

THE HEIGHT OF MALLOWS TREES

BY LOUIGI ADDARIO-BERRY* AND BENOÎT CORSINI†

Department of Mathematics and Statistics, McGill University, *louigi.addario@mcgill.ca; †benoit.corsini@mail.mcgill.ca

Random binary search trees are obtained by recursively inserting the elements $\sigma(1), \sigma(2), \dots, \sigma(n)$ of a uniformly random permutation σ of $[n] = \{1, \dots, n\}$ into a binary search tree data structure. Devroye (*J. Assoc. Comput. Mach.* **33** (1986) 489–498) proved that the height of such trees is asymptotically of order $c^* \log n$, where $c^* = 4.311\dots$ is the unique solution of $c \log((2e)/c) = 1$ with $c \geq 2$. In this paper, we study the structure of binary search trees $T_{n,q}$ built from Mallows permutations. A Mallows(q) permutation is a random permutation of $[n] = \{1, \dots, n\}$ whose probability is proportional to $q^{\text{Inv}(\sigma)}$, where $\text{Inv}(\sigma) = |\{i < j : \sigma(i) > \sigma(j)\}|$. This model generalizes random binary search trees, since Mallows(q) permutations with $q = 1$ are uniformly distributed. The laws of $T_{n,q}$ and $T_{n,q^{-1}}$ are related by a simple symmetry (switching the roles of the left and right children), so it suffices to restrict our attention to $q \leq 1$.

We show that, for $q \in [0, 1]$, the height of $T_{n,q}$ is asymptotically $(1 + o(1))(c^* \log n + n(1 - q))$ in probability. This yields three regimes of behaviour for the height of $T_{n,q}$, depending on whether $n(1 - q)/\log n$ tends to zero, tends to infinity or remains bounded away from zero and infinity. In particular, when $n(1 - q)/\log n$ tends to zero, the height of $T_{n,q}$ is asymptotically of order $c^* \log n$, like it is for random binary search trees. Finally, when $n(1 - q)/\log n$ tends to infinity, we prove stronger tail bounds and distributional limit theorems for the height of $T_{n,q}$.

CONTENTS

1. Introduction	2221
1.1. Related work	2223
1.2. Notation	2224
1.3. Mallows trees and random binary search trees	2224
1.4. Right depth and height of Mallows trees	2226
1.5. Intermediate values	2229
1.6. Distributional limits	2230
2. Connection to random binary search trees	2232
2.1. A useful function	2232
2.2. Upper tail bound	2234
2.3. Lower tail bound	2236
3. Right depth of Mallows trees	2240
3.1. The bivariate moment generating function	2240
3.2. Bounds on the left subtrees	2242
3.3. Almost sure convergence	2244
4. Intermediate values and threshold process	2245
4.1. Threshold process	2246
4.2. Bounds on M_m^B	2251
4.3. Bounds on $ T_n^B(\bar{1}^{R_m^B+1}) $	2252
4.4. Convergence in probability	2254

Received July 2020; revised December 2020.

MSC2020 subject classifications. Primary 60B15, 60C05; secondary 05A05, 60F05, 60F15, 60K35, 82B23, 82B26.

Key words and phrases. Mallows trees, Mallows permutations, binary search trees, random trees, random permutations, heights of trees.

5. Distributional limits 2258
 5.1. Central limit theorem for the right depth 2258
 5.2. Central limit theorem for the height of Mallows trees 2259
 5.3. Poisson fluctuations for the height 2262
 6. Further questions 2263
 Appendix A: Moments of the right depth 2264
 Appendix B: Asymptotics of μ_α 2265
 Acknowledgements 2270
 Funding 2270
 References 2270

1. Introduction. Let $T_\infty = \{\emptyset\} \cup \bigcup_{k \geq 0} \{\bar{0}, \bar{1}\}^k$ be the complete infinite rooted binary tree, with nodes at depth $n \geq 1$ indexed by strings $u = u_1, \dots, u_k \in \{\bar{0}, \bar{1}\}^k$, so u has parent u_1, \dots, u_{k-1} and children $u\bar{0}$ and $u\bar{1}$. For a set $V \subset T_\infty$ and node $u \in T_\infty$, we write $uV = \{uv, v \in V\}$.

For $u \in \{\bar{0}, \bar{1}\}^k$, we write $|u| = k$ and say that u has depth k . A subtree of T_∞ (or just “a tree”, for short) is a set $T \subset T_\infty$ which is connected when viewed as a subgraph of T_∞ . For any subtree T of T_∞ , the *root* of T is defined to be the unique element of T of minimum depth. For a tree T and a node $u \in T_\infty$, we write $T(u) = (uT_\infty) \cap T$ for the subtree of T rooted at u ; when $\emptyset \in T$, then $T(u) = \emptyset$ if and only if $u \notin T$. Finally, for $T \subset T_\infty$, we write $h(T) = \sup\{|u|, u \in T\} - \inf\{|u|, u \in T\}$; if T is a tree then $h(T)$ is the greatest distance of any node of T from the root of T .

For $n \geq 1$, we write $[n] = \{1, 2, \dots, n\}$. Given an injective function $f : [n] \rightarrow \mathbb{Z}_+ := \{1, 2, \dots\}$, the *binary search tree* $T\langle f \rangle$ is the subtree of T_∞ defined inductively as follows (see Figure 1 for an example). If $n = 0$, then $T\langle f \rangle := \emptyset$ is the empty tree. Otherwise, view f as a sequence of distinct integers $f = (f(1), f(2), \dots, f(n))$, and write f^- (resp., f^+) for the subsequence of f consisting of terms $f(i)$ such that $f(i) < f(1)$ (resp., $f(i) > f(1)$), listed in the same order as in f . Then set $T\langle f \rangle := \{\emptyset\} \cup (\bar{0}T\langle f^- \rangle) \cup (\bar{1}T\langle f^+ \rangle)$.

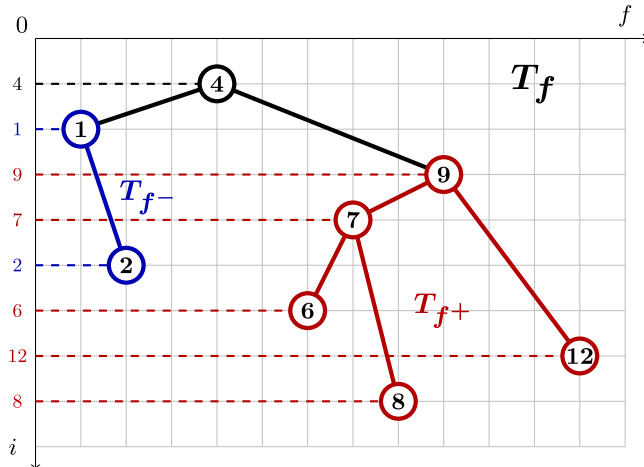


FIG. 1. The labelled tree $(T\langle f \rangle, \tau\langle f \rangle)$ for $f = (4, 1, 9, 7, 2, 6, 12, 8)$. The sequences f^- and f^+ are $(9, 7, 6, 12, 8)$, respectively. The subtree in blue corresponds to $T\langle f^- \rangle$ and the one in red corresponds to $T\langle f^+ \rangle$. The corresponding labels given by $\tau\langle f \rangle$ are written on the nodes; so, for example, $\tau\langle f \rangle(\bar{0}\bar{1}) = 2$ and $\tau\langle f \rangle(\bar{1}\bar{1}) = 12$.

We label the nodes of $T\langle f \rangle$ by the elements of $\{f(1), f(2), \dots, f(n)\}$ as follows. Set $\tau\langle f \rangle(\emptyset) = f(1)$; then, inductively, for nodes $u \in T\langle f \rangle$ with $|u| \geq 1$, set

$$\tau\langle f \rangle(u) = \begin{cases} \tau\langle f^- \rangle(v) & \text{if } u = \bar{0}v, \\ \tau\langle f^+ \rangle(v) & \text{if } u = \bar{1}v. \end{cases}$$

The definitions of $T\langle f \rangle$ and $\tau\langle f \rangle$ easily extend to injective functions $f : \mathbb{Z}^+ \mapsto \mathbb{Z}^+$, by considering the sequence $(f(i), i \geq 1)$.

In this article, we study the heights of binary search trees built from random, Mallows-distributed permutations. For $n \geq 0$ and $q \in [0, \infty)$, the Mallows distribution with parameters n and q (introduced in [23]) is the probability measure $\pi_{n,q}$ on the symmetric group \mathcal{S}_n given by

$$\pi_{n,q}(\sigma) := Z_{n,q}^{-1} \cdot q^{\text{Inv}(\sigma)}.$$

Here, $\text{Inv}(\sigma) := |\{1 \leq i < j \leq n : \sigma(i) > \sigma(j)\}|$ 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.

For a permutation $\sigma = (\sigma(1), \dots, \sigma(n))$, the reversed permutation $\sigma' = (n + 1 - \sigma(1), \dots, n + 1 - \sigma(n))$ has $\text{Inv}(\sigma') = \binom{n}{2} - \text{Inv}(\sigma)$. This implies that, if σ is a $\pi_{n,q}$ -distributed random permutation, then its reversal is $\pi_{n, \frac{1}{q}}$ -distributed. The effect of this reversal on the associated binary search trees is also easy to understand: $T_{\sigma'}$ is obtained from T_{σ} by swapping all left and right subtrees. Since the map $q \mapsto \frac{1}{q}$ bijectively sends $(1, \infty)$ to $(0, 1)$, it follows from these observations that we may as well restrict our attention to $q \in [0, 1]$. Note that when $q = 0$, $\pi_{n,q}$ assigns weight 1 to the identity permutation and when $q = 1$, $\pi_{n,q}$ is the uniform distribution on \mathcal{S}_n .

We prove the following results. In what follows, we write $T_{n,q}$ for a random tree with the distribution of $T\langle \sigma \rangle$ for σ a $\pi_{n,q}$ -distributed random permutation, and we write $\text{MT}(n, q)$ for the law of such a tree; we call $T_{n,q}$ a *Mallows tree* (with parameters n and q). Also, we let $c^* = 4.311\dots$ be the unique solution of $c \log(\frac{2e}{c}) = 1$ with $c \geq 2$.

THEOREM 1.1. *For any $[0, 1]$ -valued sequence $(q_n)_{n \geq 0}$,*

$$\frac{h(T_{n,q_n})}{n(1 - q_n) + c^* \log n} \rightarrow 1$$

in probability and in L_p for any $p > 0$.

When $q_n = 1$ for all n , the trees T_{n,q_n} are *random binary search trees*—the binary search trees corresponding to uniformly random permutations. This case of Theorem 1.1 implies that $\frac{h(T_{n,1})}{c^* \log n} \rightarrow 1$ in probability, which is a well-known result of Devroye [11].

On the other hand, when $q_n = q \in [0, 1)$ for all n , Theorem 1.1 implies that $h(T_{n,q_n}) = (1 - q + o_{\mathbb{P}}(1))n$. In this case, $T_{n,q}$ consists of a “rightward” path of length $(1 - q + o_{\mathbb{P}}(1))n$, with left subtrees of height $O_{\mathbb{P}}(\log \frac{1}{1-q})$ hanging from each of its nodes. (The notation $O_{\mathbb{P}}$ and $o_{\mathbb{P}}$ is defined in Section 1.2, below.)

When $(q_n)_{n \geq 0}$ is small enough that the first term in the denominator overwhelms the second, we are able to strengthen the above result, obtaining strong bounds on the rate of convergence.

THEOREM 1.2. *Fix any $[0, 1]$ -valued sequence $(q_n)_{n \geq 0}$ such that $n(1 - q_n)/\log n \rightarrow \infty$. Then for any $\varepsilon > 0$ and $\lambda > 0$,*

$$\mathbb{P}\left(\left| \frac{h(T_{n,q_n})}{n(1 - q_n)} - 1 \right| > \varepsilon\right) = O\left(\frac{1}{n^\lambda}\right).$$

When the first term in the denominator is dominant and also $nq_n \rightarrow \infty$, we prove a central limit theorem for the height.

THEOREM 1.3. *Fix any $[0, 1]$ -valued sequence $(q_n)_{n \geq 0}$ such that $n(1 - q_n)/\log n \rightarrow \infty$ and $nq_n \rightarrow \infty$. Then*

$$\frac{h(T_{n,q_n}) - n(1 - q_n) - c^* \log((1 - q_n)^{-1})}{\sqrt{n(1 - q_n)q_n}} \xrightarrow{d} \text{NORMAL}(0, 1).$$

Finally, when $nq_n = O(1)$, we prove a Poisson limit theorem for the height (after re-centering but without re-scaling).

THEOREM 1.4. *Let $(q_n)_{n \geq 0}$ be any $[0, 1]$ -valued sequence such that $nq_n \rightarrow \lambda \in [0, \infty)$. Then*

$$n - 1 - h(T_{n,q_n}) \xrightarrow{d} \text{POISSON}(\lambda).$$

The next subsection briefly discusses related literature on random trees and Mallows permutations. Section 1.2 then introduces some notation we need. The remainder of Section 1 describes some of the key tools used in proving Theorem 1.1–1.4, and while doing so, provides an overview of our approach to their proofs. Theorem 1.1 is proved in three parts, depending on whether $n(1 - q_n)$ is much smaller than, much larger than, or of the same order as $\log n$. The arguments for these cases are sketched in Sections 1.3–1.5, respectively. Since the proof of Theorem 1.2 essentially consists in extracting quantitative estimates from the proof of Theorem 1.1, we do not spend much space on it in the [Introduction](#). Finally, Section 1.6 describes our arguments for our distributional limit results, Theorem 1.3 and 1.4.

1.1. Related work. The Mallows permutation model was first introduced by C.L. Mallows [23] in the context of ranking theory. The study of its probabilistic properties has taken off in the past decade; previously studied properties of Mallows permutations include the length of the longest increasing subsequence [4, 6, 25], the cycle structure and pattern occurrence [10, 17, 21, 26, 27], relations to exchangeability [18, 19] and to random matchings [3], random dynamics with Mallows permutations as stationary distribution [5, 12] and thermodynamic properties of Mallows measures [33, 34].

Since our work is focused on random trees built from Mallows permutations, it is also natural to situate it in the context of the literature on random trees. This is a vast literature and we only discuss a smattering of it. As mentioned above, Devroye [11] proved that the height h_n of a random binary search tree of size n is asymptotically $(c^* + o_{\mathbb{P}}(1)) \log n$; this built on previous work of Pittel [29], who proved that $h_n/\log n \rightarrow \alpha \in (0, \infty)$ almost surely, but did not identify the constant α . Random binary search trees lie within the more general *increasing tree model*, for which the first-order behaviour of the height has been well characterized [8, 14]. Building on Devroye’s results, Reed [30] and Drmota [13] found two conceptually different proofs that the variance of h_n is bounded in n .

The study of random binary trees, random increasing trees and their ilk, is intimately connected to the properties of *branching random walk*; results on the height of random trees are often extracted (at varying levels of difficulty) from results on the maximal displacement of a corresponding branching random walk. For example, the results of [8, 9, 11] rely on the Hammersley–Kingman–Biggins theorem [7, 20, 22], which provides a law of large numbers for the maximum of branching random walks; and the arguments of Reed [30] proceeds by relating the height of binary search trees to the minimal position in a binary branching random walk with exponential step distribution. Further related results on minima in branching

random walks can be found in [1, 2]. The lecture notes [31] provide an excellent introduction to the theory of branching random walks.

Finally, Mallows trees were introduced by S. N. Evans, R. Grübel and A. Wakolbinger [16], who studied properties of the tree and generating processes. They showed, among other results, that Mallows trees are a specific case of *trickle down process*; that is to say, they can be generated in a sequential manner, by adding one leaf at a time.

1.2. *Notation.* For functions $f : \mathbb{R} \rightarrow \mathbb{R}$, $g : \mathbb{R} \rightarrow \mathbb{R}$ or $f : \mathbb{N} \rightarrow \mathbb{R}$, $g : \mathbb{N} \rightarrow \mathbb{R}$, we write $f = O(g)$ to mean $f(n) = O(g(n))$ as $n \rightarrow \infty$ unless a different limit is specified. If $f = O(g)$, then we also write $g = \Omega(f)$. We also use the notation $f = o(g)$ and its synonym $g = \omega(f)$. If $f = o(g)$, then we will also write $f \ll g$ and $g \gg f$. We write $f \sim g$ to mean that $f(n) = (1 + o(1))g(n)$ as $n \rightarrow \infty$.

