude the proof of Theorem 3.6, first note that (P1) fully characterizes the distribution of $(G_i)_{i \in \mathbb{Z}}$ and thus, thanks to the fact that our previous construction of $(G_i)_{i \in \mathbb{Z}}$ also satisfies (P2), we have proven that (P1) $\Rightarrow$ (P2). We now prove the converse result by showing that the property (P2) fully characterizes the distribution of $(G_i)_{i \in \mathbb{Z}}$. To avoid any confusion, we now write $(\tilde{G}_i)_{i \in \mathbb{Z}}$ for a general sequence of random variables satisfying (P2) and keep $(G_i)_{i \in \mathbb{Z}}$ for a sequence of independent $\text{GEOMETRIC}^+(p)$ (which thus satisfies (P1) and (P2)). We further extend the notations to define $\tilde{s}_n = \min\{i \geq 0 : \tilde{G}_0 + \ldots + \tilde{G}_i > n\}$.

Let $k \in \mathbb{Z}$ be a fixed integer and $g_{-k}, \ldots, g_k$ be a sequence of $2k+1$ integers. We want to prove that

$$\mathbb{P}\big(\tilde{G}_i = g_i, \forall |i| \leq k\big) = \mathbb{P}\big(G_i = g_i, \forall |i| \leq k\big).$$

To do so, write $\Phi_n^k(G) = (\phi_i)_{|i| \leq k}$ for the $2k+1$ terms of the sequence $\Phi_n(G) = (\phi_i)_{i \in \mathbb{Z}}$ (defined in (12)) around the central term. In particular, note that $\Phi_n^k(G)$ belongs to the σ-algebra generated

by $\{(G_i)_{i\geq 0}, (G_i)_{s_n - k \leq i \leq 0}\}$ and that s_n belongs to the σ-algebra generated by $(G_i)_{i\geq 0}$. We define in a similar way $\Phi_n^k(\tilde{G})$.

Using the fact that $(\tilde{G}_i)_{i\in\mathbb{Z}}$ satisfies (P2), for any $n \in \mathbb{N}$, we have that

$$\mathbb{P}\big(\tilde{G}_i = g_i, \forall |i| \leq k\big) = \mathbb{P}\big(\Phi_n^k(\tilde{G}) = (g_i)_{|i|\leq k}\big).$$

Now, due to the heavy tail of geometric random variables, we know that $\tilde{s}_n$ almost-surely diverges to infinity as n goes to infinity. This implies that, as n goes to infinity, we have

$$\mathbb{P}\big(\tilde{G}_i = g_i, \forall |i| \leq k\big) = \mathbb{P}\big(\Phi_n^k(\tilde{G}) = (g_i)_{|i|\leq k}, \tilde{s}_n \geq k\big) + o(1).$$

But now, thanks to the equality in distribution, we know that

$$\mathbb{P}\big(\Phi_n^k(\tilde{G}) = (g_i)_{|i|\leq k}, \tilde{s}_n \geq k\big) = \mathbb{P}\big(\Phi_n^k(G) = (g_i)_{|i|\leq k}, s_n \geq k\big).$$

Finally, using the same argument as that of $(\tilde{G}_i)_{i\in\mathbb{Z}}$ but on $(G_i)_{i\in\mathbb{Z}}$ instead, the previous equations imply that

$$\begin{aligned}
\mathbb{P}\big(\tilde{G}_i = g_i, \forall |i| \leq k\big) &= \mathbb{P}\big(\Phi_n^k(\tilde{G}) = (g_i)_{|i|\leq k}, \tilde{s}_n \geq k\big) + o(1) \\
&= \mathbb{P}\big(\Phi_n^k(G) = (g_i)_{|i|\leq k}, s_n \geq k\big) + o(1) \\
&= \mathbb{P}\big(G_i = g_i, \forall |i| \leq k\big) + o(1).
\end{aligned}$$

Since the two probabilities do not depend on n, they are equal. This proves that $(G_i)_{i\in\mathbb{Z}}$ and $(\tilde{G}_i)_{i\in\mathbb{Z}}$ have the same distribution and concludes the proof of Theorem 3.6. □

3.3. **Distribution of the two-sided infinite Mallows tree.** In this section, we finally prove Theorem 1.2 using the results of Sections 3.1 and 3.2 above.

Proof of Theorem 1.2. As mentioned in the introduction, the Mallows distribution (on any subset of $\mathbb{Z}$) is the only q-invariant distribution on permutations. Thus, for any finite subset $I \subseteq \mathbb{Z}$, by letting $\mathcal{X}_q\langle I \rangle$ be the unique permutation on $[|I|] = \{1, \ldots, |I|\}$ whose images are in the same order as that of $\mathcal{X}_q$ on $\mathcal{X}_q^{-1}(I)$, it follows that $\mathcal{X}_q\langle I \rangle$ is distributed as $\pi_{|I|,q}$ and is moreover independent of $\mathcal{X}_q$ on $\mathbb{Z} \setminus \mathcal{X}_q^{-1}(I)$. Let now $(R_i)_{i\in\mathbb{Z}}$ be the sequence of records whose indices coincides with that of the standard record representation of $\mathcal{X}_q$ and write $\mathcal{X}_q[i] = \mathcal{X}_q\langle [\![R_i + 1, R_{i+1} - 1]\!]\rangle$ for the induced permutations between consecutive records. Then, using the previous property of $\mathcal{X}_q$, conditionally given $(R_i)_{i\in\mathbb{Z}}$, the sequence $(\mathcal{X}_q^i)_{i\in\mathbb{Z}}$ is a family of independent Mallows permutations with parameter q, and their sizes are defined by the distances between consecutive records, but do not depend on the values of the records. Finally, recall the definition of redwood trees in (10) as

$$Y_q \overset{d}{=} Y\langle \mathcal{X}_q \rangle = \bigcup_{k\in\mathbb{Z}} \{\mathbf{1}^k\} \cup \mathbf{1}^k \mathbf{0} T\langle \mathcal{X}_q[k]\rangle$$

to see that it only remains to prove that the statement regarding the distribution of the sizes of the trees $(|\mathcal{X}_q(k)|)_{k\in\mathbb{Z}}$ from Theorem 1.2 is true. Indeed, the trees $(\mathcal{Y}_q(k))_{k\in\mathbb{Z}}$ exactly correspond to $(\mathbf{1}^k \mathbf{0} T\langle \mathcal{X}_q[k]\rangle)_{k\in\mathbb{Z}}$ and are thus, conditionally given their sizes, independent Mallows trees with parameter q.

Consider the distribution of $(|\mathcal{Y}_q(k)|)_{k\in\mathbb{Z}}$, which corresponds to that of $(R_{k+1} - R_k - 1)_{k\in\mathbb{Z}}$. First of all, define $(T_i)_{i\in\mathbb{Z}}$ as in Corollary 3.5, that is

$$T_i = \begin{cases} R_i & \text{if } i \geq 1 \\ 0 & \text{if } i = 0 \\ R_{i+1} - 1 & \text{if } i \leq -1. \end{cases}$$

Further let $G = (G_i)_{i\in\mathbb{Z}}$ be defined by $G_i = T_{i+1} - T_i$ for all $i \in \mathbb{Z}$. Thanks to Corollary 3.5, we know that $\Phi_n(G) = G$ for all $n \geq 0$, where Φ_n is defined in (12). Moreover, Corollary 3.3 states that $(G_i)_{i\geq 0}$ is a sequence of independent GEOMETRIC$^+(1-q)$, so G satisfies (P2) of Theorem 3.6 and thus also satisfies (P1): $G = (G_i)_{i\in\mathbb{Z}}$ is a sequence of independent GEOMETRIC$^+(1-q)$ random variables.

To conclude the proof of the theorem, observe that

$$R_{i+1} - R_i - 1 = \begin{cases} G_i - 1 & \text{if } i \geq 1 \\ G_0 + G_{-1} - 2 & \text{if } i = 0 \\ G_{i-1} - 1 & \text{if } i \leq -1 \end{cases}$$

to first see that the sizes of the subtrees $(\mathcal{Y}_q(k))_{k \in \mathbb{Z}}$ are all independent and their distribution is GEOMETRIC$(1-q)$ when $k \neq 0$. Finally, the distribution of $\mathcal{Y}_q(0)$ is equal to the sum of two independent GEOMETRIC$(1-q)$ and is thus given by

$$(14) \quad \mathbb{P}(\mathcal{Y}_q(0) = n) = \mathbb{P}(G_0 + G_{-1} = 2 + n) = \sum_{g=1}^{1+n} q^{g-1}(1-q)q^{2+n-g-1}(1-q) = (n+1)q^n(1-q)^2,$$

which happens to be exactly that of GEOMETRIC$^\star(1-q)$ as defined in (3), thus concluding the proof of the theorem. $\square$

3.4. **Local limit of Mallows trees.** We conclude this section with the proof of Theorem 1.1, showing that finite Mallows trees almost-surely locally converge to the redwood tree of the two-sided infinite Mallows permutation. We separate the proof into three sub-propositions and an intermediate lemma.