For sequences of random variables $(X_n)_{n \geq 0}$ and $(Y_n)_{n \geq 0}$, we write $X_n = O_{\mathbb{P}}(Y_n)$ if, for all $\varepsilon > 0$, there exists $K > 0$ such that

$$\limsup_{n \geq 0} \mathbb{P}(|X_n| \geq K Y_n) < \varepsilon.$$

We write $X_n = o_{\mathbb{P}}(Y_n)$ if, for all $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n| > \varepsilon Y_n) = 0.$$

For random variables X and Y , we write $X \stackrel{d}{=} Y$ if X and Y have the same distribution. We write $X \leq Y$, or equivalently $Y \geq X$, if for all $x \in \mathbb{R}$, we have

$$\mathbb{P}(X \geq x) \leq \mathbb{P}(Y \geq x);$$

in this case we say X is stochastically smaller than Y .

1.3. *Mallows trees and random binary search trees.* Many properties of random binary search trees can be extended to Mallows trees. Perhaps the most fundamental of these are the *branching property*, which means that disjoint subtrees of a Mallows tree are conditionally independent given their sizes, and the *projective consistency*, which is the fact that subtrees of Mallows trees are again Mallows trees. The following proposition, due to Evans, Grübel and Wakolbinger, [16], formalizes these properties, and additionally describes the joint distribution of the sizes of the left and right subtrees of the root in a Mallows tree.

PROPOSITION 1.5 ([16], Section 7). *For all $q \in [0, 1]$ and $n \geq 1$, for any $0 \leq k \leq n - 1$, we have*

$$\mathbb{P}(|T_{n,q}(\bar{0})| = k) = \mathbb{P}(|T_{n,q}(\bar{1})| = n - 1 - k) = \begin{cases} \frac{(1 - q)q^k}{1 - q^n} & \text{if } q \in [0, 1), \\ \frac{1}{n} & \text{if } q = 1. \end{cases}$$

Moreover, $T_{n,q}(\bar{0})$ and $T_{n,q}(\bar{1})$ are conditionally independent Mallows trees given their sizes. That is, for every $0 \leq k \leq n - 1$, for any trees t_0 and t_1 rooted at \emptyset , of respective sizes k and $n - 1 - k$, we have

$$\mathbb{P}(T_{n,q}(\bar{0}) = \bar{0}t_0, T_{n,q}(\bar{1}) = \bar{1}t_1 \mid |T_{n,q}(\bar{0})| = k) = \mathbb{P}(T_{k,q} = t_0)\mathbb{P}(T_{n-1-k,q} = t_1).$$

Conversely, these properties characterize Mallows trees.

From this proposition, one can see that the split between left and right subtree is not symmetric; the right subtree at any node is stochastically larger than its left subtree. This observation straightforwardly leads to the following proposition, stating that the rightmost path in $T_{n,q}$ is the stochastically longest path.

PROPOSITION 1.6. For all $q \in [0, 1]$ and $n \geq 1$, for all $v \in T_\infty$,

$$\mathbb{P}(v \in T_{n,q}) \leq \mathbb{P}(\bar{I}^{|v|} \in T_{n,q}).$$

We prove Proposition 1.6 in Section 2.2. With the result of this last proposition, we can bound the height of T_{n,q_n} from above by using a union bound over all nodes at a given depth, as

$$\begin{aligned} \mathbb{P}(h(T_{n,q_n}) \geq h) &\leq \sum_{v \in T_\infty : |v|=h} \mathbb{P}(v \in T_{n,q_n}) \\ &\leq 2^h \mathbb{P}(\bar{I}^h \in T_{n,q_n}). \end{aligned}$$

When $n(1 - q_n)/\log n \rightarrow 0$, this bound is tight enough to prove the upper bound of Theorem 1.1. It will also be useful in proving that $((\frac{h(T_{n,q_n})}{n(1-q_n)+c^*\log n})^p)_{n \geq 1}$ is uniformly integrable for any sequence $(q_n)_{n \geq 0}$, thereby extending the convergence in probability to the L_p convergence.

For $(q_n)_{n \geq 0}$ such that $n(1 - q_n)/\log n \rightarrow 0$, we also use a comparison argument, albeit a slightly more complicated one, to prove the lower bound. At the heart of the argument is the following computation. Let U be UNIFORM([0, 1]). Then, for any $q \in [0, 1)$, $n \geq 1$ and $0 \leq k \leq n - 1$,

$$\begin{aligned} \mathbb{P}\left(\left\lfloor \frac{\log(1 - U(1 - q^n))}{\log q} \right\rfloor = k\right) &= \mathbb{P}\left(k \leq \frac{\log(1 - U(1 - q^n))}{\log q} < k + 1\right) \\ &= \mathbb{P}\left(\frac{1 - q^k}{1 - q^n} \leq U < \frac{1 - q^{k+1}}{1 - q^n}\right) \\ &= \frac{(1 - q)q^k}{1 - q^n}. \end{aligned}$$

This identity and Proposition 1.5 together imply that we can generate a $MT(n, q)$ -distributed tree as follows. Let $(U_v)_{v \in T_\infty}$ be independent UNIFORM([0, 1]) random variables indexed by the nodes of T_∞ . Set $S_{n,q}(\emptyset) = n$. Then, for $v \in T_\infty$, inductively define

$$S_{n,q}(v\bar{0}) = \begin{cases} \left\lfloor \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)})}{\log q} \right\rfloor & \text{if } q \in (0, 1), \\ \lfloor S_{n,q}(v)U_v \rfloor & \text{if } q = 1, \\ 0 & \text{if } q = 0 \end{cases}$$

and

$$S_{n,q}(v\bar{1}) = S_{n,q}(v) - 1 - S_{n,q}(v\bar{0}).$$

Then the tree $T_{n,q} = \{v \in T_\infty : S_{n,q}(v) \geq 1\}$ is $MT(n, q)$ -distributed, and $S_{n,q}$ corresponds to the size of the subtree at any given node: $S_{n,q}(u) = |T_{n,q}(u)|$.

This construction couples the trees $(T_{n,q})$ as both q and n vary. Using this coupling, we will be able to prove the following proposition.

PROPOSITION 1.7. For all $q \in [0, 1)$ and $n \geq 1$, for any $0 \leq \ell \leq n$, with $m = \lfloor \frac{1 - q^n}{1 - q^{\ell+2}} \rfloor$, then

$$\mathbb{P}(h(T_{n,q}) \leq \ell) \leq \mathbb{P}(h(T_{m,1}) \leq \ell).$$

The proof of Proposition 1.7 can be found in Section 2.3. When $(q_n)_{n \geq 0}$ is such that $n(1 - q_n)/\log n \rightarrow 0$, this stochastic bound combined with results of Devroye [11] will yield the desired lower bound.

$b_{i,j}$	1	2	3	4	5	6	7	8	9	10	11	12	13	...
1	0	0	0	1 ₄	1	0	0	1	1	0	1	0	0	...
2	1 ₁	0	1	0	0	1	1	0	0	1	0	0	1	
3	0	0	0	1	0	0	0	0	1 ₉	1	1	0	0	...
4	0	0	0	0	0	0	1 ₇	0	1	0	0	1	1	...
5	0	1 ₂	1	0	1	0	1	1	0	0	0	0	1	...
6	1	1	0	0	0	1 ₆	0	0	1	0	0	1	0	...
7	0	0	0	1	0	1	1	0	0	0	0	1 ₁₂	0	...
8	1	0	0	1	0	0	1	1 ₈	0	1	0	0	0	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

FIG. 2. An example of the top-left corner of an infinite $\{0, 1\}$ -valued matrix $b = (b_{i,j})_{i,j \geq 1}$. The boxed 1's are in position $(i, f^b(i))$ and their column numbers are indicated as blue subscripts. For a given column i , the numbers in red correspond to positions (i, j) where $j \in F_{i-1}^b$ and can be found below the boxed 1's; the boxed 1 in row i is always the first nonred 1 in row i .

1.4. *Right depth and height of Mallows trees.* The results of the previous section, relating Mallows trees to random binary search trees, give tight bounds on the height only when $n(1 - q_n)/\log n \rightarrow 0$; in this case the two tree models show strong resemblance. When $(q_n)_{n \geq 0}$ does not satisfy this condition, the rightmost path of a Mallows tree is playing a more important role in its height. We now study the properties of this path and its connection to the rest of the tree.

Note that if $f : [n] \rightarrow \mathbb{Z}_+$ is an injective function and $f' = f|_{[n-1]}$, then $T\langle f' \rangle$ is the subtree of $T\langle f \rangle$ consisting of the nodes with labels $f(1), \dots, f(n - 1)$. More precisely, $T\langle f' \rangle = \{v \in T_f : \tau\langle f \rangle(v) \neq f(n)\}$, and $\tau\langle f' \rangle$ is the restriction of $\tau\langle f \rangle$ to $T\langle f' \rangle$. For example, in Figure 1, with $f' = (f(1), \dots, f(7)) = (4, 1, 9, 7, 2, 6, 12)$, the tree $T\langle f' \rangle$ is obtained from the depicted tree $T\langle f \rangle$ by removing the node with label 8. We next use this fact to describe an explicit construction of a nested sequence of Mallows trees, which will be useful for our analysis.

Given an infinite $\{0, 1\}$ -valued matrix $b = (b_{i,j})_{i,j \geq 1}$ with infinitely many ones in every row, define an injective function $f^b = (f^b(i), i \geq 1)$ as follows. Let $f^b(1) = \inf\{j \in \mathbb{N} : b_{1,j} = 1\}$. Having defined $(f^b(k), 1 \leq k < i)$, let $F_{i-1}^b = \{f^b(k), 1 \leq k < i\}$ and set $f^b(i) = \inf\{j \in \mathbb{N} \setminus F_{i-1}^b : b_{i,j} = 1\}$. We write T_n^b as shorthand for $T\langle f_n^b \rangle$, where $f_n^b = (f^b(i), 1 \leq i \leq n)$. An example is shown in Figure 2. For the matrix b shown in that figure, we obtain $f_8^b = (4, 1, 9, 7, 2, 6, 12, 8)$, so the tree T_8^b is precisely the binary search tree shown in Figure 1.

The infinite b -model for trees is the sequence of (labelled) trees $(T_n^b)_{n \geq 0}$, which by construction is increasing, in that T_n^b is a subtree of T_{n+1}^b for all $n \geq 0$. The corresponding sequence of labelling functions is defined by $(\tau_n^b)_{n \geq 0} := (\tau\langle f_n^b \rangle)_{n \geq 0}$. We sometimes omit the matrix b when it is clear from context. The utility of this construction is explained by the following proposition.

PROPOSITION 1.8 ([4], Lemma 2.1). *Fix $q \in [0, 1)$ and let $B = (B_{i,j})_{i,j \geq 1}$ have independent BERNOULLI($1 - q$) entries, and for $n \geq 1$, let $\sigma_n^B \in S_n$ be the permutation of $[n]$ defined by $\sigma_n^B(i) = \text{rank}\{f_n^B(i), F_n^B\}$. Then σ_n^B is $\pi_{n,q}$ -distributed for all $n \geq 0$.*

It follows that for B as in the proposition, T_n^B is $\text{MT}(n, q)$ -distributed for all n . The random trees in the sequence $(T_n^B)_{n \geq 0}$ may be viewed as the successive states of a transient Markov

chain taking values in the set of finite subtrees of T_∞ . This chain was already defined in [16], where aspects of its asymptotic behaviour were studied; however, the observation that its one-dimensional marginals are all Mallows-distributed appears to be new.

The next corollary is a direct consequence of Proposition 1.8 and the fact that $(T_n^b)_{n \geq 0}$ is increasing for all b ; its proof is omitted.

COROLLARY 1.9. *Let $n \geq 0$ and $q \in [0, 1]$. Then, for all $u \in T_\infty$,*

$$\mathbb{P}(u \in T_{n,q}) \leq \mathbb{P}(u \in T_{n+1,q}).$$

We write $T^B = \bigcup_{n \geq 0} T_n^B$ for the infinite tree which is the increasing limit of the sequence $(T_n^B)_{n \geq 0}$, and τ^B for the corresponding labelling. It is immediate from the construction that $T^B = T(f^B)$.

For B as in the proposition, the random function $f^B : \mathbb{Z}_+ \rightarrow \mathbb{Z}_+$ defined previously is a.s. a bijection, and its law is the so-called MALLOWS(q) distribution on $S_\infty = \{\sigma : \mathbb{Z}_+ \rightarrow \mathbb{Z}_+ : \sigma \text{ a permutation}\}$, introduced in [18]; the fact that f^B is MALLOWS(q)-distributed was proved in [17].

Let $M_0^B = 0$ and for $n \geq 1$, let $M_n^B = \max(f^B(1), \dots, f^B(n))$. Then set $R_0^B = 0$ and for $n \geq 1$, let $R_n^B = |\{i \in [n] : M_i^B > M_{i-1}^B\}|$ be the number of records in the sequence $(f^B(1), \dots, f^B(n))$. Note that R_n^B is precisely the *right depth* of T_n^B , that is, $R_n^B = \max\{d : \bar{1}^d \in T_n^B\}$. Also, for all $k \geq 1$, if $R_{n-1}^B = k - 1$ and $R_n^B = k$ then $\tau^B(\bar{1}^k) = M_n^B$.

For any $k \geq 0$ and any node $u \in T^B(\bar{1}^k \bar{0})$, we have $\tau^B(\bar{1}^{k-1}) < \tau^B(u) < \tau^B(\bar{1}^k)$; here, we write $\bar{1}^0 = \emptyset$ and for $k = 0$, set $\tau^B(\bar{1}^{k-1}) = 0$. Since f^B is a bijection, it follows that the subtree $T^B(\bar{1}^k \bar{0})$ contains exactly $\tau^B(\bar{1}^k) - \tau^B(\bar{1}^{k-1}) - 1$ nodes, and the labels assigned to these nodes by τ^B are precisely the elements of the set $\{\tau^B(\bar{1}^{k-1}) + 1, \dots, \tau^B(\bar{1}^k) - 1\}$. Moreover, since any infinite sequence of positive integers contains infinitely many records, necessarily T^B contains the infinite rightward path $P_R := \{\bar{1}^k, k \geq 0\}$, and the left subtrees $(T^B(\bar{1}^k \bar{0}), k \geq 0)$ hanging from P_R have respective sizes $(\tau^B(\bar{1}^k) - \tau^B(\bar{1}^{k-1}) - 1, k \geq 1)$.

Much of our analysis will be based on the decompositions of T^B and T_n^B as

$$T^B = P_R \cup \bigcup_{k \geq 0} T^B(\bar{1}^k \bar{0}),$$

$$T_n^B = \{\bar{1}^k, 0 \leq k \leq R_n^B\} \cup \bigcup_{0 \leq k \leq R_n^B} T_n^B(\bar{1}^k \bar{0}).$$

From the second decomposition, it is immediate that

$$(1.1) \quad h(T_{n,q}) \stackrel{d}{=} h(T_n^B) = \max_{0 \leq k \leq R_n^B} \{h(T_n^B(\bar{1}^k \bar{0})) + k + 1\} \leq \max_{0 \leq k \leq R_n^B} \{h(T^B(\bar{1}^k \bar{0})) + k + 1\}.$$

In order to use (1.1) to get useful information about the height, we need to understand the distributions of R_n^B and of the subtrees $T_n^B(\bar{1}^k \bar{0})$ and $T^B(\bar{1}^k \bar{0})$. The last of these is the easiest to describe. We say a random variable G is GEOMETRIC(c)-distributed if $\mathbb{P}(G = k) = (1 - c)^k c$ for $k \in \mathbb{N} = \{0, 1, 2, \dots\}$.