For the rest of this section we fix an integer $r \geq 1$ and a rooted graph H. We also assume without loss of generality that $(X_{n,q})_{n \in \mathbb{N} \cup \{\infty\}}$ is coupled according to the infinite Bernoulli model, as described in Section 2.1.1, and we write $T_{n,q} = T\langle X_{n,q} \rangle$ for $n \in \mathbb{N} \cup \{\infty\}$. The goal of this section is thus to prove that $T_{n,q}$ satisfies (11) with $\mathcal{Y}_q$, that is

$$(15) \quad \frac{1}{n} \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v, T_{n,q}) \equiv H\}} \xrightarrow{\text{a.s.}} \mathbb{P}(B_r(\mathcal{Y}_q) \equiv H).$$

We also recall that $\mathcal{Y}_q$ is defined as (o, Y_q), where $Y_q = Y\langle \mathcal{X}_q \rangle$ is the redwood tree of a two-sided infinite Mallows permutation $\mathcal{X}_q$ and o is a uniformly sampled vertex of $Y_q \cap \mathbb{T}_0$, where $\mathbb{T}_0 = \{\varnothing\} \cup \mathbf{0}\mathbb{T}$. By writing $Y_q(k) = Y_q \cap \mathbf{1}^k \mathbf{0}\mathbb{T}$ as in Theorem 1.2, this implies that

$$(16) \quad \mathbb{P}(B_r(\mathcal{Y}_q) \equiv H) = \mathbb{P}\left(B_r \left(o, \bigcup_{|k| \leq r} \{\mathbf{1}^k\} \cup Y_q(k) \right) \equiv H \right),$$

which will prove to be a useful characterization of the target limit. We also extend the notation $Y_q(k)$ to $T_{\infty,q}$ by letting $T_{\infty,q}(k) = T_{\infty,q} \cap \mathbf{1}^k \mathbf{0}\mathbb{T}$ for any $k \geq 0$. We mention here that the sequence $(T_{\infty,q}(k))_{k \geq 0}$ is known to be a sequence of independent Mallows trees with GEOMETRIC$(1-q)$ sizes (see [ABC21, Lemma 1.10]).

We now state and prove the first proposition, which allows us to changed the left-hand side in (15) by an object with less inter-dependency.

Proposition 3.7. *Using the notations introduced at the beginning of Section 3.4, we have that*

$$\frac{1}{n} \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v, T_{n,q}) \equiv H\}} - \frac{1}{n} \sum_{k=0}^{\lfloor n(1-q) \rfloor} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} \xrightarrow{\text{a.s.}} 0.$$

Proof. Recall the definition of $R_{n,q}$ from (5) and $F_{n,q}$ from (6), leading to (7):

$$\left| \left\{ 0 \leq k \leq R_{n,q} : T\langle X_{\infty,q} \rangle \cap \mathbf{1}^k \mathbb{T}_0 \neq T\langle X_{n,q} \rangle \cap \mathbf{1}^k \mathbb{T}_0 \right\} \right| \leq F_{n,q}.$$

Further let

$$T_{\infty,q}^n = T_{\infty,q} \cap \bigcup_{0 \leq k \leq R_{n,q}} \mathbf{1}^k \mathbb{T}_0$$

be the infinite tree $T_{\infty,q}$ cropped at depth $R_{n,q}$ on the rightmost path, or equivalently the tree $T_{n,q}$ where we fill in the left subtrees as they appear in their limit $T_{\infty,q}$. We recall here that, by definition, $F_{n,q} = |T_{\infty,q}^n \setminus T_{n,q}|$.

For any binary search tree T and $v \in T \cap \mathbf{1}^k \mathbb{T_0}$, we see that $B_r(v, T)$ has size at most $3 \cdot 2^r$ and so we obtain

$$\left| \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v, T_{n,q}) \equiv H\}} - \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} \right| \leq (3 \cdot 2^r) F_{n,q} \,.$$

Using the definition of $F_{n,q}$, we also see that

$$\sum_{v \in T_{\infty,q}^n} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} - \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} = \sum_{v \in T_{\infty,q}^n \setminus T_{n,q}} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} \leq F_{n,q} \,,$$

Combining the previous results leads to

$$\left| \frac{1}{n} \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v, T_{n,q}) \equiv H\}} - \frac{1}{n} \sum_{v \in T_{\infty,q}^n} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} \right|$$

$$\leq \left| \frac{1}{n} \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v, T_{n,q}) \equiv H\}} - \frac{1}{n} \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} \right| + \frac{1}{n} \sum_{v \in T_{\infty,q}^n \setminus T_{n,q}} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}}$$

$$\leq \frac{(3 \cdot 2^r + 1) F_{n,q}}{n} \,,$$

But now, thanks to Lemma 2.2 and since r is fixed, this implies that

$$(17) \qquad \frac{1}{n} \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v, T_{n,q}) \equiv H\}} - \frac{1}{n} \sum_{v \in T_{\infty,q}^n} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} \xrightarrow{\text{a.s.}} 0 \,.$$

Then, using the definition of $T_{\infty,q}^n$, we observe that

$$\sum_{v \in T_{\infty,q}^n} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} = \sum_{k=0}^{R_{n,q}} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} \,.$$

which is almost the desired result, except that we need to replace $R_{n,q}$ with $n(1-q)$.

Using Lemma 2.1, we know that, almost-surely, $|R_{n,q} - \lfloor n(1-q) \rfloor| \leq \sqrt{n} \log n$ for n large enough. Moreover, using the strong law of large number along with the fact that $(T_{\infty,q}(k))_{k \geq 0}$ are independent Mallows trees of geometric sizes, we have that

$$\frac{1}{2\sqrt{n} \log n} \sum_{k : |k - \lfloor n(1-q) \rfloor| \leq \sqrt{n} \log n} \left(1 + |T_{\infty,q}(k)| \right) \xrightarrow{\text{a.s.}} \frac{1}{1-q} \,,$$

where the limit is simply the expected value of a GEOMETRIC$^+(1-q)$ random variable. Now, for n large enough so that $|R_{n,q} - \lfloor n(1-q) \rfloor| \leq \sqrt{n} \log n$ almost-surely, we have that

$$\left| \frac{1}{n} \sum_{v \in T_{\infty,q}^n} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} - \frac{1}{n} \sum_{k=0}^{\lfloor n(1-q) \rfloor} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} \right|$$

$$= \frac{1}{n} \left| \sum_{k=0}^{R_{n,q}} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} - \sum_{k=0}^{\lfloor n(1-q) \rfloor} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v, T_{\infty,q}) \equiv H\}} \right|$$

$$\leq \frac{2\sqrt{n} \log n}{n} \cdot \left[\frac{1}{2\sqrt{n} \log n} \sum_{k : |k - \lfloor n(1-q) \rfloor| \leq \sqrt{n} \log n} \left(1 + |T_{\infty,q}(k)| \right) \right] \,.$$

If follows that

$$\frac{1}{n} \sum_{v \in T_{\infty,q}^n} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} - \frac{1}{n} \sum_{k=0}^{\lfloor n(1-q) \rfloor} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} \xrightarrow{\text{a.s.}} 0,$$

which combined with (17) provides us with the desired almost-sure convergence. $\qquad \square$

The previous proposition provides an equivalent sequence of random variables to consider in order to prove that (15) holds. We now state and prove the second proposition, showing the existence and a first characterization of the limit of the desired sequence of random variables.

Proposition 3.8. *Using the notations introduced at the beginning of Section 3.4, we have that*

$$\frac{1}{n} \sum_{k=0}^{\lfloor n(1-q) \rfloor} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} \xrightarrow{\text{a.s.}} (1-q)L,$$

where

(18) $$L = \mathbb{E}\left[\sum_{v \in \{\mathbf{1}^r\} \cup T_{\infty,q}(r)} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} \right].$$

Proof. Consider some $k \geq 0$ and let $v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)$. Since $B_r(v,T_{\infty,q})$ is the ball of radius r around v, it only depends on $(T_{\infty,q}(\ell))_{|\ell-k| \leq r}$. Moreover, if $k \geq r$, then $(T_{\infty,q}(\ell))_{|\ell-k| \geq r}$ is distributed as $(T_{\infty,q}(\ell))_{0 \leq \ell \leq 2r}$ and thus, the family of random variables

(19) $$\left(\sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} \right)_{k \geq r}$$

are identically distributed and their expectation is given by L as in (18). They are however not independent (since $r \geq 1$), but this is actually only a minor problem, since their dependency only has finite range. Indeed, for any k and k' such that $|k - k'| \geq 2r$, the random variables of (19) with indices k and k' are independent. Note that, in the case $|k - k'| = 2r$, of all the pairs of balls of radius r considered in the two random variables, exactly one of them intersects, when $v = \mathbf{1}^k$ and $v' = \mathbf{1}^{k'}$, and the intersection is $B_r(\mathbf{1}^k, T_{\infty,q}) \cap B_r(\mathbf{1}^{k'}, T_{\infty,q}) = \{\mathbf{1}^{(k+k')/2}\}$; however, since this intersection is deterministic, the two random variables are still independent.

Consider now any $r \leq k < 3r$. Thanks to the previous observation of independence and identical distribution, by applying the strong law of large numbers, we know that

$$\frac{1}{1 + \left\lfloor \frac{n(1-q)-k}{2r} \right\rfloor} \sum_{\{i: k+2ri \leq n(1-q)\}} \sum_{\{v \in \{\mathbf{1}^{k+2ri}\} \cup T_{\infty,q}(k+2ri)\}} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} \xrightarrow{\text{a.s.}} L,$$

Since this is true for any $r \leq k < 3r$ and r is fixed and finite, it follows that the sum of all the previous random variables converges almost-surely to $2rL$. Using that $\frac{1}{1 + \left\lfloor \frac{n(1-q)-k}{2r} \right\rfloor} \sim \frac{2r}{n(1-q)}$, this implies that

$$\sum_{k=r}^{3r-1} \frac{2r}{n(1-q)} \sum_{\{i: k+2ri \leq n(1-q)\}} \sum_{\{v \in \{\mathbf{1}^{k+2ri}\} \cup T_{\infty,q}(k+2ri)\}} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} \xrightarrow{\text{a.s.}} 2rL$$

and so

$$\frac{1}{n} \sum_{k=r}^{\lfloor n(1-q) \rfloor} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} \xrightarrow{\text{a.s.}} (1-q)L.$$

To conclude the proof of the proposition, use that the first r subtrees of $T_{\infty,q}$ are almost-surely finite to extend the previous sum to $0 \leq k \leq n(1-q)$ and obtain the desired almost-sure convergence. $\qquad \square$

With this second proposition, we obtain that the equivalent sequence of random variables from Proposition 3.7 admits a limit. It now only remains to prove that the limit is as desired. Before doing so, we provide a lemma which will be useful in proving the final proposition.

Lemma 3.9. *Let $(S_i)_{i\geq 0}$ be a sequence of disjoint sets such that $(|S_i|)_{i\geq 0}$ is a sequence of independent random variables distributed according to $\pi = (\pi_\ell)_{\ell \geq 1}$. Write $\mu = \sum_{\ell \geq 1} \ell \pi_\ell$ for the mean of π. Fix an integer $r \geq 1$ and a sequence of integers $(k_n)_{n \in \mathbb{N}}$ such that $k_n \to \infty$ as n goes to infinity. Finally, let U_n be a uniform random variable on the set $\bigcup_{i=0}^{k_n} S_i$. Then, for any $\ell_0, \ldots, \ell_r$, we have*

$$\lim_{n\to\infty} \mathbb{P}\Big(|S_i| = \ell_i, \forall i \leq r \;\Big|\; U_n \in S_0\Big) = \frac{\ell_0 \pi_{\ell_0}}{\mu} \prod_{i=1}^{r} \pi_{\ell_i} \, .$$

In words, conditionally given that $U_n \in S_0$, the random variables $(S_0, \ldots, S_r)$ are independent with S_0 being distributed as the size-biased version of π and the other S_i being distributed according to π.

Proof. Using the rules of conditional probability, we know that

$$\mathbb{P}\Big(|S_i| = \ell_i, \forall i \leq r \;\Big|\; U_n \in S_0\Big) = \frac{\mathbb{P}\Big(|S_i| = \ell_i, \forall i \leq r\Big)}{\mathbb{P}\Big(U_n \in S_0\Big)} \mathbb{P}\Big(U_n \in S_0 \;\Big|\; |S_i| = \ell_i, \forall i \leq r\Big) \, .$$

Since the $(S_i)_{0 \leq i \leq k_n}$ are translation invariant, we know that

$$\mathbb{P}\Big(U_n \in S_0\Big) = \frac{1}{k_n + 1} \, ,$$

which implies that

$$\mathbb{P}\Big(|S_i| = \ell_i, \forall i \leq r \;\Big|\; U_n \in S_0\Big) = \left((k_n + 1)\prod_{i=0}^{r} \pi_{\ell_i}\right) \cdot \mathbb{P}\Big(U_n \in S_0 \;\Big|\; |S_i| = \ell_i, \forall i \leq r\Big) \, .$$

Now, conditioning on the values of the other S_i (assuming that n is large enough so that $k_n \geq r$), we have that

$$\mathbb{P}\Big(U_n \in S_0 \;\Big|\; |S_i| = \ell_i, \forall i \leq r\Big) = \mathbb{E}\left[\mathbb{P}\Big(U_n \in S_0 \;\Big|\; |S_i| = \ell_i, \forall i \leq r, \big(|S_i|\big)_{r < i \leq k_n}\Big)\right]$$

$$= \mathbb{E}\left[\frac{\ell_0}{\sum_{i=0}^{r} \ell_i + \sum_{i=r+1}^{k_n} |S_i|}\right] \, ,$$

which, once plugged back into the previous equality, leads to

$$\mathbb{P}\Big(|S_i| = \ell_i, \forall i \leq r \;\Big|\; U_n \in S_0\Big) = \ell_0 \prod_{i=0}^{r} \pi_{\ell_i} \cdot \mathbb{E}\left[\frac{k_n + 1}{\sum_{i=0}^{r} \ell_i + \sum_{i=r+1}^{k_n} |S_i|}\right] \, .$$

Finally, using the fact that $k_n \to \infty$ along with the strong law of large numbers and the dominated convergence theorem (using that $|S_i| \geq 1$), we obtain that

$$\lim_{n\to\infty} \mathbb{E}\left[\frac{k_n + 1}{\sum_{i=0}^{r} \ell_i + \sum_{i=r+1}^{k_n} |S_i|}\right] = \frac{1}{\mu} \, ,$$

which provides the desired limit. □

We finally prove the third proposition used to prove Theorem 1.1, which characterizes the limit obtained in Proposition 3.8.

Proposition 3.10. *Using the notations introduced at the beginning of the section and L defined in (18), we have that*

$$(1 - q)L = \mathbb{P}\big(B_r(\mathcal{Y}_q) \equiv H\big) \, .$$

Proof. Using the definition of L from (18) and its relation with the family of random variables given in (19), we have that

$$(1-q)L = (1-q)\mathbb{E}\left[\sum_{v\in\{\mathbf{1}^r\}\cup T_{\infty,q}(r)} \mathbb{1}_{\{B_r(v,T_{\infty,q})\equiv H\}}\right]$$

$$= (1-q)\mathbb{E}\left[\frac{1}{\lfloor n(1-q)\rfloor + 1 - r}\sum_{k=r}^{\lfloor n(1-q)\rfloor}\sum_{v\in\{\mathbf{1}^k\}\cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v,T_{\infty,q})\equiv H\}}\right].$$

For more clarity in the rest of this proof, let

$$T_{n,q}^L = \bigcup_{k=r}^{\lfloor n(1-q)\rfloor} \{\mathbf{1}^k\}\cup T_{\infty,q}(k)$$

be the tree which will help us characterize $(1-q)L$; in particular, $T_{n,q}^L$ corresponds to $T_{\infty,q}^n$ defined in the proof of Proposition 3.7 where we remove the first r left subtrees. Then, combining the two previous equations, we have

$$(1-q)L = (1-q)\mathbb{E}\left[\frac{1}{\lfloor n(1-q)\rfloor + 1 - r}\sum_{v\in T_{n,q}^L} \mathbb{1}_{\{B_r(v,T_{\infty,q})\equiv H\}}\right]$$

$$= \frac{n(1-q)}{\lfloor n(1-q)\rfloor + 1 - r}\mathbb{E}\left[\frac{|T_{n,q}^L|}{n}\cdot\frac{1}{|T_{n,q}^L|}\sum_{v\in T_{n,q}^L} \mathbb{1}_{\{B_r(v,T_{\infty,q})\equiv H\}}\right].$$

We note that the term in front of the expectation converges to 1 as n goes to infinity. Moreover, $|T_{n,q}^L|$ is a sum of $\lfloor n(1-q)\rfloor - r + 1$ independent GEOMETRIC$^+(1-q)$ random variables and so it is tight and satisfies the strong law of large numbers: $|T_{n,q}^L|/n \xrightarrow{\text{a.s.}} 1$. Finally, note that

$$\mathbb{E}\left[\frac{1}{|T_{n,q}^L|}\sum_{v\in T_{n,q}^L} \mathbb{1}_{\{B_r(v,T_{\infty,q})\equiv H\}}\right] = \mathbb{P}\left(B_r(V_n,T_{\infty,q})\equiv H\right),$$

where V_n is a uniformly sampled vertex from $T_{n,q}^L$. So now, if we manage the prove that the right-hand side of the previous equation admits a limit as n goes to infinity, by the dominated convergence theorem we can conclude that

$$(1-q)L = \lim_{n\to\infty}\mathbb{P}\left(B_r(V_n,T_{\infty,q})\equiv H\right).$$

We now show that the limit indeed exists and corresponds to the desired value.