LEMMA 1.10. *Fix $q \in [0, 1]$ and let $B = (B_{i,j})_{i,j \geq 1}$ have independent BERNOULLI($1 - q$) entries. Then the random trees $(T^B(\bar{1}^k \bar{0}))_{k \geq 0}$ are independent and identically distributed with*

$$T^B(\bar{1}^k \bar{0}) \stackrel{d}{=} \bar{1}^k \bar{0} T_{G(q),q},$$

where $G(q)$ is GEOMETRIC($1 - q$)-distributed and is independent of the trees $(T_{n,q})_{n \geq 0}$. In other words, for all $k, n \geq 0$ and any tree $t \subseteq T_\infty$ with $|t| = n$, we have

$$\mathbb{P}(T^B(\bar{1}^k \bar{0}) = \bar{1}^k \bar{0}t) = q^n(1 - q) \cdot \mathbb{P}(T_{n,q} = t).$$

PROOF. As explained above, for all $k \geq 0$,

$$|T^B(\bar{1}^k \bar{0})| = \tau^B(\bar{1}^k) - \tau^B(\bar{1}^{k-1}) - 1.$$

Moreover, by construction the random variables $(\tau^B(\bar{1}^k) - \tau^B(\bar{1}^{k-1}) - 1)_{k \geq 0}$ are independent and GEOMETRIC($1 - q$)-distributed. Finally, by Proposition 1.5, conditionally on having a given size n , $T^B(\bar{1}^k \bar{0})$ is MT(n, q)-distributed, from which the result follows. \square

Combined with (1.1), this lemma yields a key distributional upper bound on $h(T_{n,q})$. We now have

$$\begin{aligned} h(T_{n,q}) &\stackrel{d}{=} \max_{0 \leq k \leq R_n^B} \{h(T_n^B(\bar{1}^k \bar{0})) + k + 1\} \leq \max_{0 \leq k \leq R_n^B} \{h(T^B(\bar{1}^k \bar{0})) + k + 1\} \\ &= R_n^B + 1 + \max_{0 \leq k \leq R_n^B} \{h(T^B(\bar{1}^{R_n^B - k} \bar{0})) - k\} \\ &\stackrel{d}{=} R_n^B + 1 + \max_{0 \leq k \leq R_n^B} \{h(T^B(\bar{1}^k \bar{0})) - k\} \\ &\leq R_n^B + 1 + \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\}. \end{aligned}$$

Considering that R_n^B corresponds to the depth of the rightmost path in T_n^B , we also have $R_n^B \leq h(T_n^B) \stackrel{d}{=} h(T_{n,q})$. In combination with (1.1), this yields that

$$(1.2) \quad R_n^B \leq h(T_{n,q}) \leq R_n^B + 1 + \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\},$$

where we recall that \leq denotes stochastic inequality. Lemma 1.10 tells us the trees whose heights appear in the final supremum are independent and $T_{G(q),q}$ -distributed. Combining this inequality with a union-bound and Proposition 1.6, we now have

$$\begin{aligned} \mathbb{P}(h(T_{n,q}) \geq h) &\leq \mathbb{P}\left(R_n^B + 1 + \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} \geq h\right) \\ &\leq \inf_{0 \leq \ell \leq h} \left\{ \mathbb{P}(R_n^B + 1 \geq h - \ell) + \sum_{k \geq 0} \mathbb{P}(h(T^B(\bar{1}^k \bar{0})) \geq \ell + k) \right\} \\ &= \inf_{0 \leq \ell \leq h} \left\{ \mathbb{P}(R_n^B + 1 \geq h - \ell) + \sum_{k \geq 0} \mathbb{P}(h(T_{G(q),q}) \geq \ell + k) \right\} \\ &\leq \inf_{0 \leq \ell \leq h} \left\{ \mathbb{P}(R_n^B + 1 \geq h - \ell) + \sum_{k \geq 0} 2^{k+\ell} \mathbb{P}(R_{G(q)}^B \geq \ell + k) \right\}, \end{aligned}$$

where in the final line $G(q)$ should be understood to be independent of the random variables in B . The preceding argument gives us a way to derive upper tail bounds on $(h(T_{n,q}))$ exclusively by controlling the upper tails of the random variables $(R_n^B, n \geq 1)$. The next proposition is our key tool for doing so.

PROPOSITION 1.11. Fix $q \in [0, 1)$ and let $B = (B_{i,j})_{i,j \geq 1}$ be an array of independent BERNOULLI($1 - q$) random variables. Then the sequence $(R_n^B, M_n^B)_{n \geq 0}$ is a Markov chain with transition probabilities given by

$$\mathbb{P}(R_{n+1}^B = r + k, M_{n+1}^B = m + \ell \mid R_n^B = r, M_n^B = m) = \begin{cases} q^{m+\ell-n-1}(1 - q) & \text{if } k = 1, \ell \geq 1, \\ 1 - q^{m-n} & \text{if } k = 0, \ell = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, for $x, y \in \mathbb{C}$ such that $q|y| < 1$, we have

$$\mathbb{E}[x^{R_n^B+1} y^{M_n^B}] = y^n \prod_{1 \leq k \leq n} \frac{q + (1 - q)x - q^k}{1 - q^k y}.$$

We prove Proposition 1.11 in Section 3.1. We do not in fact make use of the bivariate moment generating function (MGF) of R_n^B and M_n^B , but only the MGFs of R_n^B and M_n^B , which can of course be easily recovered from the bivariate MGF. However, it seems interesting to record the explicit form of the bivariate MGF. Moreover, although the MGF of R_n^B has already been computed [32], (1.135), we did not find a reference for the MGF of M_n^B , and we also did not see a derivation of the MGF of M_n^B which is substantially shorter than that of the bivariate MGF.

Using the moment generating function to control the behaviour of $R_n^B - \mathbb{E}[R_n^B]$ yields Chernoff-type bounds for both the upper and lower tail. The bounds are strong enough that they allow us to prove both the upper and lower bounds of Theorem 1.2, implying the bounds of Theorem 1.1 when $(q_n)_{n \geq 0}$ is such that $n(1 - q_n)/\log n \rightarrow \infty$. For the upper bound, the key consequence of Proposition 1.11 is the following proposition, which allows us to control the right-hand side of (1.2). It will also be used in the analysis for other ranges of $(q_n)_{n \geq 0}$.

PROPOSITION 1.12. There exist universal constants M, C and λ such that, for all $q \in [0, 1)$ and $\xi \in \mathbb{R}$, we have

$$\mathbb{P}\left(\sup_{k \geq 0} \{h(T^B(\bar{\Gamma}^k \bar{0})) - k\} \geq c^* \log\left(\frac{1}{1 - q}\right) + M \sqrt{\log\left(\frac{1}{1 - q}\right) + \xi}\right) \leq C e^{-\lambda \xi}.$$

As above, c^* is the unique solution of $c \log\left(\frac{2e}{c}\right) = 1$ with $c \geq 2$. The proof of this proposition can be found in Section 3.2.

1.5. *Intermediate values.* The most technical part of the proof of Theorem 1.1 appears when $n(1 - q_n)/\log n = \Theta(1)$. In this situation, the proof uses a combination of the techniques from the two previous cases, when $n(1 - q_n)/\log n \rightarrow 0$ and when $n(1 - q_n)/\log n \rightarrow \infty$.

In order to obtain the identities and bounds in (1.2), we decomposed T_n^B into its rightmost path $(\bar{\Gamma}^k, 0 \leq k \leq R_n^B)$, together with the left subtrees hanging from each of its nodes. Because $T_n^B(\bar{\Gamma}^{R_n^B+1}) = \emptyset$, this decomposition can be rewritten as

$$T_n^B = \{\bar{\Gamma}^k, 0 \leq k < R_n^B + 1\} \cup \left(\bigcup_{0 \leq k < R_n^B + 1} T_n^B(\bar{\Gamma}^k \bar{0}) \right) \cup T_n^B(\bar{\Gamma}^{R_n^B+1}).$$

For any $0 \leq d \leq R_n^B + 1$, we may similarly decompose T_n^B along the initial segment $(\bar{\Gamma}^k, 0 \leq k < d)$ of the rightmost path to obtain

$$T_n^B = \{\bar{\Gamma}^k, 0 \leq k < d\} \cup \left(\bigcup_{0 \leq k < d} T_n^B(\bar{\Gamma}^k \bar{0}) \right) \cup T_n^B(\bar{\Gamma}^d).$$

From this decomposition, it is immediate that, for all $d \geq 0$,

$$h(T_{n,q}) \stackrel{d}{=} h(T_n^B) \leq \max \left\{ \max_{0 \leq k < d} \{h(T_n^B(\bar{1}^k \bar{0})) + k + 1\}, d + h(T_n^B(\bar{1}^d)) \right\},$$

with equality whenever $d \leq R_n^B + 1$. This inequality implies that, for all $d \geq 0$, we have

$$\begin{aligned} h(T_{n,q}) &\leq \max \left\{ \max_{0 \leq k < d} \{h(T^B(\bar{1}^k \bar{0})) + k + 1\}, d + h(T_n^B(\bar{1}^d)) \right\} \\ (1.3) \quad &\stackrel{d}{=} d + \max \left\{ \max_{0 \leq k < d} \{h(T^B(\bar{1}^k \bar{0})) - k\}, h(T_n^B(\bar{1}^d)) \right\} \\ &\leq d + \max \left\{ \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\}, h(T_n^B(\bar{1}^d)) \right\}; \end{aligned}$$

The first term in the maximum is the same supremum as appears in (1.2), and Proposition 1.12 gives an essentially sharp upper tail bound for this term of the form $c^* \log n + O_{\mathbb{P}}(\sqrt{\log n})$. Because we aim to prove that $\frac{h(T_{n,q_n})}{n(1-q_n) + c^* \log n} \rightarrow 1$ in probability, a natural choice for d is then $n(1 - q_n)$. This choice indeed gives the desired bound for $h(T_n^B(\bar{1}^d))$, due to the following proposition.

PROPOSITION 1.13. *Let $(q_n)_{n \geq 0}$ be such that $n(1 - q_n)/\log n = \Theta(1)$. Then*

$$\frac{h(T_{n,q_n}(\bar{1}^{\lfloor n(1-q_n) \rfloor}))}{c^* \log n} \rightarrow 1$$

in probability.

The proof of this proposition can be found in Section 4.4, and boils down to showing that

$$(1 - q_n) \cdot |T_{n,q_n}(\bar{1}^{\lfloor n(1-q_n) \rfloor})| = o_{\mathbb{P}}(\log |T_{n,q_n}(\bar{1}^{\lfloor n(1-q_n) \rfloor})|),$$

then applying the bounding technique described in Section 1.3. This proposition, combined with the upper bound on the supremum, proves the upper bound of Theorem 1.1 when $n(1 - q_n)/\log n = \Theta(1)$.

To prove the corresponding lower bound, we use the simple inequality

$$h(T_{n,q}) \stackrel{d}{=} h(T_n^B) \geq d + h(T_n^B(\bar{1}^d)),$$

which holds for any $0 \leq d \leq R_n^B + 1$. Again taking $d = \lfloor n(1 - q_n) \rfloor$, which is at most R_n^B with high probability, Proposition 1.13 then yields that $h(T_{n,q_n}) \geq n(1 - q_n) + (c^* + o_{\mathbb{P}}(1)) \log n$, which is the lower bound of Theorem 1.1 when $n(1 - q_n)/\log n = \Theta(1)$.

1.6. *Distributional limits.* In Section 1.4, we described the strong connection between $h(T_{n,q_n})$ and the right depth R_n^B whenever $(q_n)_{n \geq 0}$ is such that $n(1 - q_n)/\log n \rightarrow \infty$; in this regime, we can transfer many results on the asymptotic behaviour of $(R_n^B)_{n \geq 0}$ to the sequence $(h(T_{n,q_n}))_{n \geq 0}$.

If not only $n(1 - q_n)/\log n \rightarrow \infty$ but, more strongly, $nq_n \rightarrow \lambda \in [0, \infty)$, then it is straightforward to prove that $\mathbb{P}(h(T_n^B) = R_n^B) = 1 - o(1)$. In this case, by studying the characteristic function of $n - 1 - R_n^B$, it follows fairly easily that

$$n - 1 - R_n^B \xrightarrow{d} \text{POISSON}(\lambda),$$

from which Theorem 1.4 follows. The details of this argument appear in Section 5.3.

If we assume now that $nq_n \rightarrow \infty$, by analyzing the moment generating function of R_n^B given in Proposition 1.11, we can prove a central limit theorem for the right depth; this is stated in the following proposition.

PROPOSITION 1.14. *Let $(q_n)_{n \geq 0}$ be such that $nq_n \rightarrow \infty$. Then*

$$\frac{R_n^B - \sum_{1 < k \leq n} \frac{1 - q_n}{1 - q_n^k}}{\sqrt{n(1 - q_n)q_n + \log(n \wedge \frac{1}{1 - q_n})}} \xrightarrow{d} \text{NORMAL}(0, 1).$$

In particular, if $n(1 - q_n)/\log n \rightarrow \infty$, then this implies that

$$\frac{R_n^B - n(1 - q_n) - \log((1 - q_n)^{-1})}{\sqrt{n(1 - q_n)q_n}} \xrightarrow{d} \text{NORMAL}(0, 1).$$

The proof can be found in Section 5.1. If we furthermore assume that $n(1 - q_n)/(\log n)^2 \rightarrow \infty$, then the $\log((1 - q_n)^{-1})$ term in the numerator of this proposition can be removed; in this case, using (1.2) to compare $h(T_{n,q_n})$ with R_n^B gives tight enough bounds to establish the conclusion of Theorem 1.3.

For the remaining regime of $(q_n)_{n \geq 0}$, that is, when $n(1 - q_n)/\log n \rightarrow \infty$ and $n(1 - q_n)/(\log n)^2 = O(1)$, both terms $n(1 - q_n)$ and $\log((1 - q_n)^{-1})$ contribute to the asymptotic behaviour of R_n^B . In this regime, the proof of the central limit theorem for $h(T_{n,q_n})$ requires a similar technique to the one of the proof of Theorem 1.1 in the intermediate case (when $n(1 - q_n)/\log n = \Theta(1)$), but with a different choice of d .

Previously, we deterministically chose $d = \lfloor n(1 - q_n) \rfloor$. In the current setting, we instead require d to be a random variable defined as follows. For $n \geq 0$, let $m = m(n)$ be the smallest integer such that $m(1 - q_n) + \log m \geq n(1 - q_n)$ and let $D = D(n) = R_m^B + 1$. The same chain of reasoning that yielded (1.2) and (1.3) now gives us the bounds

$$(1.4) \quad h(T_n^B(\bar{1}^{D(n)})) \leq h(T_n^B) - D(n) \leq \max \left\{ \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\}, h(T_n^B(\bar{1}^{D(n)})) \right\}.$$

It is not hard to see that $m(n) = n - (1 + o(1)) \frac{\log n}{1 - q_n}$ and then we use Proposition 1.14 to prove that

$$\frac{R_{m(n)}^B - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \xrightarrow{d} \text{NORMAL}(0, 1).$$

Moreover, Proposition 1.12 straightforwardly implies that the supremum on the right-hand side of (1.4) is at most $c^* \log n + O_{\mathbb{P}}(\sqrt{\log n})$. Finally, the next proposition gives the last ingredient to conclude the proof of Theorem 1.3 in the case $n(1 - q_n)/\log n \rightarrow \infty$ and $n(1 - q_n)/(\log n)^2 = O(1)$.

PROPOSITION 1.15. *Let $(q_n)_{n \geq 0}$ be such that $\log(n(1 - q_n)) = O(\sqrt{\log n})$ and $n(1 - q_n) = \omega(\sqrt{\log n})$. For $n \geq 0$, let $m = m(n) = \min\{\ell \geq 0 : \ell(1 - q_n) + \log \ell \geq n(1 - q_n)\}$. Then the sequence of random variables*

$$\left(\frac{h(T_n^B(\bar{1}^{R_m^B + 1})) - c^* \log n}{\sqrt{\log n}} \right)_{n \geq 2}$$

is tight.

The rather technical proof of Proposition 1.15 can be found in Section 5.2. Proposition 1.15 implies that $h(T_n^B(\bar{1}^{R_m^B + 1}))/c^* \log n \rightarrow 1$ in probability, which suggests some relation to Proposition 1.13. However, we did not see a simple way to give a unified statement. The reason for the hypothesis on $(q_n)_{n \geq 0}$ in Proposition 1.15 is mainly due to an error term of order $O(\sqrt{\log n})$ in several of the asymptotic estimates which arise in our analysis.

Combining Proposition 1.14 and 1.13, the bounds in (1.2) and (1.4), and a subsequence argument as for Theorem 1.1, we can conclude the proof of Theorem 1.3. This also concludes the sketch of the proofs of the three theorems.

2. Connection to random binary search trees. In this section, we prove Theorem 1.1 in the case when $(q_n)_{n \geq 0}$ is such that $n(1 - q_n)/\log n \rightarrow 0$. As explained in Section 1.3, the proof will be divided into upper and lower bound. We also prove the following proposition, which gives a bound on the second order term for the convergence in probability and will be useful for Theorem 1.3.

PROPOSITION 2.1. *Let $(q_n)_{n \geq 0}$ be taking values in $[0, 1)$ such that $n(1 - q_n)/\log n \rightarrow 0$. Then, for any sequence $(\gamma_n)_{n \geq 0}$ such that $\gamma_n/(n(1 - q_n) \vee \sqrt{\log n}) \rightarrow \infty$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{P}(|h(T_{n,q_n}) - c^* \log n| \geq \gamma_n) = 0.$$

2.1. A useful function. Before proving Proposition 2.1, we introduce and study a very useful function, which will play a role in the analysis for all the regimes of $(q_n)_{n \geq 0}$. For $n \geq 0$, $q \in [0, 1)$ and $\alpha \geq 1$, let

$$\mu_\alpha(n, q) := \sum_{1 < k \leq n} \left(\frac{1 - q}{1 - q^k} \right)^\alpha.$$

The following fact gives a first indication of the importance of μ_α .

FACT 2.2. For all $n \geq 0$ and $q \in [0, 1)$, with $B = (B_{i,j})_{i,j \geq 1}$ having independent BERNULLI($1 - q$) entries, we have

$$\mathbb{E}[R_n^B] = \mu_1(n, q)$$

and

$$\text{Var}[R_n^B] = \mu_1(n, q) - \mu_2(n, q).$$

More generally, there is a formula relating the expectation of $(R_n^B)^\alpha$ to the functions μ_i for $i \leq \alpha$, as given in the next proposition.

PROPOSITION 2.3. *Let $n \geq 0$, $q \in [0, 1)$, and $B = (B_{i,j})_{i,j \geq 1}$ have independent BERNULLI($1 - q$) entries. Write $S_\beta = \{(s_1, \dots, s_\beta) : s_1 + 2s_2 + \dots + \beta s_\beta = \beta\}$ and for $s = (s_1, \dots, s_\beta) \in S_\beta$, let $|s| = s_1 + \dots + s_\beta$. Then, for all integer $\alpha \geq 1$, we have*

$$\mathbb{E}[(R_n^B)^\alpha] = \sum_{1 \leq \beta \leq \alpha \wedge (n-1)} \sum_{s \in S_\beta} (-1)^{\beta + |s|} \beta! \begin{Bmatrix} \alpha \\ \beta \end{Bmatrix} \prod_{1 \leq i \leq \beta} \frac{\mu_i(n, q)^{s_i}}{i^{s_i} s_i!},$$

where $(\begin{Bmatrix} \alpha \\ \beta \end{Bmatrix})_{\alpha, \beta \geq 1}$ are Stirling numbers of the second kind.

The proofs of Fact 2.2 and Proposition 2.3 can be found in Appendix A. Because $\mu_1(n, q)$ relates to the expected value of R_n^B , it can also be used to give concentration bounds on the right depth, as shown below.

LEMMA 2.4. *Fix $n \geq 0$ and $q \in [0, 1)$. Then, for any $c > 1$, we have*

$$\mathbb{P}(R_n^B > c\mu_1(n, q)) \leq \exp\left(\left[c \log\left(\frac{e}{c}\right) - 1\right]\mu_1(n, q)\right)$$

and

$$\mathbb{P}(R_n^B < c^{-1}\mu_1(n, q)) \leq \exp(\lceil c^{-1} \log(ce) - 1 \rceil \mu_1(n, q)).$$

PROOF. The proof simply follows from Chernoff’s bounds. For all $t > 0$, by Proposition 1.11, we have

$$\begin{aligned} \mathbb{P}(R_n^B > c\mu_1(n, q)) &\leq \frac{\mathbb{E}[e^{tR_n^B}]}{e^{tc\mu_1(n, q)}} \\ &= e^{-tc\mu_1(n, q)} \prod_{1 < k \leq n} \left(1 + (e^t - 1) \frac{1 - q}{1 - q^k}\right). \end{aligned}$$

By the convexity of the exponential function, we have $1 + x \leq e^x$, hence

$$\prod_{1 < k \leq n} \left(1 + (e^t - 1) \frac{1 - q}{1 - q^k}\right) \leq \prod_{1 < k \leq n} e^{(e^t - 1) \frac{1 - q}{1 - q^k}} = \exp((e^t - 1)\mu_1(n, q)).$$

Using this bound in the previous equation, we obtain

$$\mathbb{P}(R_n^B > c\mu_1(n, q)) \leq \exp(-tc\mu_1(n, q) + (e^t - 1)\mu_1(n, q)).$$

The optimal value for t corresponds to $e^t = c$ and yields the first bound of the lemma. Similarly, for any $t > 0$, we have

$$\begin{aligned} \mathbb{P}(R_n^B < c^{-1}\mu_1(n, q)) &\leq e^{tc^{-1}\mu_1(n, q)} \prod_{1 < k \leq n} \left(1 + (e^{-t} - 1) \frac{1 - q}{1 - q^k}\right) \\ &\leq \exp(tc^{-1}\mu_1(n, q) + (e^{-t} - 1)\mu_1(n, q)), \end{aligned}$$

and the optimal bound is given by $e^{-t} = c^{-1}$, yielding the second bound. \square

In order to use these results, we need to better understand the behaviour of μ_α . The following two propositions describe the asymptotic growth of μ_1 , and of μ_α for $\alpha > 1$, respectively.

PROPOSITION 2.5. Consider any sequence $(q_n)_{n \geq 0}$ taking values in $[0, 1)$. Then we have

$$\mu_1(n, q_n) = n(1 - q_n) + \log\left(n \wedge \frac{1}{1 - q_n}\right) + O\left(\sqrt{\log\left(n \wedge \frac{1}{1 - q_n}\right)}\right).$$

PROPOSITION 2.6. Let $\alpha > 1$. Consider any sequence $(q_n)_{n \geq 0}$ taking values in $[0, 1)$. Then we have

$$\mu_\alpha(n, q_n) = n(1 - q_n)^\alpha + \zeta(\alpha) - 1 + O\left(\left((1 - q_n) \vee \frac{1}{n}\right)^{\frac{\alpha - 1}{\alpha + 1}}\right) = n(1 - q_n)^\alpha + O(1),$$

where $\zeta(\alpha)$ is the Riemann zeta function: $\zeta(\alpha) = \sum_{k \geq 1} \frac{1}{k^\alpha}$.

The second equality in Proposition 2.6 directly follows from the first one as $(1 - q_n) \vee \frac{1}{n} \leq 1$. The proofs of these two propositions can be found in Appendix B. Since the details of the proofs are tedious, we provide a brief sketch of the idea. The proof technique heavily relies on the fact that $\frac{1 - q}{1 - q^k} \sim \frac{1}{k}$ whenever $k(1 - q) = o(1)$, and that $\frac{1 - q}{1 - q^k} \sim (1 - q)$ whenever $k(1 - q) = \omega(1)$. Pretending for the moment that the first asymptotic holds whenever $k(1 - q) \leq 1$ and that the second one holds whenever $k(1 - q) > 1$, and assuming that $1 \ll \frac{1}{1 - q} \ll n$, it follows that

$$\mu_\alpha(n, q) = \sum_{1 < k \leq n} \left(\frac{1 - q}{1 - q^k}\right)^\alpha$$

$$\begin{aligned} &\simeq \sum_{1 < k \leq \frac{1}{1-q}} \frac{1}{k^\alpha} + \sum_{\frac{1}{1-q} < k \leq n} (1-q)^\alpha \\ &\simeq n(1-q)^\alpha - 1 + \sum_{1 \leq k \leq \frac{1}{1-q}} \frac{1}{k^\alpha}; \end{aligned}$$

the second term asymptotically simplifies to $\log(\frac{1}{1-q})$ when $\alpha = 1$ and to $\zeta(\alpha)$ when $\alpha > 1$.

The fact that when $\alpha > 1$, the second-order term of the development of μ_α is $\zeta(\alpha) - 1$, is interesting to us, and we can imagine it has already appeared in the literature; however, we were unable to find a reference.

2.2. Upper tail bound. To prove the upper bound of Proposition 2.1, we start by proving Proposition 1.6, as this stochastic bound is at the heart of the proof.

PROOF OF PROPOSITION 1.6. We must show that, for all $v \in T_\infty$, $n \geq 0$ and $q \in [0, 1]$, we have

$$\mathbb{P}(v \in T_{n,q}) \leq \mathbb{P}(\bar{1}^{|v|} \in T_{n,q}).$$

First of all, if $q = 0$, then $T_{n,q} = \{\bar{1}^k, 0 \leq k \leq n\}$, and the inequality is clearly true. Secondly, if $q = 1$, then $T_{n,q}$ is a random binary search tree, and by symmetry we have

$$\mathbb{P}(v \in T_{n,q}) = \mathbb{P}(u \in T_{n,q})$$

for all u, v such that $|u| = |v|$, which also proves the inequality.

Fix now $q \in (0, 1)$. We prove the result by induction on n . For $n \leq 1$, the assertion holds because either $T_{n,q}$ is empty or $T_{n,q} = \{\emptyset\}$. Consider now some node $v \in T_\infty$. Write $T_{n,q}^L$ and $T_{n,q}^R$, respectively, for the left and right subtrees of $T_{n,q}$ re-rooted at \emptyset ; in other words, $T_{n,q}(\bar{0}) = \bar{0}T_{n,q}^L$ and $T_{n,q}(\bar{1}) = \bar{1}T_{n,q}^R$.

Assume first that $v = \bar{1}v'$. In this case, $v \in T_{n,q}$ if and only if $v' \in T_{n,q}^R$. Moreover, by Proposition 1.5, we know that, conditioned on its size, $T_{n,q}^R$ is a Mallows tree; since $|T_{n,q}^R| < n$, by induction we thus have

$$\mathbb{P}(v \in T_{n,q}) = \mathbb{P}(v' \in T_{n,q}^R) \leq \mathbb{P}(\bar{1}^{|v'|} \in T_{n,q}^R) = \mathbb{P}(\bar{1}^{|v|} \in T_{n,q}),$$

which proves the statement in this case.

Assume next that $v = \bar{0}v'$. In this case, $v \in T_{n,q}$ if and only if $v' \in T_{n,q}^L$. Moreover, by Proposition 1.5, for $0 \leq k \leq n - 1$, we have

$$\mathbb{P}(|T_{n,q}^L| \leq k) = \frac{1 - q^{k+1}}{1 - q^n} \geq q^{n-k-1} \frac{1 - q^{k+1}}{1 - q^n} = \frac{q^{n-k-1} - q^n}{1 - q^n} = \mathbb{P}(|T_{n,q}^L| \geq n - k - 1).$$

Since $\mathbb{P}(|T_{n,q}^L| \geq n - k - 1) = \mathbb{P}(|T_{n,q}^R| \leq k)$, it follows that $\mathbb{P}(|T_{n,q}^L| \leq k) \geq \mathbb{P}(|T_{n,q}^R| \leq k)$, which means that $|T_{n,q}^L| \leq |T_{n,q}^R|$. By Corollary 1.9, it follows that

$$\mathbb{P}(v \in T_{n,q}) = \mathbb{P}(v' \in T_{n,q}^L) \leq \mathbb{P}(v' \in T_{n,q}^R) \leq \mathbb{P}(\bar{1}^{|v|} \in T_{n,q}),$$

proving the assertion in this case.

Finally, if $v = \emptyset$, then $v = \bar{1}^{|v|}$ and $\mathbb{P}(v \in T_{n,q}) = \mathbb{P}(\bar{1}^{|v|} \in T_{n,q})$. This completes the proof. □

Before proving Proposition 2.1, we state and prove the following proposition, which allows us to extend the results from convergence in probability to convergence in L_p in Theorem 1.1.

PROPOSITION 2.7. For any sequence $(q_n)_{n \geq 0}$ and any $p > 0$, the family of random variables

$$\left(\left(\frac{h(T_{n,q_n})}{n(1 - q_n) + c^* \log n} \right)^p \right)_{n \geq 1}$$

is uniformly integrable.

PROOF. First of all, by Proposition 2.5 we have $n(1 - q_n) + c^* \log n = \Theta(\mu_1(n, q_n))$; hence, it suffices to prove that $(h(T_{n,q_n})/\mu_1(n, q_n))^p$ is uniformly integrable, and we now do this.

Fix $a \in \mathbb{R}$. Combine Proposition 1.6 with a union bound to obtain

$$\begin{aligned} \mathbb{P}\left(\left(\frac{h(T_{n,q_n})}{\mu(n, q_n)}\right)^p \geq a\right) &= \mathbb{P}(h(T_{n,q_n}) \geq a^{\frac{1}{p}} \mu(n, q_n)) \\ (2.1) \qquad \qquad \qquad &= \mathbb{P}(\exists v \in T_{n,q_n} \text{ such that } |v| = \lceil a^{\frac{1}{p}} \mu_1(n, q_n) \rceil) \\ &\leq 2^{a^{\frac{1}{p}} \mu(n, q_n) + 1} \mathbb{P}(R_n^B \geq a^{\frac{1}{p}} \mu_1(n, q_n)). \end{aligned}$$

Using Chernoff’s bound together with the moment generating function of R_n^B from Proposition 1.11, we have

$$\mathbb{P}(R_n^B \geq a^{\frac{1}{p}} \mu_1(n, q_n)) \leq e^{-ta^{\frac{1}{p}} \mu_1(n, q_n)} \mathbb{E}[e^{tR_n^B}] \leq \exp(-ta^{\frac{1}{p}} \mu_1(n, q_n) + (e^t - 1)\mu_1(n, q_n)).$$

Taking $t = \log 2 + 1$ and using this bound in (2.1), it follows that

$$\mathbb{P}\left(\left(\frac{h(T_{n,q_n})}{\mu(n, q_n)}\right)^p \geq a\right) \leq 2 \exp(-[a^{\frac{1}{p}} + 1 - 2e]\mu_1(n, q_n)).$$

Thus, for $a > (2e - 1)^p$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\left(\frac{h(T_{n,q_n})}{\mu(n, q_n)}\right)^p \geq a\right) = 0,$$

which proves the claimed uniform integrability. \square

This result implies that the convergence in L_p in Theorem 1.1 follows from the convergence in probability. Therefore, from now on we can only focus our attention on proving the latter type of convergence.

Using a similar argument as for the previous proof, we conclude this section by proving the upper tail bound of Proposition 2.1.

PROOF OF PROPOSITION 2.1 (UPPER TAIL). Fix $n \geq 0$ and $h = h_n > \mu_1(n, q_n)$. By combining Proposition 1.6 with a union bound, we have

$$\begin{aligned} \mathbb{P}(h(T_{n,q_n}) \geq h) &= \mathbb{P}(\exists v \in T_\infty : |v| = h, v \in T_{n,q_n}) \\ &\leq 2^h \mathbb{P}(\bar{1}^h \in T_{n,q_n}) \\ &= 2^h \mathbb{P}(R_n^B \geq h). \end{aligned}$$

Because $h > \mu_1(n, q_n)$, it follows from Lemma 2.4 that

$$\mathbb{P}(R_n^B \geq h) \leq \exp\left(\left[\frac{h}{\mu_1(n, q_n)} \log\left(\frac{e\mu_1(n, q_n)}{h}\right) - 1\right] \mu_1(n, q_n)\right),$$

hence

$$(2.2) \quad \mathbb{P}(h(T_{n,q_n}) \geq h) \leq \exp\left(h \log\left(\frac{2e\mu_1(n, q_n)}{h}\right) - \mu_1(n, q_n)\right).$$

Fix a sequence $(\gamma_n)_{n \geq 0}$ such that $\gamma_n = \omega(n(1 - q_n) \vee \sqrt{\log n})$ and with $\gamma_n = o(\log n)$; this second hypothesis on $(\gamma_n)_{n \geq 0}$ suffices to prove Proposition 2.1 as the probabilities we aim to bound are decreasing functions of γ_n . By Proposition 2.5, we have $\mu_1(n, q_n) = n(1 - q_n) + \log n + O(\sqrt{\log n}) = \log n + o(\gamma_n)$. Taking $h_n = \lfloor c^* \log n + \gamma_n \rfloor$ and since $c^* > 1$, we thus have that $h_n > \mu_1(n, q_n)$ for n large enough, so (2.2) gives

$$(2.3) \quad \mathbb{P}(h(T_{n,q_n}) \geq h_n) \leq \exp\left(h_n \log\left(\frac{2e\mu_1(n, q_n)}{h_n}\right) - \mu_1(n, q_n)\right).$$

To bound the right-hand side of (2.3), we note that $c^* \mu_1(n, q_n) = c^* \log n + o(\gamma_n) = h_n - (1 + o(1))\gamma_n$, so

$$\begin{aligned} \log\left(\frac{2e\mu_1(n, q_n)}{h_n}\right) &= \log\left(\frac{2e}{c^*} \cdot \frac{c^* \mu_1(n, q_n)}{h_n}\right) \\ &= \log\left(\frac{2e}{c^*}\right) + \log\left(1 - (1 + o(1))\frac{\gamma_n}{h_n}\right) \\ &= \frac{1}{c^*} - (1 + o(1))\frac{\gamma_n}{h_n}, \end{aligned}$$

the last identity holding since $c^* \log(\frac{2e}{c^*}) = 1$ and since $\gamma_n = o(h_n)$. Together with (2.3), this yields

$$\mathbb{P}(h(T_{n,q_n}) \geq h_n) \leq \exp\left(\frac{h_n}{c^*} - (1 + o(1))\gamma_n - \mu_1(n, q_n)\right) = \exp\left(\left(\frac{1}{c^*} - 1 + o(1)\right)\gamma_n\right).$$

Because $h_n \leq c^* \log n + \gamma_n$ and $\gamma_n \rightarrow \infty$ this concludes the proof of the upper bound. \square

2.3. Lower tail bound. We now prove the lower tail bound of Proposition 2.1. In order to do so, we use the coupling explained in Section 1.3: $S_{n,q}(\emptyset) = n$ and given $S_{n,q}(v)$, we have

$$S_{n,q}(v\bar{0}) = \begin{cases} \left\lfloor \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)})}{\log q} \right\rfloor & \text{if } q \in (0, 1), \\ \lfloor S_{n,q}(v)U_v \rfloor & \text{if } q = 1, \\ 0 & \text{if } q = 0 \end{cases}$$

and

$$S_{n,q}(v\bar{1}) = S_{n,q}(v) - 1 - S_{n,q}(v\bar{0}).$$

As we saw in Section 1.3, it follows from Proposition 1.5 that $\{v \in T_\infty : S_{n,q}(v) \geq 1\}$ is $MT(n, q)$ -distributed.