Using that the subtrees $(T_{\infty,q}(k))_{k\geq 0}$ are independent and identically distributed, we have

$$\mathbb{P}\left(B_r(V_n,T_{\infty,q})\equiv H\right) = \sum_{k=r}^{\lfloor n(1-q)\rfloor}\mathbb{P}\left(B_r(V_n,T_{\infty,q})\equiv H \mid V_n\in T_{n,q}^L\cap\mathbf{1}^k\mathbb{T_0}\right)\mathbb{P}\left(V_n\in T_{n,q}^L\cap\mathbf{1}^k\mathbb{T_0}\right)$$

$$= \mathbb{P}\left(B_r(V_n,T_{\infty,q})\equiv H \mid V_n\in T_{n,q}^L\cap\mathbf{1}^r\mathbb{T_0}\right)$$

$$= \mathbb{P}\left(B_r(V_n,T_{\infty,q})\equiv H \mid V_n\in\{\mathbf{1}^k\}\cup T_{\infty,q}(\mathbf{1}^r)\right).$$

The event $V_n\in\{\mathbf{1}^k\}\cup T_{\infty,q}(\mathbf{1}^r)$ only depends on the size of $|T_{\infty,q}(\mathbf{1}^r)|$ in comparison to the other left subtrees $(|T_{\infty,q}(k)|)_{r<k\leq n(1-q)}$ and is independent of their structure as well as that of the other left subtrees of $T_{\infty,q}$. Thus, Lemma 3.9 tells us that, as we make n tends to infinity, the distribution of $(|T_{\infty,q}(k)|)_{0\leq k\leq 2r}$ conditionally given that $V_n\in\{\mathbf{1}^k\}\cup T_{\infty,q}(\mathbf{1}^r)$ converges to a sequence of independent

random variables, all distributed according to a GEOMETRIC$(1-p)$, except for $|T_{\infty,q}(r)|$, for which the formula applied with $\pi_\ell = q^{\ell-1}(1-q)$ and $\mu = 1/(1-q)$ gives us

$$\mathbb{P}\Big(|T_{\infty,q}(r)| = \ell \mid V_n \in \{\mathbf{1}^k\} \cup T_{\infty,q}(\mathbf{1}^r)\Big) \longrightarrow \frac{(\ell+1)q^\ell(1-q)}{1/(1-q)} = (\ell+1)q^\ell(1-q)\,;$$

this is exactly the distribution of a GEOMETRIC$^\star(1-q)$ as given in (3).

Putting everything together, the distribution of the subtrees $(T_{\infty,q}(k))_{0 \le k \le 2r}$ conditionally given that $V_n \in \{\mathbf{1}^k\} \cup T_{\infty,q}(\mathbf{1}^r)$ asymptotically converges in distribution to a set of $(2r+1)$ independent Mallows trees of geometric sizes, except for the r-th tree whose distribution is size-biased. But now, this is exactly the distribution of $(Y_q(k))_{|k| \le r}$ (by shifting the indices), meaning that

$$\lim_{n \to \infty} \mathbb{P}\Big(B_r(V_n, T_{\infty,q}) \equiv H\Big) = \mathbb{P}\big(B_r(\mathcal{Y}_q) \equiv H\big)\,,$$

where we used (16) to prove that the right-hand side is as desired. This concludes the proof of the proposition. $\square$

With the third proposition proven, we now combine all previous results to prove Theorem 1.1.

Proof of Theorem 1.1. We aim to prove that (15) holds, that is

$$\frac{1}{n} \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v,T_{n,q}) \equiv H\}} \xrightarrow{\text{a.s.}} \mathbb{P}\big(B_r(\mathcal{Y}_q) \equiv H\big)\,.$$

To do so, we first combine Proposition 3.8 with Proposition 3.10 to obtain that a modified sequence converges as desired:

$$\frac{1}{n} \sum_{k=0}^{\lfloor n(1-q) \rfloor} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} \xrightarrow{\text{a.s.}} \mathbb{P}\big(B_r(\mathcal{Y}_q) \equiv H\big)\,.$$

But now, thanks to Proposition 3.7, we have that

$$\frac{1}{n} \sum_{v \in T_{n,q}} \mathbb{1}_{\{B_r(v,T_{n,q}) \equiv H\}} - \frac{1}{n} \sum_{k=0}^{\lfloor n(1-q) \rfloor} \sum_{v \in \{\mathbf{1}^k\} \cup T_{\infty,q}(k)} \mathbb{1}_{\{B_r(v,T_{\infty,q}) \equiv H\}} \xrightarrow{\text{a.s.}} 0\,.$$

The proof of the theorem simply follows from combining the two previous convergence results. $\square$

4. THE OTHER TYPES OF CONVERGENCE

After proving the local limit of Mallows trees in Section 3, we now focus on three other types of convergence and prove Theorem 1.3, 1.4, and 1.5 respectively in Section 4.1, 4.2, and 4.3 below. For the rest of this section, and without loss of generality, we assume that $(X_{n,q})_{n \in \mathbb{N} \cup \{\infty\}}$ are coupled according to the infinite Bernoulli model defined in Section 2.1.1.

4.1. **Rooted limit of Mallows trees.** The aim of this section is to prove Theorem 1.3, which corresponds to the limit of Mallows trees according to the rooted topology defined in Section 2.3.1. This result straightforwardly follows from the coupling of the infinite Bernoulli model as defined in Section 2.1.1.

Proof of Theorem 1.3. From the definition of $X_{\infty,q}$ as the limit of $(X_{n,q})_{n \in \mathbb{N}}$, we know that, for any $r \ge 0$, the subtree of $T\langle X_{n,q} \rangle$ rooted at $\mathbf{1}^r \mathbf{0}$ converges to that of $T\langle X_{\infty,q} \rangle$. More precisely, the sequence $(T\langle X_{n,q} \rangle \cap \mathbf{1}^r \mathbf{0} \mathbb{T})_{n \in \mathbb{N}}$ is increasing and satisfies

$$\bigcup_{n \in \mathbb{N}} \Big(T\langle X_{n,q} \rangle \cap \mathbf{1}^r \mathbf{0} \mathbb{T}\Big) = T\langle X_{\infty,q} \rangle \cap \mathbf{1}^r \mathbf{0} \mathbb{T}\,,$$

which also implies that

$$\bigcup_{n \in \mathbb{N}} \Big(T\langle X_{n,q} \rangle \cap \mathbf{1}^r \mathbb{T}_\mathbf{0}\Big) = T\langle X_{\infty,q} \rangle \cap \mathbf{1}^r \mathbb{T}_\mathbf{0}\,,$$

Since $T\langle X_{\infty,q}\rangle \cap \mathbf{1}^r\mathbb{T}_\mathbf{0}$ is finite, there exists $N_0(r) < \infty$ such that, for all $n \geq N_0(r)$, we have

$$T\langle X_{n,q}\rangle \cap \mathbf{1}^r\mathbb{T}_\mathbf{0} = T\langle X_{\infty,q}\rangle \cap \mathbf{1}^r\mathbb{T}_\mathbf{0}\,.$$

But now, by letting $N(r) = \max\{N_0(k) : 0 \leq k \leq r\}$, we have that $N(r) < \infty$ almost-surely, and for all $n \geq N(r)$

$$\bigcup_{k=0}^{r} T\langle X_{n,q}\rangle \cap \mathbf{1}^k\mathbb{T}_\mathbf{0} = \bigcup_{k=0}^{r} T\langle X_{\infty,q}\rangle \cap \mathbf{1}^k\mathbb{T}_\mathbf{0}\,.$$

In particular, for any $n \geq N(r)$, we also have

$$B_r\big(\varnothing, T\langle X_{n,q}\rangle\big) = B_r\big(\varnothing, T\langle X_{\infty,q}\rangle\big)$$

and so, for any rooted tree t, it follows that

$$\lim_{n\to\infty} \mathbb{P}\Big(B_r\big(\varnothing, T\langle X_{n,q}\rangle\big) \equiv t\Big) = \mathbb{P}\Big(B_r\big(\varnothing, T\langle X_{\infty,q}\rangle\big) \equiv t\Big),$$

which is exactly the definition of the rooted convergence. $\qquad\square$

It is interesting to note that, in the case of the aforementioned coupling, we actually have that

$$\mathbb{1}_{\{B_r(\varnothing, T\langle X_{n,q}\rangle)\equiv t\}} \xrightarrow{\text{a.s.}} \mathbb{1}_{\{B_r(\varnothing, T\langle X_{\infty,q}\rangle)\equiv t\}}\,,$$

which could be seen as an *almost-sure rooted convergence*. However, the almost-sure rooted convergence simply means that, for n large enough, we almost-surely have that

$$B_r(\varnothing, T\langle X_{n,q}\rangle) = B_r(\varnothing, T\langle X_{\infty,q}\rangle)\,,$$

which was the case when considering the coupling of the infinite Bernoulli model but is not necessarily an interesting topology in general.

4.2. **Strong Gromov-Hausdorff-Prokhorov limit of Mallows trees.** The aim of this section is to prove Theorem 1.4, which corresponds to the limit of Mallows trees according to the strong Gromov-Hausdorff-Prokhorov (GHP) topology. We start by providing a quick definition of this topology (we refer to [Mie09, Section 6] for further details), which we then use to provide a simpler condition sufficient to prove the limit from Theorem 1.4. We conclude this section with the proof of the theorem.