For any node $v \in T_\infty$, write $X_{(v,v\bar{0})} = U_v$ and $X_{(v,v\bar{1})} = 1 - U_v$. and let $P_v = \prod_{e < v} X_e$, where $e < v$ denotes the set of edges on the path from \emptyset to v . The following proposition gives useful bounds for $S_{n,q}$ using P .

LEMMA 2.8. *Let $n \geq 1$ and $q \in (0, 1)$. Then, a.s. for all $v \in T_\infty$, we have*

$$\frac{\log(1 - P_v(1 - q^n))}{\log q} - |v| \leq S_{n,q}(v) \leq n - \frac{\log(q^n + P_v(1 - q^n))}{\log q}.$$

Moreover, these inequalities naturally extend to $q = 1$ as follows: almost surely

$$nP_v - |v| \leq S_{n,1}(v) \leq nP_v.$$

PROOF. We only consider the case $q \in (0, 1)$, since in the case $q = 1$ the assertion was already proven in [11], and can be obtained from our formula by taking the limit $q \rightarrow 1$.

We prove this lemma by induction on the depth of v . For $v = \emptyset$, $P_v = 1$ and the inequalities hold. Assume now it holds for some $v \in T_\infty$. We start by proving that the lower bound holds for the left child of v . By definition, we know that

$$S_{n,q}(v\bar{0}) = \left\lfloor \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)}))}{\log q} \right\rfloor \geq \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)}))}{\log q} - 1,$$

and the right-hand side is increasing in $S_{n,q}(v)$. By the induction hypothesis, we know that $S_{n,q}(v) \geq \frac{\log(1 - P_v(1 - q^n))}{\log q} - |v|$. Replacing $S_{n,q}(v)$ by its lower bound into the previous formula gives us

$$\begin{aligned} S_{n,q}(v\bar{0}) &\geq \frac{\log(1 - U_v(1 - q^{\frac{\log(1 - P_v(1 - q^n))}{\log q} - |v|}))}{\log q} - 1 \\ &= \frac{\log(1 - U_v + U_v q^{-|v|}(1 - P_v(1 - q^n)))}{\log q} - 1. \end{aligned}$$

Since $\log q < 0$ and $(1 - U_v) \leq q^{-|v|}(1 - U_v)$, it follows that

$$\begin{aligned} S_{n,q}(v\bar{0}) &\geq \frac{\log(q^{-|v|}(1 - U_v) + U_v q^{-|v|}(1 - P_v(1 - q^n)))}{\log q} - 1 \\ &= \frac{\log(1 - U_v P_v(1 - q^n))}{\log q} - |v| - 1, \end{aligned}$$

which is the desired lower bound.

We now prove that the upper bound holds for the right child of v . Using the definition again, we have

$$\begin{aligned} S_{n,q}(v\bar{1}) &= S_{n,q}(v) - 1 - S_{n,q}(v\bar{0}) \\ &= S_{n,q}(v) - 1 - \left\lfloor \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)}))}{\log q} \right\rfloor \\ &= \left\lceil S_{n,q}(v) - \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)}))}{\log q} \right\rceil - 1 \\ &\stackrel{\text{a.s.}}{=} \left\lceil S_{n,q}(v) - \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)}))}{\log q} \right\rceil, \end{aligned}$$

the last equality following from the fact that $\frac{\log(1 - U_v(1 - q^{S_{n,q}(v)}))}{\log q}$ is a.s. not an integer. From this equality, we obtain

$$S_{n,q}(v\bar{1}) \leq S_{n,q}(v) - \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)}))}{\log q},$$

and the right-hand side is increasing in $S_{n,q}(v)$, which can be checked by direct computation. Applying the induction hypothesis, we obtain

$$\begin{aligned} S_{n,q}(v\bar{1}) &\leq n - \frac{\log(q^n + P_v(1 - q^n))}{\log q} - \frac{\log(1 - U_v(1 - q^{n - \frac{\log(q^n + P_v(1 - q^n))}{\log q}}))}{\log q} \\ &= n - \frac{\log(q^n + (1 - U_v)P_v(1 - q^n))}{\log q}, \end{aligned}$$

which is the desired upper bound.

For the two remaining bounds, note that, for any integer $S \geq 1$ and $U \in (0, 1)$, we have

$$\begin{aligned}
 (2.4) \quad & \left(S - \frac{\log(1 - U(1 - q^S))}{\log q} \right) - \left(\frac{\log(1 - (1 - U)(1 - q^S))}{\log q} \right) \\
 &= \frac{S \log q - \log(q^S + U(1 - U)(1 - q^S)^2)}{\log q} \\
 &= \frac{1}{\log \frac{1}{q}} \log \left(1 + \frac{U(1 - U)(1 - q^S)^2}{q^S} \right) \geq 0.
 \end{aligned}$$

From this inequality, it follows that

$$\begin{aligned}
 S_{n,q}(v\bar{1}) &= \left\lfloor S_{n,q}(v) - \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)}))}{\log q} \right\rfloor \\
 &\geq \left\lfloor \frac{\log(1 - (1 - U_v)(1 - q^{S_{n,q}(v)}))}{\log q} \right\rfloor,
 \end{aligned}$$

and then the same technique as the one used to bound $S_{n,q}(v\bar{0})$ from below gives us that

$$S_{n,q}(v\bar{1}) \geq \frac{\log(1 - (1 - U_v)P_v(1 - q^n))}{\log q} - |v| - 1,$$

which yields the desired lower bound. Similarly, (2.4) gives that

$$\begin{aligned}
 S_{n,q}(v\bar{0}) &= \left\lfloor \frac{\log(1 - U_v(1 - q^{S_{n,q}(v)}))}{\log q} \right\rfloor \\
 &\leq \left\lfloor S_{n,q}(v) - \frac{\log(1 - (1 - U_v)(1 - q^{S_{n,q}(v)}))}{\log q} \right\rfloor,
 \end{aligned}$$

from which we deduce that

$$S_{n,q}(v\bar{0}) \leq n - \frac{\log(q^n + U_v P_v(1 - q^n))}{\log q}$$

by the same technique used to bound $S_{n,q}(v\bar{1})$ from above. This last inequality concludes the induction and the proof of the lemma. \square

With this results, we can now compare $(S_{n,q})$ for different values of n and q using P . This leads to Proposition 1.7.

PROOF OF PROPOSITION 1.7. We must prove that

$$\mathbb{P}(h(T_{n,q}) \leq \ell) \leq \mathbb{P}(h(T_{m,1}) \leq \ell)$$

where $m = \lfloor \frac{1 - q^n}{1 - q^{\ell+2}} \rfloor$. First of all, if $q = 0$, then $h(T_{n,q}) = n$ and $m = \lfloor \frac{1 - q^n}{1 - q^{\ell+2}} \rfloor = 1$, so the inequality holds. Assume now that $q \in (0, 1)$. Because $\{v \in T_\infty : S_{n,q}(v) \geq 1\}$ is $\text{MT}(n, q)$ -distributed, we have

$$\mathbb{P}(h(T_{n,q}) \leq \ell) = \mathbb{P}(\forall v \in T_\infty \text{ with } |v| = \ell + 1 : S_{n,q}(v) < 1).$$

Using the lower bound in Lemma 2.8, this implies that

$$\begin{aligned}
 \mathbb{P}(h(T_{n,q}) \leq \ell) &\leq \mathbb{P}\left(\forall v \in T_\infty \text{ with } |v| = \ell + 1 : \frac{\log(1 - P_v(1 - q^n))}{\log q} - |v| < 1\right) \\
 &= \mathbb{P}\left(\forall v \in T_\infty \text{ with } |v| = \ell + 1 : P_v < \frac{1 - q^{\ell+2}}{1 - q^n}\right).
 \end{aligned}$$

Consider now some $m \geq 1$ and use the same technique but with the upper bound of Lemma 2.8:

$$\begin{aligned} \mathbb{P}(h(T_{m,1}) \leq \ell) &= \mathbb{P}(\forall v \in T_\infty \text{ with } |v| = \ell + 1 : S_{m,1}(v) < 1) \\ &\geq \mathbb{P}\left(\forall v \in T_\infty \text{ with } |v| = \ell + 1 : P_v < \frac{1}{m}\right). \end{aligned}$$

The proposition now follows by defining $m = \lfloor \frac{1-q^n}{1-q^{\ell+2}} \rfloor$, since $\frac{1}{m} \geq \frac{1-q^{\ell+2}}{1-q^n}$. \square

We now have all the results necessary to prove the lower tail bound of Proposition 2.1.

PROOF OF PROPOSITION 2.1 (LOWER TAIL). Consider $(q_n)_{n \geq 0}$ and $(\gamma_n)_{n \geq 0}$ as in the proposition and let $m_n = \lfloor \frac{1-(q_n)^n}{1-(q_n)^{\ell_n+2}} \rfloor$ where $\ell_n = \lfloor c^* \log n - \gamma_n \rfloor$. We again assume without loss of generality that $\gamma_n = o(\log n)$.

Let us first study the asymptotic behaviour of $\log m_n$. Note that $\lfloor \frac{1-q^n}{1-q^{\ell+2}} \rfloor$ is monotone in q provided $\ell + 2 \leq n$. Because $n(1 - q_n) < \log n$ for n large enough, it follows that

$$\frac{1 - (1 - \frac{\log n}{n})^n}{1 - (1 - \frac{\log n}{n})^{\ell_n+2}} - 1 \leq m_n \leq \lim_{q \rightarrow 1} \frac{1 - q^n}{1 - q^{\ell+2}} = \frac{n}{\ell_n + 2}.$$

Moreover, $(1 - \frac{\log n}{n})^n = o(1)$ and $(1 - \frac{\log n}{n})^{\ell_n+2} = e^{-(1+o(1))\frac{c^*(\log n)^2}{n}} = 1 - (1 + o(1)) \times \frac{c^*(\log n)^2}{n}$, which implies that

$$\log\left(\frac{1 - (1 - \frac{\log n}{n})^n}{1 - (1 - \frac{\log n}{n})^{\ell_n+2}} - 1\right) = \log\left((1 + o(1))\frac{n}{c^*(\log n)^2} - 1\right) = \log n + O(\log \log n).$$

It follows that

$$\log n + O(\log \log n) \leq \log m_n \leq \log n - \log(\ell_n + 2) = \log n + O(\log \log n),$$

hence $\log m_n = \log n + O(\log \log n)$.

Let us now study $(T_{m_n,1})_{n \geq 0}$. Recall that when $q = 1$, a $MT(n, q)$ -distributed tree has the distribution of a random binary search tree. We use the results of Reed [30] and Drmota [13], who prove that

$$\mathbb{E}[h(T_{m_n,1})] = c^* \log m_n + O(\log \log m_n)$$

and

$$\text{Var}[h(T_{m_n,1})] = O(1).$$

Next,

$$\mathbb{P}(h(T_{m_n,1}) \leq \ell_n) = \mathbb{P}(h(T_{m_n,1}) - \mathbb{E}[h(T_{m_n,1})] \leq \ell_n - \mathbb{E}[h(T_{m_n,1})])$$

and

$$\ell_n - \mathbb{E}[h(T_{m_n,1})] = c^* \log n - \gamma_n - c^* \log m_n + O(\log \log m_n) = (-1 + o(1))\gamma_n,$$

the last equality holding since $\gamma_n = \omega(\sqrt{\log n}) = \omega(\log \log n)$. Applying Chebyshev's inequality, we obtain

$$\mathbb{P}(h(T_{m_n,1}) \leq \ell_n) \leq \frac{\text{Var}[h(T_{m_n,1})]}{(\ell_n - \mathbb{E}[h(T_{m_n,1})])^2} = O\left(\frac{1}{(\gamma_n)^2}\right) = o(1).$$

The lower bound of Proposition 2.1 follows by applying Proposition 1.7 to obtain

$$\mathbb{P}(h(T_{n,q_n}) \leq c^* \log n - \gamma_n) = \mathbb{P}(h(T_{n,q_n}) \leq \ell_n) \leq \mathbb{P}(h(T_{m_n,1}) \leq \ell_n) = o(1). \quad \square$$

3. Right depth of Mallows trees. In this section, we study the right depth R_n^B of a Mallows tree and use its properties to prove Theorem 1.2.

3.1. *The bivariate moment generating function.* In this section, we prove Proposition 1.11, which we recall provides a formula for the bivariate moment generating function of R_n^B and M_n^B .

PROOF OF PROPOSITION 1.11. Fix $q \in [0, 1)$ and let $B = (B_{i,j})_{i,j \geq 1}$ have independent BERNOLLI($1 - q$) entries. Let us first study the transition probabilities of the random process $(R_n^B, M_n^B)_{n \geq 0}$. Define $\mathcal{L}_n = \sigma(B_{i,j}, 1 \leq i \leq n, j \geq 1)$ for the σ -algebra generated by the first n rows of the matrix B and note that R_n^B and M_n^B are \mathcal{L}_n -measurable. Since $(R_n^B)_{n \geq 0}$ corresponds to the number of records of the sequence $(M_n^B)_{n \geq 0}$, we have

$$\begin{aligned}
 (3.1) \quad & \mathbb{P}(R_{n+1}^B = r, M_{n+1}^B = m \mid R_n^B = r, M_n^B = m, \mathcal{L}_n) \\
 &= \mathbb{P}(f^B(n+1) < M_n^B \mid R_n^B = r, M_n^B = m, \mathcal{L}_n) \\
 &= \mathbb{P}(\exists j \in [M_n^B - 1] \setminus F_n^B, B_{n+1,j} = 1 \mid R_n^B = r, M_n^B = m, \mathcal{L}_n) \\
 &= 1 - \mathbb{P}(\forall j \in [M_n^B - 1] \setminus F_n^B, B_{n+1,j} = 0 \mid R_n^B = r, M_n^B = m, \mathcal{L}_n).
 \end{aligned}$$

Moreover, since R_n^B, M_n^B and F_n^B are \mathcal{L}_n -measurable, and $(B_{n+1,j})_{j \geq 1}$ is independent of \mathcal{L}_n , we have

$$\begin{aligned}
 (3.2) \quad & \mathbb{P}(\forall j \in [M_n^B - 1] \setminus F_n^B, B_{n+1,j} = 0 \mid R_n^B = r, M_n^B = m, \mathcal{L}_n) \\
 &= \mathbb{E}[q^{|[M_n^B - 1] \setminus F_n^B|} \mid R_n^B = r, M_n^B = m, \mathcal{L}_n] \\
 &= q^{m-n};
 \end{aligned}$$

the final equality holding since

$$\begin{aligned}
 & |[M_n^B - 1] \setminus F_n^B| = |[M_n^B - 1] \setminus \{f^B(1), \dots, f^B(n)\}| \\
 &= |[M_n^B] \setminus \{f^B(1), \dots, f^B(n)\}| \\
 &= M_n^B - n.
 \end{aligned}$$

Combining (3.1) and (3.2) shows that the desired transition probability has the claimed value when $k = 0$ and $\ell = 0$. Similarly, for $\ell \geq 1$, we have

$$\begin{aligned}
 & \mathbb{P}(R_{n+1}^B = r + 1, M_{n+1}^B = m + \ell \mid R_n^B = r, M_n^B = m, \mathcal{L}_n) \\
 &= \mathbb{P}(f^B(n+1) = m + \ell \mid R_n^B = r, M_n^B = m, \mathcal{L}_n) \\
 &= \mathbb{P}(\forall j \in [M_n^B + \ell - 1] \setminus F_n^B, B_{n+1,j} = 0 \text{ and } B_{n+1, M_n^B + \ell} = 1 \mid R_n^B = r, M_n^B = m, \mathcal{L}_n) \\
 &= q^{m+\ell-n-1}(1 - q).
 \end{aligned}$$

This establishes the Markov property and proves that the transition probability is as claimed.