A *weighted metric space* is a triplet (E, d, μ) where (E, d) is a compact metric space and (E, μ) is a probability space. For a pair of weighted metric spaces $\mathcal{E}_1 = (E_1, d_1, \mu_1)$ and $\mathcal{E}_2 = (E_2, d_2, \mu_2)$, we define the GHP distance as follows. First, a *correspondence* between E_1 and E_2 is a set $C \subseteq E_1 \times E_2$ such that $\{x_1 : \exists(x_1, x_2) \in C\} = E_1$ and $\{x_2 : \exists(x_1, x_2) \in C\} = E_2$ and the *distortion* of C is defined by

$$\text{dis}(C) := \sup\left\{\big|d_1(x_1, x_1') - d_2(x_2, x_2')\big| : (x_1, x_2), (x_1', x_2') \in C\right\}.$$

Second, a *coupling* between μ_1 and μ_2 is a probability measure M on $E_1 \times E_2$ such that the marginal measures of M on E_1 and E_2 are respectively μ_1 and μ_2. Now, the GHP distance on the pair $\mathcal{E}_1$ and $\mathcal{E}_2$ is defined by

$$d_{\text{GHP}}(\mathcal{E}_1, \mathcal{E}_2) := \inf\left\{\epsilon > 0 : \exists(C, M), M(C) \geq 1 - \epsilon, \text{dis}(C) \leq 2\epsilon\right\},$$

where the pair (C, M) is composed of a correspondence C between E_1 and E_2 and coupling M between μ_1 and μ_2.

While to transform d_{GHP} into a metric on a proper complete separable metric space we need to consider weighted metric spaces up to isomorphism, the definition of d_{GHP} on any pair of weighted metric space is enough for the sake of this study. Thus, given a sequence of random weighted metric spaces $(\mathcal{E}_n)_{n\in\mathbb{N}}$ and a metric space $\mathcal{E}$, we say that it converges according to the strong GHP topology, and write it $\mathcal{E}_n \xrightarrow{\text{GHP}} \mathcal{E}$ if

$$d_{\text{GHP}}(\mathcal{E}_n, \mathcal{E}) \xrightarrow{\text{a.s.}} 0\,.$$

We call this convergence *strong* in comparison to a weaker one where the convergence would occur in probability (or equivalently in distribution, since the limit is constant). We now provide a useful condition to check whether $(\mathcal{E}_n)_{n \in \mathbb{N}}$ converges to the limit in Theorem 1.4, that is $\mathcal{E}_u = ([0,1], d, \mu_{[0,1]})$ where $d(x,y) = |x-y|$ and $\mu_{[0,1]}$ is the uniform measure on $[0,1]$.

Lemma 4.1. *Let $(\mathcal{E}_n)_{n \in \mathbb{N}} = ((E_n, d_n, \mu_n))_{n \in \mathbb{N}}$ be a sequence of metric spaces such that, for all $n \in \mathbb{N}$, E_n is a finite set of size n and μ_n is the uniform measure on E_n. Assume that, for all $n \in \mathbb{N}$, there exists an ordering of the elements of $E_n = \{x_{i,n} : 1 \leq i \leq n\}$ such that*

$$(20) \qquad \max\left\{ \left| d_n(x_{i,n}, x_{j,n}) - \frac{|i-j|}{n} \right| : 1 \leq i, j \leq n \right\} \xrightarrow{\text{a.s.}} 0 .$$

Then we have the convergence according to the strong GHP topology

$$\mathcal{E}_n \xrightarrow{\text{GHP}} \mathcal{E}_u$$

Proof. Using the definition of the strong GHP convergence, we see that it suffices to find a sequence $(C_n)_{n \in \mathbb{N}}$ of correspondences and a sequence $(M_n)_{n \geq 1}$ of matching such that, C_n is a correspondence between E_n and $[0,1]$ satisfying

$$\text{dis}(C_n) \xrightarrow{\text{a.s.}} 0$$

and M_n is a matching between μ_n and $\mu_{[0,1]}$ satisfying

$$M_n(C_n) \xrightarrow{\text{a.s.}} 1 .$$

To do so we rely on the ordering of E_n provided as assumption of the lemma.

Consider first the correspondence defined by

$$C_n = \left\{ \left(x_{i,n}, \left[\frac{i-1}{n}, \frac{i}{n} \right] \right) : 1 \leq i \leq n \right\} .$$

It is indeed a correspondence between E_n and $[0,1]$ since $E_n = \{x_{i,n} : 1 \leq i \leq n\}$ and $[0,1] = \bigcup_{i=1}^n [(i-1)/n, i/n]$. Now, C_n is a union of n disjoint segments of $E_n \times [0,1]$, so choose M_n to be the uniform probability measure on C_n. By the definition of C_n, M_n is indeed a matching between μ_n and $\mu_{[0,1]}$ since $M_n(\{x_{i,n}\} \times [0,1]) = M_n(\{x_{i,n} \times [(i-1)/n, i/n]\}) = \frac{1}{n}$ and, for any Lebesgue measurable set A of $[0,1]$,

$$M_n(E_n \times A) = \sum_{i=1}^n M_n(\{x_{i,n}\} \times (A \cap [(i-1)/n, i/n]))$$

$$= \sum_{i=1}^n \mu_{[0,1]} \big(A \cap [(i-1)/n, i/n] \big)$$

$$= \mu_{[0,1]}(A) ,$$

where the last equality uses that the union of the singletons $\{1/n, 2/n, \ldots, (n-1)/n\}$ has $\mu_{[0,1]}$-measure 0. This proves that $(C_n)_{n \in \mathbb{N}}$ and $(M_n)_{n \in \mathbb{N}}$ are correspondences and matching as desired. Moreover, by definition $M_n(C_n) = 1$, so it only remains to prove that $\text{dis}(C_n) \xrightarrow{\text{a.s.}} 0$. This directly follows from the assumption of the lemma since

$$\text{dis}(C_n) = \sup\left\{ \left| d_n(x_{i,n}, x_{j,n}) - \frac{|x-y|}{n} \right| : 1 \leq i, j \leq n, i-1 \leq x \leq i, j-1 \leq y \leq j \right\}$$

$$\leq \frac{2}{n} + \max\left\{ \left| d_n(x_{i,n}, x_{j,n}) - \frac{|i-j|}{n} \right| : 1 \leq i, j \leq n \right\} .$$

The right-hand side of this equation converges almost-surely to 0 and so $\text{dis}(C_n) \xrightarrow{\text{a.s.}} 0$, proving that $d_{\text{GHP}}(\mathcal{E}_n, \mathcal{E}_u) \xrightarrow{\text{a.s.}} 0$ as desired. $\qquad \square$

We now use the condition stated in this lemma to prove the convergence result from Theorem 1.4.

Proof of Theorem 1.4. In order to prove that the theorem holds, we simply show that $(T\langle X_{n,q}\rangle, ((1-q)n)^{-1} \cdot d_n, \mu_n)$ satisfies the conditions of Lemma 4.1. First of all, note that $T\langle X_{n,q}\rangle$ is indeed a set of size n and that μ_n is, by definition, the uniform measure on this set. It thus only remains to prove that there exists an ordering of the elements of $T\langle X_{n,q}\rangle$ satisfying (20).

Fix $n \in \mathbb{N}$ for now and organize $T\langle X_{n,q}\rangle$ into its sequence of left subtrees $T_{0,n}, \ldots, T_{r_n,n}$. More precisely, we have $r_n = R_{n,q}$ as defined in (5) and

$$T_{k,n} = T\langle X_{n,q}\rangle \cap \mathbf{1}^k \mathbb{T}_{\mathbf{0}} \subseteq \{\mathbf{1}^k\} \cup \left(T\langle X_{\infty,q}\rangle \cap \mathbf{1}^k \mathbf{0}\mathbb{T} \right).$$

Now, for any $0 \le k \le r$, let $x^k = (x^{i,k} : 1 \le i \le |T_{k,n}|)$ be an arbitrary ordering of the set $T_{k,n}$ and define $(x_{i,n})_{1 \le i \le n}$ as the concatenation of the $(x^k)_{0 \le k \le r_n}$, that is

$$(x_{1,n}, \ldots, x_{n,n}) = \left(x^{0,1}, \ldots, x^{0,|T_{0,n}|}, x^{1,1}, \ldots, x^{r_n,|T_{r_n,n}|} \right).$$

In words, the sequence $(x_{i,n})_{1 \le i \le n}$ goes in order through the left subtrees of $T\langle X_{n,q}\rangle$ and explore each such left subtree in an arbitrary order. We now prove that this sequence satisfies (20).

Consider $1 \le i, j \le n$. Note that $x_{i,n}$ belongs to the left subtree $T_{k,n}$ where $k = k_n(i)$ is the minimal integer satisfying

$$i \le |T_{0,n}| + \ldots + |T_{k,n}|.$$

Similarly, $x_{j,n}$ belongs to the left subtree $T_{\ell,n}$ where $\ell = k_n(j)$ is minimal so that

$$j \le |T_{0,n}| + \ldots + |T_{\ell,n}|.$$

Since the minimal distance between two vertices of $T_{k,n}$ and $T_{\ell,n}$ is attained when considering $\mathbf{1}^k$ and $\mathbf{1}^\ell$, whose distance is $|k - \ell|$, we have that

$$|k - \ell| \le d_n(x_{i,n}, x_{j,n}) \le |k - \ell| + |T_{k,n}| + |T_{\ell,n}|.$$

By taking the maximum over all the left subtree sizes, we obtain that

$$(21) \quad \max\left\{ \left| \frac{d_n(x_{i,n}, x_{j,n})}{(1-q)n} - \frac{|k_n(i) - k_n(j)|}{(1-q)n} \right| : 1 \le i, j \le n \right\} \le \frac{2}{(1-q)n} \max\left\{ |T_{k,n}| : 0 \le k \le r_n \right\}.$$

We now show that the right-hand side of this equation converges to 0 almost-surely.

Using the definition of $T_{k,n}$ and the coupling of the infinite Bernoulli model, we have that

$$T_{k,n} = T\langle X_{n,q}\rangle \cap \mathbf{1}^k \mathbb{T}_{\mathbf{0}} \subseteq \{\mathbf{1}^k\} \cup \left(T\langle X_{\infty,q}\rangle \cap \mathbf{1}^k \mathbf{0}\mathbb{T} \right).$$

We also recall from [ABC21, Lemma 1.10] that the family of random variables $(|T\langle X_{\infty,q}\rangle \cap \mathbf{1}^k \mathbf{0}\mathbb{T}|)_{k \ge 0}$ is a sequence of independent GEOMETRIC$(1-q)$. Moreover, using that $R_{n,q} \le n$, it follows that

$$\max\left\{ |T_{k,r}| : 0 \le k \le r_n \right\} \le 1 + \max\left\{ |T\langle X_{\infty,q}\rangle \cap \mathbf{1}^k \mathbf{0}\mathbb{T}| : 0 \le k \le n \right\}.$$

Thus, for any $\epsilon > 0$, we have

$$\mathbb{P}\left(\frac{2}{(1-q)n} \max\left\{ |T_{k,r}| : 0 \le k \le r_n \right\} > \epsilon \right)$$

$$\le \mathbb{P}\left(\max\left\{ |T\langle X_{\infty,q}\rangle \cap \mathbf{1}^k \mathbf{0}\mathbb{T}| : 0 \le k \le n \right\} \ge \frac{\epsilon(1-q)n}{2} - 1 \right)$$

$$= 1 - \left(1 - q^{\lceil \frac{\epsilon(1-q)n}{2} - 1 \rceil} \right)^{n+1}$$

$$= (1 + o(1)) n q^{\frac{\epsilon(1-q)n}{2}},$$

and the almost-sure convergences follows from applying the Borel-Cantelli lemma. This proves that the right-hand side of (21) converges alsmot-surely to 0 and so

$$(22) \qquad \max\left\{ \left| \frac{d_n(x_{i,n}, x_{j,n})}{(1-q)n} - \frac{|k_n(i) - k_n(j)|}{(1-q)n} \right| : 1 \le i, j \le n \right\} \xrightarrow{\text{a.s.}} 0$$

This is almost exactly (20) and we now prove that $|k_n(i) - \ell_n(j)|/(1-q)$ can be replaced with $|i - j|$ by proving that

$$(23) \qquad \max\left\{ \left| \frac{i}{n} - \frac{k_n(i)}{(1-q)n} \right| : 1 \le i \le n \right\} \xrightarrow{\text{a.s.}} 0 \,.$$

Recall the definition of $k = k_n(i)$ as the unique integer such that

$$\left| T_{0,n} \right| + \ldots + \left| T_{k-1,n} \right| < i \le \left| T_{0,n} \right| + \ldots + \left| T_{k,n} \right| \,.$$

First, using that all the left subtrees have size at least 1, we see that $k \le i$. Let now $(a_n)_{n \in \mathbb{N}}$ be a sequence of integers such that $a_n/n \to 0$ but $a_n/\log n \to \infty$ (for example $a_n = \lfloor \log^2 n \rfloor$ or $a_n = \lfloor \sqrt{n} \rfloor$). Then, the error in (23) by removing the first a_n term is negligible, as we observed that for $i \le a_n$ we had $k \le a_n$, implying

$$(24) \quad \left| \max\left\{ \left| \frac{i}{n} - \frac{k_n(i)}{(1-q)n} \right| : 1 \le i \le n \right\} - \max\left\{ \left| \frac{i}{n} - \frac{k_n(i)}{(1-q)n} \right| : a_n \le i \le n \right\} \right| \le \frac{a_n}{n} + \frac{a_n}{(1-q)n} \,.$$

Thus, it suffices to prove that the convergence in (23) is true for $a_n \le i \le n$.

Recall the definition of $F_{n,q}$ from (6), to see that, for any $k \ge 0$

$$\left| T_{0,n} \right| + \ldots + \left| T_{k-1,n} \right| \ge k + \sum_{\ell=0}^{k-1} \left| T\langle X_{\infty,q} \rangle \cap \mathbf{1}^\ell \mathbf{0} \mathbb{T} \right| - F_{n,q} \,.$$

Combining this inequality with the fact that $T\langle X_{n,q} \rangle$ is a subtree of $T\langle X_{\infty,q} \rangle$ and the definition of $k(i)$, we obtain

$$(25) \qquad k(i) + \sum_{\ell=0}^{k(i)-1} \left| T\langle X_{\infty,q} \rangle \cap \mathbf{1}^\ell \mathbf{0} \mathbb{T} \right| - F_{n,q} < i \le k(i) + 1 + \sum_{\ell=0}^{k(i)} \left| T\langle X_{\infty,q} \rangle \cap \mathbf{1}^\ell \mathbf{0} \mathbb{T} \right| \,.$$

Now, using the fact that $(|T\langle X_{\infty,q} \rangle \cap \mathbf{1}^k \mathbf{0} \mathbb{T}|)_{k \ge 0}$ is a family of independent GEOMETRIC$(1-q)$ random variables along with the strong law of large number implies that, as k goes to infinity, we have

$$\frac{1}{k} \sum_{\ell=0}^{k} \left| T\langle X_{\infty,q} \rangle \cap \mathbf{1}^\ell \mathbf{0} \mathbb{T} \right| \xrightarrow{\text{a.s.}} \frac{q}{1-q} \,.$$

Moreover, using the upper bound of (25) with $i = a_n$, we can verify that

$$\frac{a_n}{k_n(a_n)} \le 1 + \frac{1}{k_n(a_n)} + \frac{1}{k_n(a_n)} \sum_{\ell=0}^{k_n(a_n)} \left| T\langle X_{\infty,q} \rangle \cap \mathbf{1}^\ell \mathbf{0} \mathbb{T} \right| \,,$$

and so $k_n(a_n) \xrightarrow{\text{a.s.}} \infty$. Since $i \mapsto k_n(i)$ is increasing this implies that, uniformly over all $a_n \le i \le n$, we have

$$\frac{1}{k_n(i)} \sum_{\ell=0}^{k_n(i)} \left| T\langle X_{\infty,q} \rangle \cap \mathbf{1}^\ell \mathbf{0} \mathbb{T} \right| \xrightarrow{\text{a.s.}} \frac{q}{1-q} \,.$$

Combining this result with the upper bound of (25), it follows that, for n large enough, for any $a_n \le i \le n$, we almost-surely have that

$$\frac{i}{k_n(i)} \le 1 + \frac{1}{k_n(i)} + \frac{1}{k_n(i)} \sum_{\ell=0}^{k_n(i)} \left| T\langle X_{\infty,q} \rangle \cap \mathbf{1}^\ell \mathbf{0} \mathbb{T} \right|$$

$$\le 1 + 1 + \frac{2q}{1-q}$$

$$= \frac{2}{1-q} \,,$$

and so $k_n(i) \ge (1-q)i/2 \ge (1-q)a_n/2$.

Recall the bounds from (25) and use the previous convergence results related to $k_n(i)$ so see that

$$\left|\frac{i}{k_n(i)} - \frac{1}{1-q}\right| \leq \left|\frac{1}{k_n(i)}\sum_{\ell=0}^{k_n(i)}\left|T\langle X_{\infty,q}\rangle \cap \mathbf{1}^\ell \mathbf{0}\mathbb{T}\right| - \frac{q}{1-q}\right|$$

$$+ \frac{1}{k_n(i)} + \frac{\left|T\langle X_{\infty,q}\rangle \cap \mathbf{1}^{k_n(i)}\mathbf{0}\mathbb{T}\right| + F_{n,q}}{k_n(i)}.$$

Thus, thanks to $k_n(i) \geq k_n(a_n) \geq (1-q)a_n/2$ almost surely for n large enough, the fact that the left subtrees are almost surely finite, and the result from Lemma 2.2, satisfied here since $a_n/\log n \to \infty$, we eventually obtain that

$$\frac{i}{k_n(i)} \xrightarrow{\text{a.s.}} \frac{1}{1-q};$$

moreover, this convergence is uniform over $a_n \leq i \leq n$, meaning that

$$\max\left\{\left|\frac{i}{k_n(i)} - \frac{1}{1-q}\right| : a_n \leq i \leq n\right\} \xrightarrow{\text{a.s.}} 0.$$

From this convergence, we finally obtain that

$$(26) \qquad \max\left\{\left|\frac{i}{n} - \frac{k_n(i)}{(1-q)n}\right| : a_n \leq i \leq n\right\} \leq \max\left\{\frac{k_n(i)}{n} \cdot \left|\frac{i}{k_n(i)} - \frac{1}{1-q}\right| : a_n \leq i \leq n\right\}$$

$$\leq \max\left\{\left|\frac{i}{k_n(i)} - \frac{1}{1-q}\right| : a_n \leq i \leq n\right\} \xrightarrow{\text{a.s.}} 0,$$

where we use that $k_n(i) \leq i \leq n$ for the second inequality. We now have all the results to prove the theorem.