We now move on to the derivation of the bivariate characteristic function, and we do so by induction. The equation does hold when $n = 0$, as can be straightforwardly checked, but for the sake of the proof, it is more natural to start the induction at $n = 1$. In this case, we have

$$\begin{aligned}
 \mathbb{E}[x^{R_1^B+1}y^{M_1^B}] &= \sum_{j \geq 1} \mathbb{E}[x^{R_1^B+1}y^{M_1^B} \mid f^B(1) = j] \mathbb{P}(f^B(1) = j) \\
 &= \sum_{j \geq 1} x^1 y^j q^{j-1} (1 - q)
 \end{aligned}$$

$$\begin{aligned} &= \frac{xy(1-q)}{1-xy} \\ &= y^1 \frac{q + (1-q)x - q^1}{1-q^1y}, \end{aligned}$$

which is the desired formula. Assume now that the formula holds for some $n \geq 1$. First, by conditioning on (R_n^B, M_n^B) , we have

$$(3.3) \quad \mathbb{E}[x^{R_{n+1}^B+1}y^{M_{n+1}^B}] = \sum_{r,m} \mathbb{E}[x^{R_{n+1}^B+1}y^{M_{n+1}^B} \mid R_n^B = r, M_n^B = m] \mathbb{P}(R_n^B = r, M_n^B = m).$$

For the first term inside the sum, we have

$$\begin{aligned} &\mathbb{E}[x^{R_{n+1}^B+1}y^{M_{n+1}^B} \mid R_n^B = r, M_n^B = m] \\ &= \sum_{k \in \{0,1\}} \sum_{\ell \geq 0} x^{r+k+1}y^{m+\ell} \mathbb{P}(R_{n+1}^B = r+k, M_{n+1}^B = m+\ell \mid R_n^B = r, M_n^B = m). \end{aligned}$$

We now split the sum according to whether $\ell = 0$ or $\ell \geq 1$. Note that if $\ell = 0$, then $k = 0$, and if $\ell \geq 1$, then $k = 1$. Using the previously derived transition probabilities, we obtain

$$\begin{aligned} &\mathbb{E}[x^{R_{n+1}^B+1}y^{M_{n+1}^B} \mid R_n^B = r, M_n^B = m] \\ &= x^{r+1}y^m(1-q^{m-n}) + \sum_{\ell \geq 1} x^{r+2}y^{m+\ell}q^{m+\ell-n-1}(1-q) \\ &= x^{r+1}y^m(1-q^{m-n}) + \frac{x^{r+2}y^{m+1}q^{m-n}(1-q)}{1-xy}. \end{aligned}$$

Plugging this back into (3.3), we have

$$\begin{aligned} &\mathbb{E}[x^{R_{n+1}^B+1}y^{M_{n+1}^B}] \\ &= \sum_{r,m} \left[x^{r+1}y^m(1-q^{m-n}) + \frac{x^{r+2}y^{m+1}q^{m-n}(1-q)}{1-xy} \right] \mathbb{P}(R_n^B = r, M_n^B = m) \\ &= \mathbb{E}[x^{R_n^B+1}y^{M_n^B}] - \frac{1}{q^n} \mathbb{E}[x^{R_n^B+1}(qy)^{M_n^B}] + \frac{xy(1-q)}{q^n(1-xy)} \mathbb{E}[x^{R_n^B+1}(qy)^{M_n^B}] \\ &= \mathbb{E}[x^{R_n^B+1}y^{M_n^B}] + \frac{1}{q^n} \left[\frac{xy(1-q)}{1-xy} - 1 \right] \mathbb{E}[x^{R_n^B+1}(qy)^{M_n^B}]. \end{aligned}$$

Using the induction hypothesis, this yields the equation

$$\begin{aligned} &\mathbb{E}[x^{R_{n+1}^B+1}y^{M_{n+1}^B}] \\ &= \left[y^n \prod_{1 \leq k \leq n} \frac{q + (1-q)x - q^k}{1 - q^k y} \right] + \frac{1}{q^n} \left[\frac{xy(1-q)}{1-xy} - 1 \right] \\ &\quad \times \left[(qy)^n \prod_{1 \leq k \leq n} \frac{q + (1-q)x - q^k}{1 - q^k(qy)} \right] \\ &= y^n \prod_{1 \leq k \leq n} \frac{q + (1-q)x - q^k}{1 - q^k y} + \frac{(xy(1-q) - 1 + qy)y^n}{1-xy} \prod_{1 \leq k \leq n} \frac{q + (1-q)x - q^k}{1 - q^{k+1}y}. \end{aligned}$$

Finally, factorizing by $y^n \prod_{1 \leq k \leq n} \frac{q+(1-q)x-q^k}{1-q^k y}$, it follows that

$$\begin{aligned} \mathbb{E}[x^{R_{n+1}^B+1} y^{M_{n+1}^B}] &= \left[y^n \prod_{1 \leq k \leq n} \frac{q+(1-q)x-q^k}{1-q^k y} \right] \left[1 + \frac{xy(1-q)-1+qy}{1-q^{n+1}y} \right] \\ &= \left[y^n \prod_{1 \leq k \leq n} \frac{q+(1-q)x-q^k}{1-q^k y} \right] \left[\frac{y(q+(1-q)x-q^{n+1})}{1-q^{n+1}y} \right] \\ &= y^{n+1} \prod_{1 \leq k \leq n+1} \frac{q+(1-q)x-q^k}{1-q^k y}, \end{aligned}$$

establishing the second assertion of the proposition. \square

With this result under our belt, we are able to prove bounds on their height in the case when $n(1-q_n)/\log n \rightarrow \infty$.

3.2. *Bounds on the left subtrees.* Before proving Theorem 1.2, we prove Proposition 1.12, whose bounds will be useful in in the proof of all three theorems of this paper.

PROOF OF PROPOSITION 1.12. To begin, let us briefly explain why

$$(3.4) \quad M := c^* \times \sup_{n \geq 2, q \in (0, 1)} \left\{ \frac{\mu_1(n, q) - n(1-q) - \log(n \wedge \frac{1}{1-q})}{\sqrt{\log(n \wedge \frac{1}{1-q})}} \right\}.$$

is finite. Assuming it is not, we can find a sequence $(n_k, q_k)_{k \geq 0}$ such that the term inside the supremum of (3.4) goes to infinity. By extracting subsequences, we can assume that $(q_k)_{k \geq 0}$ converges in $[0, 1]$ and that $(n_k)_{k \geq 0}$ is constant equal to some $n \geq 2$ or diverges to ∞ .

If $n_k = n$, the numerator is bounded and the only way for this term to diverge is to have $\log(n \wedge \frac{1}{1-q_k}) \rightarrow 0$, hence $q_k \rightarrow 0$. In this case, we have

$$\begin{aligned} \mu_1(n_k, q_k) - n_k(1-q_k) - \log\left(n_k \wedge \frac{1}{1-q_k}\right) &= \sum_{1 \leq i \leq n} \frac{1-q_k}{1-q_k^i} - n(1-q_k) - \log\left(\frac{1}{1-q_k}\right) \\ &= O(q_k), \end{aligned}$$

and since $\log(n_k \wedge \frac{1}{1-q_k}) \sim q_k$, the term inside the supremum of (3.4) actually converges to 0 and not infinity. On the other hand, if $n_k \rightarrow \infty$, we can apply Proposition 2.5 to bound the numerator in (3.4) and see that the supremum is again finite.

We will prove the bound stated in the proposition for this value of M . By the definition of M , for all $n \geq 2$ and $q \in (0, 1)$, we have

$$\mu_1(n, q) - n(1-q) - \log\left(n \wedge \frac{1}{1-q}\right) \leq \frac{M}{c^*} \sqrt{\log\left(n \wedge \frac{1}{1-q}\right)};$$

moreover, this inequality remains true when $n = 1$ or $q = 0$ as both sides equal 0.

Now, without loss of generality, fix $\xi > 0$, write

$$\mathbb{P}_{LS} := \mathbb{P}\left(\sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} \geq c^* \log\left(\frac{1}{1-q}\right) + M \sqrt{\log\left(\frac{1}{1-q}\right)} + \xi\right),$$

and let $\ell = \lceil c^* \log(\frac{1}{1-q}) + M \sqrt{\log(\frac{1}{1-q})} + \xi \rceil$. By taking a union bound over k , we obtain

$$(3.5) \quad \mathbb{P}_{LS} \leq \sum_{k \geq 0} \mathbb{P}(h(T^B(\bar{1}^k \bar{0})) \geq \ell + k).$$

By Lemma 1.10, $T^B(\bar{1}^k\bar{0}) \stackrel{d}{=} T_{G(q),q}$, so

$$\mathbb{P}(h(T^B(\bar{1}^k\bar{0})) \geq \ell + k) = \mathbb{P}(h(T_{G(q),q}) \geq \ell + k).$$

Consider now some $\lambda > 0$ and divide the probability according to whether $G(q) \geq \frac{\lambda(k+\xi)+2-q}{1-q}$ or $G(q) < \frac{\lambda(k+\xi)+2-q}{1-q}$, to obtain

$$\begin{aligned} \mathbb{P}(h(T^B(\bar{1}^k\bar{0})) \geq \ell + k) &= \mathbb{P}\left(h(T_{G(q),q}) \geq \ell + k, G(q) \geq \frac{\lambda(k+\xi)+2-q}{1-q}\right) \\ (3.6) \qquad \qquad \qquad &+ \mathbb{P}\left(h(T_{G(q),q}) \geq \ell + k, G(q) < \frac{\lambda(k+\xi)+2-q}{1-q}\right). \end{aligned}$$

For the first term, drop the first event and use that $G(q)$ is geometric to obtain

$$\mathbb{P}\left(h(T_{G(q),q}) \geq \ell + k, G(q) \geq \frac{\lambda(k+\xi)+2-q}{1-q}\right) \leq q^{\frac{\lambda(k+\xi)+2-q}{1-q}} \leq e^{-\lambda(k+\xi)-1},$$

where the last inequality follows from the fact that $0 \leq q \leq e^{q-1}$. For the second term, use the increasing property of $(T_{n,q})_{n \geq 0}$ from Corollary 1.9 to bound $G(q)$ by its maximal value, and then drop the second event, to obtain

$$\mathbb{P}\left(h(T_{G(q),q}) \geq \ell + k, G(q) < \frac{\lambda(k+\xi)+2-q}{1-q}\right) \leq \mathbb{P}(h(T_{m,q}) \geq \ell + k),$$

where $m = \lfloor \frac{\lambda(k+\xi)+2-q}{1-q} \rfloor$. Putting this back into (3.6) gives us

$$(3.7) \qquad \mathbb{P}(h(T^B(\bar{1}^k\bar{0})) \geq \ell + k) \leq e^{-\lambda(k+\xi)-1} + \mathbb{P}(h(T_{m,q}) \geq \ell + k).$$

In order to bound \mathbb{P}_{LS} , we now need to bound $\mathbb{P}(h(T_{m,q}) \geq \ell + k)$. Taking a union bound over all nodes at depth $\ell + k$ and then applying Proposition 1.6, we obtain

$$\mathbb{P}(h(T_{m,q}) \geq \ell + k) \leq 2^{\ell+k} \mathbb{P}(R_m^B \geq \ell + k).$$

Using Chernoff's bound with the moment generating function of R_m^B from Proposition 1.11, it follows that, for any $t > 0$,

$$\mathbb{P}(R_m^B \geq \ell + k) \leq e^{-t(\ell+k)} \mathbb{E}[e^{tR_m^B}] \leq \exp(-t(\ell+k) + (e^t - 1)\mu_1(m, q)).$$

Putting this back into the previous inequality and taking $t = \log c^*$ gives us

$$\mathbb{P}(h(T_{m,q}) \geq \ell + k) \leq \exp\left(\log\left(\frac{2}{c^*}\right)(\ell+k) + (c^* - 1)\mu_1(m, q)\right).$$

Since $m \geq \frac{1}{1-q}$ and by using the definition of M , we know that

$$\mu_1(m, q) \leq m(1-q) + \log\left(\frac{1}{1-q}\right) + \frac{M}{c^*} \sqrt{\log\left(\frac{1}{1-q}\right)}.$$

From this inequality, we obtain

$$\begin{aligned} \mathbb{P}(h(T_{m,q}) \geq \ell + k) &\leq \exp\left(\log\left(\frac{2}{c^*}\right)(\ell+k) + (c^* - 1)\left(m(1-q) + \log\left(\frac{1}{1-q}\right) + \frac{M}{c^*} \sqrt{\log\left(\frac{1}{1-q}\right)}\right)\right). \end{aligned}$$

From their respective definitions, we know that $\ell \geq c^* \log\left(\frac{1}{1-q}\right) + M\sqrt{\log\left(\frac{1}{1-q}\right)} + \xi$ and that $m \leq \frac{\lambda(k+\xi)+2-q}{1-q}$. Using these bounds together with the fact that $\log\left(\frac{2}{c^*}\right) = \frac{1}{c^*} - 1 < 0$ in the preceding inequality, we obtain

$$\begin{aligned} \mathbb{P}(h(T_{m,q}) \geq \ell + k) &\leq \exp\left(-\frac{c^* - 1}{c^*} \left(c^* \log\left(\frac{1}{1-q}\right) + M\sqrt{\log\left(\frac{1}{1-q}\right)} + \xi + k\right)\right) \\ &\quad + (c^* - 1) \left(\lambda(k + \xi) + 2 - q + \log\left(\frac{1}{1-q}\right) + \frac{M}{c^*} \sqrt{\log\left(\frac{1}{1-q}\right)}\right) \\ &\leq \exp\left(-\frac{(c^* - 1)(1 - c^*\lambda)}{c^*} (\xi + k) + 2(c^* - 1)\right). \end{aligned}$$

Using this bound in (3.7), and then plugging the result into (3.5), we obtain

$$\mathbb{P}_{\text{LS}} \leq \sum_{k \geq 0} \left[\exp(-\lambda(k + \xi) - 1) + \exp\left(-\frac{(c^* - 1)(1 - c^*\lambda)}{c^*} (\xi + k) + 2(c^* - 1)\right) \right].$$

Choosing $\lambda = \frac{c^* - 1}{(c^*)^2}$ so that $\lambda = \frac{(c^* - 1)(1 - c^*\lambda)}{c^*}$, this bound becomes

$$\mathbb{P}_{\text{LS}} \leq e^{-\lambda\xi} \sum_{k \geq 0} [\exp(-\lambda k - 1) + \exp(-\lambda k + 2(c^* - 1))] = \left(\frac{e^{-1} + e^{2(c^* - 1)}}{1 - e^{-\lambda}}\right) e^{-\lambda\xi},$$

proving the proposition. \square

3.3. Almost sure convergence. To conclude this section, we prove Theorem 1.2, from which we deduce Theorem 1.1 in the case when $n(1 - q_n)/\log n \rightarrow \infty$ by combining it with Proposition 2.7.

PROOF OF THEOREM 1.2. Let $(q_n)_{n \geq 0}$ be a sequence such that $n(1 - q_n)/\log n \rightarrow \infty$ and fix $\varepsilon > 0$ and $\lambda > 0$. We prove that

$$\mathbb{P}\left(\left|\frac{h(T_{n,q_n})}{n(1 - q_n)} - 1\right| > \varepsilon\right) = O\left(\frac{1}{n^\lambda}\right)$$

by bounding the lower tail and the upper tail separately.

We start with the lower tail as the technique for the upper tail bound is similar but more involved. Using R_n^B as a stochastic lower bound for $h(T_{n,q_n})$ as in (1.2), we have

$$\mathbb{P}(h(T_{n,q_n}) < (1 - \varepsilon)n(1 - q_n)) \leq \mathbb{P}(R_n^B < (1 - \varepsilon)n(1 - q_n)).$$

Since $n(1 - q_n) = \omega(\log n)$, by applying Proposition 2.6 we know that $\mu_1(n, q_n) \sim n(1 - q_n)$. It follows that, for n large enough, we have $(1 - \varepsilon)n(1 - q_n)/\mu_1(n, q_n) < 1$ and we can apply the second bound of Lemma 2.4 to the previous inequality to obtain

$$\begin{aligned} &\mathbb{P}(h(T_{n,q_n}) < (1 - \varepsilon)n(1 - q_n)) \\ &\leq \exp\left(\left[\frac{(1 - \varepsilon)n(1 - q_n)}{\mu_1(n, q_n)} \log\left(\frac{e\mu_1(n, q_n)}{(1 - \varepsilon)n(1 - q_n)}\right) - 1\right] \mu_1(n, q_n)\right) \\ &= \exp\left(\left[(1 - \varepsilon) \log\left(\frac{e}{1 - \varepsilon}\right) - 1 + o(1)\right] n(1 - q_n)\right). \end{aligned}$$

By convexity, $(1 - \varepsilon) \log\left(\frac{e}{1 - \varepsilon}\right) < 1$. Since $n(1 - q_n) = \omega(\log n)$, this proves that

$$\mathbb{P}(h(T_{n,q_n}) < (1 - \varepsilon)n(1 - q_n)) = O\left(\frac{1}{n^\gamma}\right),$$

which is the desired lower bound for $h(T_{n,q_n})$.