First of all, thanks to (26), we know that

$$\max\left\{\left|\frac{i}{n} - \frac{k_n(i)}{(1-q)n}\right| : a_n \leq i \leq n\right\} \xrightarrow{\text{a.s.}} 0.$$

We further recall that $a_n/n \to \infty$ which, along with (24)

$$\left|\max\left\{\left|\frac{i}{n} - \frac{k_n(i)}{(1-q)n}\right| : 1 \leq i \leq n\right\} - \max\left\{\left|\frac{i}{n} - \frac{k_n(i)}{(1-q)n}\right| : a_n \leq i \leq n\right\}\right| \leq \frac{a_n}{n} + \frac{a_n}{(1-q)n},$$

implies that (23) holds:

$$\max\left\{\left|\frac{i}{n} - \frac{k_n(i)}{(1-q)n}\right| : 1 \leq i \leq n\right\} \xrightarrow{\text{a.s.}} 0.$$

Now, we also proved (22), stating that

$$\max\left\{\left|\frac{d_n(x_{i,n}, x_{j,n})}{(1-q)n} - \frac{|k_n(i) - k_n(j)|}{(1-q)n}\right| : 1 \leq i, j \leq n\right\} \xrightarrow{\text{a.s.}} 0.$$

Moreover, for any $1 \leq i, j \leq n$, we have

$$\left|\frac{d_n(x_{i,n}, x_{j,n})}{(1-q)n} - \frac{|i-j|}{n}\right| \leq \left|\frac{d_n(x_{i,n}, x_{j,n})}{(1-q)n} - \frac{|k_n(i) - k_n(j)|}{(1-q)n}\right| + \left|\frac{k_n(i)}{(1-q)n} - \frac{i}{n}\right| + \left|\frac{k_n(j)}{(1-q)n} - \frac{j}{n}\right|.$$

Thus, combining the three previous equations, we see that

$$\max\left\{\left|\frac{d_n(x_{i,n}, x_{j,n})}{(1-q)n} - \frac{|i-j|}{n}\right| : 1 \leq i, j \leq n\right\} \xrightarrow{\text{a.s.}} 0,$$

which is exactly the condition (20) from Lemma 4.1. This finally prove that $(T\langle X_{n,q}\rangle, ((1-q)n)^{-1} \cdot d_n, \mu_n)$ satisfies all the conditions of Lemma 4.1 and thus implies that Theorem 1.4 holds. $\qquad\square$

4.3. **Subtree size limit of Mallows trees.** The aim of this section is to prove Theorem 1.5, which corresponds to the limit of Mallows trees according to the subtree size topology, first introduced in [Grü23]. We start by providing some background on this topology, before proving the desired limit.

First, we define the set of functions on $\mathbb{T}$ with mass 1 at the root, and whose mass split at every node:

$$\Psi := \left\{ \psi : \mathbb{T} \to [0,1] \ \Big| \ \psi(\varnothing) = 1 \ \& \ \forall v \in \mathbb{T}, \psi(v) = \psi(v\mathbf{0}) + \psi(v\mathbf{1}) \right\}$$

Now, given a finite binary search tree T, we can consider the corresponding function τ_T defined by

$$\tau_T(v) := \frac{|T \cap v\mathbb{T}|}{|T|} \, .$$

Note that τ_T does not belong to Ψ since $\tau_T(\mathbf{0}) + \tau_T(\mathbf{1}) = 1 - \frac{1}{|T|}$, but its limit will, provided that $|T| \to \infty$. This is the essence of the subtree size convergence.

Consider a sequence of random binary search trees $(T_n)_{n \in \mathbb{N}}$ such that $|T_n| \to \infty$ and let $\psi \in \Psi$. We say that the $(T_n)_{n \in \mathbb{N}}$ *converge with respect to the subtree sizes* to ψ, and write it $T_n \xrightarrow{\text{ssc}} \psi$ if $(\tau_{T_n})_{n \in \mathbb{N}}$ converges in distribution towards ψ. In other words, $(T_n)_{n \in \mathbb{N}}$ converge with respect to the subtree sizes to ψ if and only if, for all $r \geq 1$ and $v_1, \ldots, v_r \in \mathbb{T}$, we have

$$\left(\tau_{T_n}(v_i) \right)_{1 \leq i \leq r} \longrightarrow \left(\psi(v_i) \right)_{1 \leq i \leq r} \, ,$$

where the convergence occurs in distribution.

Naturally, following the observations from Section 2.3 and the remark at the end of Section 4.1, we could consider this type of convergence as the *weak* convergence with respect to the subtree sizes, and could also define the *almost-sure convergence with respect to the subtree sizes* (or the *convergence in probability with respect to the subtree sizes*) by replacing the previous convergence in distribution with an almost-sure convergence (or a convergence in probability). Actually, in the case of the coupling of the infinite Bernoulli model defined in Section 2.1.1, we will show that the previous convergence occurs almost-surely, thus corresponding to an almost-sure convergence with respect to the subtree sizes. This also naturally proves the regular subtree size convergence with respect to the weak topology.

Proof of Theorem 1.5. We want to show that $(T\langle X_{n,q} \rangle)_{n \in \mathbb{N}}$ converges to ψ_R defined in (4), that is, for all $v_1, \ldots, v_r$, we have

$$\left(\tau_{n,q}(v_i) \right)_{1 \leq i \leq r} \longrightarrow \left(\psi_r(v_i) \right)_{1 \leq i \leq r} \, ,$$

where $\tau_{n,q} = \tau_{T\langle X_{n,q} \rangle}$. Fix now $r \geq 1$ and consider the sequence $(v_0, \ldots, v_R)$ defined by $R = 2^{r+1} - 2$, $v_i = \mathbf{1}^i$ for $0 \leq i \leq r$, and the set $\{v_i : 0 \leq i \leq R\}$ corresponds to all the nodes of $\mathbb{T}$ at depth at most R. Note that the ordering of v_i for $i > r$ will not matter. With this sequence defined, we will now show that

$$\left(\tau_{n,q}(v_i) \right)_{0 \leq i \leq R} \xrightarrow{\text{a.s.}} \left(\psi_R(v_i) \right)_{0 \leq i \leq R} = \big(\underbrace{1, \ldots, 1}_{r+1}, 0, \ldots, 0 \big) \, ,$$

which will then imply the desired convergence with respect to the subtree sizes. Since this sequence is finite, the previous convergence result is actually equivalent to proving that

$$\tau_{n,q}(v_i) \xrightarrow{\text{a.s.}} \begin{cases} 1 & \text{if } 0 \leq i \leq r \\ 0 & \text{otherwise} \, . \end{cases}$$

We now focus on proving this last statement.

First, for any $0 \leq i \leq r$, we have $v_i = \mathbf{1}^i$ and so

$$\tau_{n,q}(v_i) = \frac{1}{n} \Big| T\langle X_{n,q} \rangle \cap \mathbf{1}^i \mathbb{T} \Big| = \frac{1}{n} \left[n - \sum_{j=0}^{i-1} \Big(1 + \big| T\langle X_{n,q} \rangle \cap \mathbf{1}^j \mathbf{0}\mathbb{T} \big| \Big) \right] \, .$$

Using the coupling of the infinite Bernoulli model, we see that $T\langle X_{n,q}\rangle$ is a subtree of $T\langle X_{\infty,q}\rangle$, and so we can bound the previous term as follows

$$\tau_{n,q}(v_i) \geq 1 - \frac{1}{n}\sum_{j=0}^{i-1}\left(1 + \left|T\langle X_{\infty,q}\rangle \cap \mathbf{1}^j\mathbf{0}\mathbb{T}\right|\right) \geq 1 - \frac{1}{n}\sum_{j=0}^{r}\left(1 + \left|T\langle X_{\infty,q}\rangle \cap \mathbf{1}^j\mathbf{0}\mathbb{T}\right|\right).$$

Moreover, we know that the sequence $(|T\langle X_{\infty,q}\rangle \cap \mathbf{1}^k\mathbf{0}\mathbb{T}|)_{k\geq 0}$ is a family of independent GEOMETRIC$(1-q)$ random variables [ABC21, Lemma 1.10], so the term

$$E = \sum_{j=0}^{r}\left(1 + \left|T\langle X_{\infty,q}\rangle \cap \mathbf{1}^j\mathbf{0}\mathbb{T}\right|\right)$$

satisfies

$$\mathbb{E}[E] = (r+1)\left(1 + \frac{q}{1-q}\right) = \frac{1+r}{1-q} < \infty,$$

and is thus almost-surely finite. This implies that $E/n \xrightarrow{\text{a.s.}} 0$ and so, using that $\tau_{n,q}(v_i) \leq 1$, we obtain that

$$\tau_{n,q}(v_i) \xrightarrow{\text{a.s.}} 1$$

for any $0 \leq i \leq r$.

Now, for any $i > r$, there exists some $0 \leq j \leq r$ such that $v_i \in \mathbf{1}^j\mathbf{0}\mathbb{T}$, and so

$$\tau_{n,q}(v_i) \leq \frac{1}{n}\left|T\langle X_{n,q}\rangle \cap \mathbf{1}^j\mathbf{0}\mathbb{T}\right| \leq \frac{E}{n}.$$

Thus, by using that E is almost-surely finite, we obtain that

$$\tau_{n,q}(v_i) \xrightarrow{\text{a.s.}} 0$$

for any $i > r$, thus proving the desired convergence result. $\qquad\square$

5. Discussions

In this section we briefly discuss how we can extend the results of this article to the case $q > 1$ and mentiond related open problems.