Let us now prove that the upper bound also holds. Using the second stochastic inequality given in (1.2), we have

$$\mathbb{P}(h(T_{n,q_n}) > (1 + \varepsilon)n(1 - q_n)) \leq \mathbb{P}\left(R_n^B + 1 + \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} > (1 + \varepsilon)n(1 - q_n)\right).$$

Using that $X + Y > x + y$ implies that $X > x$ or $Y > y$ to bound the probability on the right, we obtain

$$\begin{aligned} \mathbb{P}(h(T_{n,q_n}) > (1 + \varepsilon)n(1 - q_n)) &\leq \mathbb{P}(R_n^B > (1 + \varepsilon/2)n(1 - q_n)) \\ &\quad + \mathbb{P}\left(1 + \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} > \varepsilon n(1 - q_n)/2\right). \end{aligned}$$

For the first term, using the first bound of Lemma 2.4 and the same arguments as for the lower bound, it follows that

$$\mathbb{P}(R_n^B > (1 + \varepsilon/2)n(1 - q_n)) = O\left(\frac{1}{n^\nu}\right).$$

For the second term, apply Proposition 1.12 with

$$\xi = \xi_n = \varepsilon n(1 - q_n)/2 - 1 - c^* \log\left(\frac{1}{1 - q_n}\right) - M \sqrt{\log\left(\frac{1}{1 - q_n}\right)}$$

to obtain

$$\mathbb{P}\left(1 + \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} > \varepsilon n(1 - q_n)/2\right) \leq C e^{-\lambda \xi_n}.$$

Since $n(1 - q_n)/\log n \rightarrow \infty$, it follows that $\log \frac{1}{1 - q_n} = O(\log n)$ and then $\xi_n \sim \varepsilon n(1 - q_n)/2$, which proves that

$$\mathbb{P}\left(1 + \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} > \varepsilon n(1 - q_n)/2\right) = O\left(\frac{1}{n^\nu}\right).$$

This yields the desired upper bound for $h(T_{n,q_n})$ and concludes the proof of Theorem 1.2. □

4. Intermediate values and threshold process. In this section, we will prove Proposition 1.13, which will allow us to conclude the proof of Theorem 1.1 in the last case, that is, when $n(1 - q_n)/\log n = \Theta(1)$. In order to do so, we prove the following proposition, which in fact handles a somewhat wider range of asymptotic behaviour for the sequence $(q_n)_{n \geq 0}$. The bounds in Proposition 4.1, below, are actually tight enough that they will also be used in Section 5.1 to prove Proposition 1.15, which is a key input of the central limit theorem for the height of Mallows trees.

PROPOSITION 4.1. *Let $(q_n)_{n \geq 0}$ be such that $\log(n(1 - q_n)) = O(\sqrt{\log n})$ and $n(1 - q_n) = \omega(\sqrt{\log n})$. For $n \geq 0$, let $m = m(n) = \min\{\ell \geq 0 : \ell(1 - q_n) + \log \ell \geq n(1 - q_n)\}$. Then, for all $(\beta_n)_{n \geq 0}$ such that $\beta_n = \omega(\sqrt{\log n})$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{P}(e^{-\beta_n} \leq |T_n^B(\bar{1}^{R_m^B+1})|(1 - q_n) \leq \beta_n) = 1.$$

From this proposition, with $m = m(n)$ as previously defined, we will show that

$$|T_n^B(\bar{1}^{R_m^B+1})|(1 - q_n) = o_{\mathbb{P}}(\log(|T_n^B(\bar{1}^{R_m^B+1})|))$$

which allows us to apply Proposition 2.1 to prove that

$$h(T_n^B(\bar{1}^{R_m^B+1})) = (c^* + o_{\mathbb{P}}(1)) \log n.$$

It is not hard to show that $R_m^B = (1 + o_{\mathbb{P}}(1))n(1 - q_n)$. Once we establish this, we will be able to prove Proposition 1.13 by comparing the sizes of $T_n^B(\bar{1}^{R_m^B+1})$ and $T_n^B(\bar{1}^{\lfloor n(1-q_n) \rfloor})$.

In order to study the size of $T_n^B(\bar{1}^{R_m^B+1})$, recall the definition of $(T_n^B)_{n \geq 0}$ from Section 1.4 and note that, for all $d \geq 0$,

$$|T_n^B(\bar{1}^{d+1})| = |\{i \in [n] : f^B(i) > \tau^B(\bar{1}^d)\}|,$$

where τ^B is the labelling function of the tree T^B . Moreover, since $\tau^B(\bar{1}^{R_m^B}) = M_m^B$ where $M_m^B = \max(f^B(i), i \in [m])$, we have

$$|T_n^B(\bar{1}^{R_m^B+1})| = |\{i \in [n] : f^B(i) > M_m^B\}|.$$

In order to study the size of the random set on the right-hand side of this equation, we define the *threshold process* N^B as follows. For all $n \in \mathbb{N}$ and $s \in \mathbb{N}$, let

$$(4.1) \quad N^B(n, s) := |\{i \in [n] : f^B(i) > s\}|.$$

The two preceding displays show that

$$|T_n^B(\bar{1}^{R_m^B+1})| = N^B(n, M_m^B),$$

which will be our key tool for bounding $|T_n^B(\bar{1}^{R_m^B+1})|$. The analysis of this identity is made easier by the following lemma, which partially decouples M^B and N^B .

LEMMA 4.2. *For all $n \geq 0$, $q \in [0, 1)$ and $0 \leq m \leq n$, we have*

$$|T_n^B(\bar{1}^{R_m^B+1})| \stackrel{d}{=} N^{B^*}(n - m, M_m^B - m),$$

where $B = (B_{i,j})$ and $B^* = (B_{i,j}^*)$ are independent matrices, each with independent BERNULLI($1 - q$) entries.

Using this distributional identity, the proof of Proposition 4.1 will be divided into the following three steps:

Step 1 We study the threshold process $N^B(n, s)$ for all values of n and s and prove bounds for its upper and lower tail probabilities (see Propositions 4.5 and 4.6).

Step 2 We prove that $M_{m(n)}^B = n + O_{\mathbb{P}}(\sqrt{\log n}/(1 - q_n))$ using Proposition 1.11.

Step 3 We combine these two results to prove Proposition 4.1.

The rest of the section is organized as follows. The results of steps 1, 2 and 3 are respectively stated and proven in Sections 4.1, 4.2 and 4.3. Finally, in Section 4.4, we prove Proposition 1.13 and deduce Theorem 1.1 in the case when $n(1 - q_n)/\log n = \Theta(1)$.

4.1. *Threshold process.* This section is focused on the behaviour of $N^B(n, s)$ as defined in (4.1). From the definition, we see that N^B is increasing in n and decreasing in s . Since f^B is bijective, it is also straightforward to verify that $(n - s)_+ \leq N(n, s) \leq n$ where $x_+ = \max(x, 0)$. The next proposition gives further properties related to the distribution of N .

PROPOSITION 4.3. Fix $q \in [0, 1)$, and let $B = (B_{i,j})_{i,j \geq 0}$ be an array of independent BERNOULLI($1 - q$) random variables. For $k \geq 0$, write $\mathcal{L}_k(B) = \sigma(B_{i,j}, 1 \leq i \leq k, j \geq 1)$ for the σ -algebra generated by the first k rows of B . Then, for any integers $n, s \geq 1$, for all $\ell \geq 0$, we have

$$\mathbb{P}(N^B(n, s) = \ell \mid f^B(1) \leq s, \mathcal{L}_k(B)) \stackrel{d}{=} \mathbb{P}(N^B(n - 1, s - 1) = \ell \mid \mathcal{L}_{k-1}(B))$$

and

$$\mathbb{P}(N^B(n, s) = \ell \mid f^B(1) > s, \mathcal{L}_k(B)) \stackrel{d}{=} \mathbb{P}(N^B(n - 1, s) = \ell - 1 \mid \mathcal{L}_{k-1}(B)).$$

PROOF. Write $B^{(i,j)}$ for the minor of B obtained by deleting the i th row and the j th column, and note that $B^{(i,j)} \stackrel{d}{=} B$. Moreover, given that $f^B(1) = r$, the rest of the values $f^B(2), f^B(3), \dots$ becomes independent of the first row and the r th column of B . Hence, for $1 \leq r \leq s$, we have

$$\begin{aligned} \mathbb{P}(N^B(n, s) = \ell \mid f^B(1) = r, \mathcal{L}_k(B)) &= \mathbb{P}(N^{B^{(1,r)}}(n - 1, s - 1) = \ell \mid \mathcal{L}_k(B)) \\ &\stackrel{d}{=} \mathbb{P}(N^B(n - 1, s - 1) = \ell \mid \mathcal{L}_{k-1}(B)), \end{aligned}$$

where the second equality holds since $\mathcal{L}_k(B)$ is generated by $\mathcal{L}_{k-1}(B^{(1,r)})$ and $\sigma(\{B_{1,j}, j \geq 1\} \cup \{B_{i,r}, i \geq 1\})$, and $B^{(1,r)}$ is independent of the second of these σ -algebras. Similarly, for $r > s$, we have

$$\begin{aligned} \mathbb{P}(N^B(n, s) = \ell \mid f^B(1) = r, \mathcal{L}_k(B)) &= \mathbb{P}(N^{B^{(1,r)}}(n - 1, s) = \ell - 1 \mid \mathcal{L}_k(B)) \\ &\stackrel{d}{=} \mathbb{P}(N^B(n - 1, s) = \ell - 1 \mid \mathcal{L}_{k-1}(B)). \end{aligned}$$

This proves the two desired equalities. \square

Applying this proposition, we can now prove Lemma 4.2.

PROOF OF LEMMA 4.2. By definition, we know that

$$|T_n^B(\bar{1}^{R_m^B+1})| = N^B(n, M_m^B).$$

Let $s \geq 1$ be an integer. Conditioning on the value of M_m^B and applying Proposition 4.3 with $k = 1$, we obtain

$$\begin{aligned} \mathbb{P}(N^B(n, M_m^B) = \ell \mid M_m^B = s) &= \mathbb{P}(N^B(n, s) = \ell \mid M_m^B = s, f^B(1) \leq s) \\ &= \mathbb{E}[\mathbb{P}(N^B(n, s) = \ell \mid M_m^B = s, f^B(1) \leq s, \mathcal{L}_1)] \\ &= \mathbb{P}(N^B(n - 1, s - 1) = \ell \mid M_{m-1}^B = s - 1). \end{aligned}$$

Applying this identity $m - 1$ times, we obtain that

$$\mathbb{P}(N^B(n, M_m^B) = \ell \mid M_m^B = s) = \mathbb{P}(N^B(n - m + 1, s - m + 1) = \ell \mid M_1^B = s - m + 1).$$

For the last step, since $\{f^B(1) \leq s - m + 1, M_1^B = s - m + 1\} = \{f^B(1) = s - m + 1\}$, it follows that

$$\begin{aligned} \mathbb{P}(N^B(n, M_m^B) = \ell \mid M_m^B = s) &= \mathbb{P}(N^B(n - m + 1, s - m + 1) = \ell \mid M_1^B = s - m + 1, f^B(1) \leq s - m + 1) \\ &= \mathbb{P}(N^B(n - m + 1, s - m + 1) = \ell \mid f^B(1) = s - m + 1) \\ &= \mathbb{P}(N^B(n - m, s - m) = \ell). \end{aligned}$$

Thus, the proof of the lemma is immediate by observing that

$$\mathbb{P}(N^{B^*}(n - m, M_m^B - m) = \ell \mid M_m^B = s) = \mathbb{P}(N^{B^*}(n - m, s - m) = \ell). \quad \square$$

In order to bound the size of $T_n^B(\bar{1}^{R_m^B+1})$ using the threshold process, we now state and prove an exact formula for the probability mass function of $N^B(n, s)$. For the remainder of the section, we write $N(n, s) = N^B(n, s)$.

PROPOSITION 4.4. *Let $n, s \geq 0$. Then, for all $\ell \in \mathbb{N}$, we have*

$$\mathbb{P}(N(n, s) = \ell) = q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\prod_{s+1-n+\ell \leq i \leq s} (1 - q^i) \right) \sum_{A \subseteq [n]: |A|=\ell} q^{\sum_{a \in A} (a-1)}.$$

PROOF. First, note that the right-hand side of the equality is 0 if $\ell \leq n - s - 1$ or $\ell \geq n + 1$ since either the product $\prod_{s+1-n+\ell \leq i \leq s} (1 - q^i)$ equals 0 or the sum $\sum_{A \subseteq [n]: |A|=\ell}$ is empty. For such ℓ , $\mathbb{P}(N(n, s) = \ell) = 0$ as well, so the claimed equality holds when $\ell \leq n - s - 1$ or $\ell > n$. We now prove that the equality holds for $(n - s)_+ \leq \ell \leq n$ by induction on $n + s$.

First, if $n + s = 0$, then $N(n, s) = 0$ and the right-hand side is equal to 1 if and only if $\ell = 0$, which proves the formula.

Fix some $n, s \geq 0$ and assume the formula holds for any n', s' such that $n' + s' < n + s$. Let ℓ be such that $(n - s)_+ \leq \ell \leq n$. By considering the possible values for $f^B(1)$ and using the two formulas in Proposition 4.3, we obtain

$$\begin{aligned} \mathbb{P}(N(n, s) = \ell) &= \mathbb{P}(N(n, s) = \ell \mid f^B(1) \leq s) \mathbb{P}(f^B(1) \leq s) \\ &\quad + \mathbb{P}(N(n, s) = \ell \mid f^B(1) > s) \mathbb{P}(f^B(1) > s) \\ &= \mathbb{P}(N(n - 1, s - 1) = \ell) \mathbb{P}(f^B(1) \leq s) \\ &\quad + \mathbb{P}(N(n - 1, s) = \ell - 1) \mathbb{P}(f^B(1) > s). \end{aligned}$$

Using the definition of f^B , we know that

$$\mathbb{P}(f^B(1) \leq s) = \sum_{1 \leq j \leq s} \mathbb{P}(f^B(1) = j) = 1 - q^s$$

and

$$\mathbb{P}(f^B(1) > s) = q^s.$$

This gives us that

$$(4.2) \quad \mathbb{P}(N(n, s) = \ell) = (1 - q^s) \mathbb{P}(N(n - 1, s - 1) = \ell) + q^s \mathbb{P}(N(n - 1, s) = \ell - 1).$$

Using the induction hypothesis, we know that

$$\mathbb{P}(N(n - 1, s - 1) = \ell) = q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\prod_{s+1-n+\ell \leq i \leq s-1} (1 - q^i) \right) \sum_{A \subseteq [n-1]: |A|=\ell} q^{\sum_{a \in A} (a-1)}.$$

Hence, multiplying by $(1 - q^s)$ on both sides and putting it into the product, we obtain

$$\begin{aligned} (1 - q^s) \mathbb{P}(N(n - 1, s - 1) = \ell) \\ = q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\prod_{s+1-n+\ell \leq i \leq s} (1 - q^i) \right) \sum_{A \subseteq [n-1]: |A|=\ell} q^{\sum_{a \in A} (a-1)}. \end{aligned}$$

Similarly, we have

$$\begin{aligned} q^s \mathbb{P}(N(n-1, s) = \ell - 1) &= q^s \cdot q^{(s-n+1)(\ell-1) + \frac{\ell(\ell-1)}{2}} \left(\prod_{s+1-n+\ell \leq i \leq s} (1 - q^i) \right) \sum_{A \subseteq [n-1]; |A| = \ell - 1} q^{\sum_{a \in A} (a-1)} \\ &= q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\prod_{s+1-n+\ell \leq i \leq s} (1 - q^i) \right) \sum_{A \subseteq [n-1]; |A| = \ell - 1} q^{n-1 + \sum_{a \in A} (a-1)}. \end{aligned}$$