5.1. Ledwood trees. All the results from this work are applied to Mallows trees with $q \in [0, 1)$. As mentioned in Section 1.1, this choice makes sense since the case $q = 1$ behaves very differently and the case $q > 1$ can be obtained by symmetry. We now explain how to extend the results to the case $q > 1$ and will further discuss the case $q = 1$ in Section 5.2.

A known result regarding Mallows trees is that, if $T_{n,q}$ is a Mallows trees of size n with parameter $q \in (0, \infty)$, then $\overline{T}_{n,q}$ obtained by swapping all left and right subtrees is also a Mallows tree of size n but with parameter $1/q$. In a similar manner, of $\mathcal{X}_q$ is a two-sided infinite Mallows permutation with parameter q, then $-\mathcal{X}_q$ is a two-sided infinite Mallows permutation with parameter $1/q$. Finally, we can define ledwood trees as subtrees of $\overline{\mathbb{Y}}$ where $\overline{\mathbb{Y}}$ is defined in the same way as $\mathbb{Y}$ except that Υ is now considered the inverse of $\mathbf{0}$ instead of that of $\mathbf{1}$. Under this new definition, for any redwood tree Y, we can define the ledwood tree $\overline{Y}$ where we transform each $\mathbf{0}$ into $\mathbf{1}$ and vice-versa, but do not change Υ. Similarly, we can extend $Y\langle\cdot\rangle$ to $\overline{Y}\langle\cdot\rangle$ either by adapting the definition from (10), or by letting $\overline{Y}\langle\mathcal{X}\rangle = \overline{Y\langle-\mathcal{X}\rangle}$.

With all these notations, if $T_{n,q}$ is a Mallows tree of size n with parameter $q > 1$, then $\overline{T}_{n,q}$ is a Mallows tree of size n with parameter $1/q < 1$ and so, by Theorem 1.1, we have

$$\overline{T}_{n,q} \xrightarrow{\text{a.s.}} \mathcal{Y}_{1/q} = \left(o, Y\langle\mathcal{X}_{1/q}\rangle\right) \in \mathbb{Y}.$$

But now, $\overline{Y\langle\mathcal{X}_{1/q}\rangle} \stackrel{d}{=} \overline{Y\langle-\mathcal{X}_q\rangle} = \overline{Y}\langle\mathcal{X}_q\rangle$, leading to

$$T_{n,q} \xrightarrow{\text{a.s.}} \left(o, \overline{Y}\langle\mathcal{X}_q\rangle\right),$$

which is exactly the extension of Theorem 1.1 to the case $q > 1$ simply obtained by replacing the redwood tree of $\mathcal{X}_q$ (when $q < 1$) by its ledwood tree (when $q > 1$). Observe that the local limit of Mallows trees for q and $1/q$ are isomorphic to each other (up to the right coupling), so the statement of Theorem 1.1 could also be generalized to

$$T_{n,q} \xrightarrow{\text{a.s.}} \mathcal{Y}_{\min\{q,1/q\}}$$

for any $q \neq 1$, thus not having to define the concept of ledwood trees.

All other results extend in a similar manner, and we only briefly describe the changes below.

- Theorem 1.2: swap the role of $\mathbf{0}$ and $\mathbf{1}$, keeping Υ as it is but with the convention that $\Upsilon = \mathbf{0}^{-1}$ so $\mathbf{0}^k = \mathbf{1}^{|k|}$ for $k < 0$ and $\mathbf{1}^k$ is only defined for $k \geq 0$.
- Theorem 1.3: here there is a small problem since one-sided infinite Mallows permutations are not define when $q > 1$; thus we can either show that the limit when $q > 1$ is just the symmetric of the limit when $q < 1$, or we can define *infinite Mallows functions* as bijective functions from $\mathbb{N}$ to $-\mathbb{N} = \{\ldots, -2, -1\}$ whose weight is proportional to $q^{\text{Inv}(\sigma)}$, having in mind that this is only properly defined when $q > 1$.
- Theorem 1.4: the limit in this case actually does not change, we just need to replace the distance by $((1 - 1/q)n)^{-1} \cdot d_n$ instead of $((1 - q)n)^{-1} \cdot d_n$.
- Theorem 1.5: we replace the limit by its symmetric, that is $\psi_L = \mathbb{1}_{v \in \{\mathbf{0}^k : k \geq 0\}}$.

5.2. **Open problems.** From the previous results and observations, we see that only the case $q = 1$ remains to be studied in order to obtain a full characterization of the limit of Mallows trees. However, this case behaves very differently than when $q \neq 1$ and we thus believe that such result would be harder to prove. Moreover, since we expect a sharp change of behaviour in the limit of Mallows trees between the cases $q \neq 1$ and $q = 1$, we would also likely observe various regimes if we let $q = q_n$ depend on n (as it was observed in the case of the height of the trees for example [ABC21]). Characterizing these limits and how they depend on the relation between q and n are likely challenging questions.

A possibly easier follow-up project would be to extend these results to the case of random regenerative permutations [PT19]. These permutations, which are known to generalize Mallows permutations, show strong regenerative properties, which happen to be one of the most important property in the study of the limit of Mallows trees. While this study would naturally build off the tools developed here, it is not clear whether the limit could also be characterized in two different ways: as a sequence of trees attached to a straight line and as a redwood tree built off a two-sided infinite permutation. For this reason we decided to focus on the case of Mallows permutations here and keep the case of random regenerative permutations for a possible later work.

Acknowledgement

References

[ABC21] Louigi Addario-Berry and Benoît Corsini. The height of Mallows trees. *The Annals of Probability*, 49(5):2220–2271, 2021.

[AS04]   David Aldous and J Michael Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures*, pages 1–72. Springer, 2004.

[BB17]   Riddhipratim Basu and Nayantara Bhatnagar. Limit theorems for longest monotone subsequences in random Mallows permutations. *Annales de L'Institut Henri Poincare Section (B) Probability and Statistics*, 53(4):1934–1951, 2017.

[BP15]   Nayantara Bhatnagar and Ron Peled. Lengths of monotone subsequences in a Mallows permutation. *Probability Theory and Related Fields*, 161(3-4):719–780, 2015.

[BS01]   Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic Journal of Probability*, 6:1–13, 2001.

[CDE18] Harry Crane, Stephen DeSalvo, and Sergi Elizalde. The probability of avoiding consecutive patterns in the Mallows distribution. *Random Structures & Algorithms*, 53(3):417–447, 2018.

[DM10]   Amir Dembo and Andrea Montanari. Ising models on locally tree-like graphs. *The Annals of Applied Probability*, pages 565–592, 2010.

[EGW12]   Steven N Evans, Rudolf Grübel, and Anton Wakolbinger. Trickle-down processes and their boundaries. *Electron. J. Probab*, 17(1):1–58, 2012.

[GO10]   Alexander Gnedin and Grigori Olshanski. q-exchangeability via quasi-invariance. *The Annals of Probability*, 38(6):2103–2135, 2010.

[GO12]   Alexander Gnedin and Grigori Olshanski. The two-sided infinite extension of the Mallows model for random permutations. *Advances in Applied Mathematics*, 48(5):615–639, 2012.

[GP18]   Alexey Gladkich and Ron Peled. On the cycle structure of Mallows permutations. *The Annals of Probability*, 46(2):1114–1169, 2018.

[Grü23]   Rudolf Grübel. A note on limits of sequences of binary trees. *Discrete Mathematics & Theoretical Computer Science*, 25(Analysis of Algorithms), 2023.

[He22]   Jimmy He. A central limit theorem for descents of a Mallows permutation and its inverse. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 58, pages 667–694. Institut Henri Poincaré, 2022.

[HMV22]   Jimmy He, Tobias Müller, and Teun W. Verstraaten. Cycles in Mallows random permutations. *Random Structures & Algorithms*, 2022.

[Hof23]   Remco van der Hofstad. *Random graphs and complex networks, volume II*. 2023.

[Mac98]   Ian Grant Macdonald. *Symmetric functions and Hall polynomials*. Oxford University Press, 1998.

[Mal57]   Colin L Mallows. Non-null ranking models. I. *Biometrika*, 44(1/2):114–130, 1957.

[Mie09]   Grégory Miermont. Tessellations of random maps of arbitrary genus. *Annales scientifiques de l'Ecole normale supérieure*, 42(5):725–781, 2009.

[MS13]   Carl Mueller and Shannon Starr. The length of the longest increasing subsequence of a random Mallows permutation. *Journal of Theoretical Probability*, 26(2):514–540, 2013.

[Muk16]   Sumit Mukherjee. Fixed points and cycle structure of random permutations. *Electronic Journal of Probability*, 21, 2016.

[Pin21]   Ross G Pinsky. Permutations avoiding a pattern of length three under Mallows distributions. *Random Structures & Algorithms*, 58(4):676–690, 2021.

[PT19]   JIM PITMAN and WENPIN TANG. Regenerative random permutations of integers. *The Annals of Probability*, 47(3):1378–1416, 2019.