Putting the previous formulas into (4.2), we obtain

$$\begin{aligned} \mathbb{P}(N(n, s) = \ell) &= q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\prod_{s+1-n+\ell \leq i \leq s} (1 - q^i) \right) \sum_{A \subseteq [n-1]; |A| = \ell} q^{\sum_{a \in A} (a-1)} \\ &\quad + q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\prod_{s+1-n+\ell \leq i \leq s} (1 - q^i) \right) \sum_{A \subseteq [n-1]; |A| = \ell - 1} q^{n-1 + \sum_{a \in A} (a-1)}. \end{aligned}$$

In order to conclude, note that

$$\sum_{A \subseteq [n-1]; |A| = \ell - 1} q^{n-1 + \sum_{a \in A} (a-1)} = \sum_{A \subseteq [n-1]; |A| = \ell - 1} q^{\sum_{a \in A \cup \{n\}} (a-1)},$$

which implies that

$$\sum_{A \subseteq [n-1]; |A| = \ell} q^{\sum_{a \in A} (a-1)} + \sum_{A \subseteq [n-1]; |A| = \ell - 1} q^{n-1 + \sum_{a \in A} (a-1)} = \sum_{A \subseteq [n]; |A| = \ell} q^{\sum_{a \in A} (a-1)},$$

and this proves the desired formula for $\mathbb{P}(N(n, s) = \ell)$. The induction and the proposition follow. \square

We conclude this section on the threshold process with upper and lower tail bounds for N . Both of these bounds use the following inequalities:

$$(4.3) \quad \sum_{A \subseteq [n]; |A| = \ell} q^{\sum_{a \in A} (a-1)} \leq \frac{1}{\ell!} \sum_{a_1, \dots, a_\ell \geq 0} q^{a_1 + \dots + a_\ell} \leq \frac{1}{\ell!} \left(\frac{1}{1 - q} \right)^\ell.$$

PROPOSITION 4.5. *Let $n, s \geq 0$. For any integer ξ such that $q^{\xi+s-n} \leq \xi(1 - q)$, we have*

$$\mathbb{P}(N(n, s) \geq \xi) \leq \frac{n}{\xi!} q^{\frac{\xi^2}{2}} \left(\frac{q^{s-n}}{1 - q} \right)^\xi.$$

PROOF. This proof will be very straightforward using (4.3). Fix some ξ respecting the given conditions. If $\xi > n$, the inequality holds as the left-hand side equals 0. Assume now that $\xi \leq n$. Using Proposition 4.4, we have

$$\mathbb{P}(N(n, s) \geq \xi) = \sum_{\xi \leq \ell \leq n} q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\prod_{s+1-n+\ell \leq i \leq s} (1 - q^i) \right) \sum_{A \subseteq [n]; |A| = \ell} q^{\sum_{a \in A} (a-1)}.$$

Applying (4.3) along with the fact that $(1 - q^i) \leq 1$ when $s + 1 - n + \ell \leq i \leq s$, it follows that

$$(4.4) \quad \mathbb{P}(N(n, s) \geq \xi) \leq \sum_{\xi \leq \ell \leq n} q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \frac{1}{\ell!} \left(\frac{1}{1 - q} \right)^\ell = \sum_{\xi \leq \ell \leq n} \frac{q^{\frac{\ell(\ell+1)}{2}}}{\ell!} \left(\frac{q^{s-n}}{1 - q} \right)^\ell.$$

To conclude the proof, we now show that the summands on the right are decreasing in ℓ . To see this, note that

$$\frac{q^{\frac{\ell(\ell+1)}{2}}}{\ell!} \left(\frac{q^{s-n}}{1-q}\right)^\ell = \frac{q^{\ell+s-n}}{\ell(1-q)} \times \frac{q^{\frac{\ell(\ell-1)}{2}}}{(\ell-1)!} \left(\frac{q^{s-n}}{1-q}\right)^{\ell-1} \leq \frac{q^{\frac{\ell(\ell-1)}{2}}}{(\ell-1)!} \left(\frac{q^{s-n}}{1-q}\right)^{\ell-1},$$

the last inequality holding since the function $\ell \mapsto \frac{q^\ell}{\ell}$ is decreasing and since we assumed that $q^{\xi+s-n} \leq \xi(1-q)$. Bounding all summands on the right-hand side of (4.4) by the $\ell = \xi$ term, we obtain

$$\mathbb{P}(N(n, s) \geq \xi) \leq (n - \xi + 1) \frac{q^{\frac{\xi(\xi+1)}{2}}}{\xi!} \left(\frac{q^{s-n}}{1-q}\right)^\xi \leq \frac{n}{\xi!} q^{\frac{\xi^2}{2}} \left(\frac{q^{s-n}}{1-q}\right)^\xi,$$

where the second inequality uses that $n - \xi + 1 \leq n$ and $q^{\frac{\xi}{2}} \leq 1$. \square

PROPOSITION 4.6. *Let $n, s \geq 0$. For any integer ξ such that $q^{\xi+s-n} \geq \xi(1-q)^2$, we have*

$$\mathbb{P}(N(n, s) \leq \xi) \leq \frac{2(1-q)^{n-s}}{((\xi-1)_+)((s-n)_+)!} q^{\frac{\xi^2}{2}} \left(\frac{q^{s-n}}{(1-q)^2}\right)^\xi.$$

PROOF. The proof will be very similar to the previous one as we will first bound the probability with a sum over ℓ and then consider the largest term. Fix some ξ respecting the given condition. If $\xi < (n-s)_+$, then the inequality holds as the left-hand side equals 0. Assume now that $\xi \geq (n-s)_+$. First, by applying Proposition 4.4 along with (4.3), we obtain

$$\begin{aligned} &\mathbb{P}(N(n, s) \leq \xi) \\ (4.5) \quad &= \sum_{(n-s)_+ \leq \ell \leq \xi} q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\prod_{s+1-n+\ell \leq i \leq s} (1-q^i)\right) \sum_{A \subseteq [n]: |A|=\ell} q^{\sum_{a \in A} (a-1)} \\ &\leq \sum_{(n-s)_+ \leq \ell \leq \xi} q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \frac{1}{\ell!} \left(\frac{1}{1-q}\right)^\ell \left(\prod_{s+1-n+\ell \leq i \leq s} (1-q^i)\right). \end{aligned}$$

Write $\tilde{s} = \min(s, \lfloor \frac{1}{1-q} \rfloor)$ and use that $1 - q^i \leq 1$ and that $1 - q^i \leq i(1 - q)$, to obtain

$$\begin{aligned} \prod_{s+1-n+\ell \leq i \leq s} (1 - q^i) &\leq \prod_{s+1-n+\ell \leq i \leq \tilde{s}} i(1 - q) = \frac{\tilde{s}!(1 - q)^{\tilde{s}-s+n-\ell}}{(s - n + \ell)!} \\ &\leq \frac{1}{((s - n)_+)!} \left(\frac{1}{1 - q}\right)^{s-n+\ell}, \end{aligned}$$

where the last inequality follows from the bounds $(s - n + \ell)! \geq ((s - n)_+)$! and

$$\tilde{s}!(1 - q)^{\tilde{s}} = \prod_{1 \leq k \leq \tilde{s}} [k(1 - q)] \leq 1.$$

Put this bound back into (4.5) to obtain

$$(4.6) \quad \mathbb{P}(N(n, s) \leq \xi) \leq \frac{(1 - q)^{n-s}}{((s - n)_+)!} \sum_{(n-s)_+ \leq \ell \leq \xi} \frac{1}{\ell!} q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\frac{1}{1 - q}\right)^{2\ell}.$$

Looking for the largest term in the sum again, we note that

$$\frac{1}{\ell!} q^{(s-n)\ell + \frac{\ell(\ell+1)}{2}} \left(\frac{1}{1 - q}\right)^{2\ell} = \frac{q^{s-n+\ell}}{\ell(1 - q)^2} \times \frac{1}{(\ell - 1)!} q^{(s-n)(\ell-1) + \frac{\ell(\ell-1)}{2}} \left(\frac{1}{1 - q}\right)^{2(\ell-1)},$$

and $q^{s-n+\ell} \geq \ell(1-q)^2$ for all $\ell \leq \xi$ by using the assumption on ξ . This implies that we can bound all terms in the sum in (4.6) from above by the $\ell = \xi$ term, and obtain

$$\mathbb{P}(N(n, s) \leq \xi) \leq [\xi + 1 - (n - s)_+] \frac{(1 - q)^{n-s}}{((s - n)_+)! \xi!} q^{(s-n)\xi + \frac{\xi(\xi+1)}{2}} \left(\frac{1}{1 - q}\right)^{2\xi}.$$

The desired bound follows by using that $\xi + 1 - (n - s)_+ \leq 2\xi$ and $q^{\frac{\xi}{2}} \leq 1$. \square

4.2. *Bounds on M_m^B .* For the remainder of Section 4, we define

$$m = m(n) = \min\{\ell \geq 0 : \ell(1 - q_n) + \log \ell \geq n(1 - q_n)\}.$$

Under the assumption that $n(1 - q_n) = \omega(\sqrt{\log n})$, this definition implies that

$$(4.7) \quad \mu_1(m(n), q_n) \sim n(1 - q_n),$$

by using the asymptotic estimate for $\mu_1(m(n), q_n)$ from Proposition 2.5.

The goal of this section is to prove the following proposition.

PROPOSITION 4.7. *Let $(q_n)_{n \geq 0}$ be such that $\log(n(1 - q_n)) = O(\sqrt{\log n})$ and $n(1 - q_n) = \omega(\sqrt{\log n})$. For $n \geq 0$, let $m = m(n) = \min\{\ell \geq 0 : \ell(1 - q_n) + \log \ell \geq n(1 - q_n)\}$. Then, for all $(\alpha_n)_{n \geq 0}$ such that $\alpha_n = \omega(\sqrt{\log n})$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(|M_{m(n)}^B - n| > \frac{\alpha_n}{1 - q_n}\right) = 0.$$

PROOF. For the rest of the proof, we omit B and n from the notation. We also write $p = p_n = 1 - q_n$. By assumption, $\log(np) = O(\sqrt{\log n})$, so $p \rightarrow 0$ and also $np^2 \rightarrow 0$. We will use these facts in the proof.

We prove the proposition by establishing upper and lower tail bounds for M_m separately. For both bounds, we will use the following inequality obtained from the moment generating function of M_m given in Proposition 1.11: for all $t \in \mathbb{R}$ such that $qe^t < 1$, we have

$$(4.8) \quad \mathbb{E}[e^{tM_m}] = \prod_{1 \leq k \leq m} \left(1 + (e^t - 1) \frac{1}{1 - q^k e^t}\right) \leq \exp\left((e^t - 1) \sum_{1 \leq k \leq m} \frac{1}{1 - q^k e^t}\right).$$

For the upper tail bound and for all $t > 0$ such that $qe^t < 1$, use Markov’s inequality and (4.8) to obtain

$$\mathbb{P}\left(M_m > n + \frac{\alpha}{p}\right) \leq \exp\left(-t \left[n + \frac{\alpha}{p}\right] + (e^t - 1) \sum_{1 \leq k \leq m} \frac{1}{1 - q^k e^t}\right).$$

Let $t = -\frac{1}{2} \log q$, so that $e^t = 1/\sqrt{q}$. Using that $q^k e^t = q^{k-\frac{1}{2}} \leq q^{k-1}$, it follows that

$$\sum_{1 \leq k \leq m} \frac{1}{1 - q^k e^t} \leq \frac{1}{1 - \sqrt{q}} + \sum_{2 \leq k \leq m} \frac{1}{1 - q^{k-1}} = \frac{1}{1 - \sqrt{q}} + \frac{1}{p} + \frac{1}{p} \mu_1(m - 1, q).$$

From (4.7), we know that $\mu_1(m - 1, q) = \mu_1(m, q) - \frac{1-q}{1-q^m} = np + O(\sqrt{\log n})$, since $\frac{1-q}{1-q^m} \leq 1$; this implies that

$$\sum_{1 \leq k \leq m} \frac{1}{1 - q^k e^t} \leq n + O\left(\frac{\sqrt{\log n}}{p}\right).$$

Moreover, since $p = 1 - q \rightarrow 0$, we have $t = \frac{\alpha}{2} + O(p^2)$ and $e^t - 1 = \frac{\alpha}{2} + O(p^2)$. Combining this with the previous bound on the sum, we obtain

$$\begin{aligned} \mathbb{P}\left(M_m > n + \frac{\alpha}{p}\right) &\leq \exp\left(\left(\frac{p}{2} + O(p^2)\right)\left[-n - \frac{\alpha}{p} + n + O\left(\frac{\sqrt{\log n}}{p}\right)\right]\right) \\ &= \exp\left(-\frac{\alpha}{2} + o(\alpha)\right); \end{aligned}$$

the last equality follows from the fact that $np^2 \rightarrow 0$ and that $\alpha = \omega(\sqrt{\log n})$. This proves the desired upper tail bound for M_m .

For the lower tail bound, using Markov’s inequality and (4.8) again, for all $t > 0$ we have

$$\mathbb{P}\left(M_m < n - \frac{\alpha}{p}\right) \leq \exp\left(t\left[n - \frac{\alpha}{p}\right] + (e^{-t} - 1) \sum_{1 \leq k \leq m} \frac{1}{1 - q^k e^{-t}}\right).$$

Let $t = -\log q$, so $e^{-t} = q$. Arguing similarly to the proof of the upper tail, we see that $t = p + O(p^2)$ and $e^{-t} - 1 = -p$. Moreover, by (4.7), it follows that

$$\sum_{1 \leq k \leq m} \frac{1}{1 - q^k e^{-t}} = \sum_{2 \leq k \leq m+1} \frac{1}{1 - q^k} = \frac{1}{p} \mu_1(m + 1, q) = n + O\left(\frac{\sqrt{\log n}}{p}\right).$$

This gives us that

$$\mathbb{P}\left(M_m < n - \frac{\alpha}{p}\right) \leq \exp\left((p + O(p^2))\left[n - \frac{\alpha}{p} - n + O\left(\frac{\sqrt{\log n}}{p}\right)\right]\right) = \exp(-\alpha + o(\alpha)),$$

which concludes the proof of the lower tail bound and the proposition. \square

4.3. *Bounds on $|T_n^B(\bar{1}^{R_m^B+1})|$.* With the results from the two previous sections, we now have all the tools required to prove Proposition 4.1.

PROOF OF PROPOSITION 4.1. By using the distributional identity from Lemma 4.2, it suffices to prove that

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{e^{-\beta_n}}{1 - q_n} \leq N^{B^*}(n - m(n), M_{m(n)}^B - m(n)) \leq \frac{\beta_n}{1 - q_n}\right) = 1.$$

We write $p_n = 1 - q_n$ and note that $p_n \rightarrow 0$ as $n \rightarrow \infty$. We now prove that the corresponding upper and lower tail probabilities converge to 0 as $n \rightarrow \infty$:

$$\text{UB} := \mathbb{P}\left(N^{B^*}(n - m(n), M_{m(n)}^B - m(n)) > \frac{\beta_n}{p_n}\right) \rightarrow 0$$

and

$$\text{LB} := \mathbb{P}\left(N^{B^*}(n - m(n), M_{m(n)}^B - m(n)) < \frac{e^{-\beta_n}}{p_n}\right) \rightarrow 0.$$

For the rest of the proof, $\alpha = (\alpha_n)_{n \geq 0}$ refers to a sequence such that $\alpha_n = \omega(\sqrt{\log n})$ and $\alpha_n = o(\beta_n)$. In other words, we have $\sqrt{\log n} \ll \alpha_n \ll \beta_n$. From now on, we omit n and the superscript B and B^* from the notation, since the random variables $N = (N(n, s))_{n, s \geq 0} = (N^{B^*}(n, s))_{n, s \geq 0}$ are independent of the random variables $M = (M_m)_{m \geq 0} = (M_m^B)_{m \geq 0}$.

For the upper tail, divide the probability according to the values of M_m as follows:

$$\begin{aligned} \text{UB} &= \mathbb{P}\left(N(n - m, M_m - m) > \frac{\beta}{p}, M_m > n - \frac{\alpha}{p}\right) \\ &\quad + \mathbb{P}\left(N(n - m, M_m - m) > \frac{\beta}{p}, M_m \leq n - \frac{\alpha}{p}\right). \end{aligned}$$

Applying Proposition 4.7, we know that the second term converges to 0. Recall now that $N(n, s)$ is decreasing in s . By independence of N and M , taking $s = \lfloor n - \frac{\alpha}{p} \rfloor = n - \lceil \frac{\alpha}{p} \rceil$, it follows that

$$(4.9) \quad \mathbb{P}\left(N(n - m, M_m - m) > \frac{\beta}{p}, M_m \geq n - \frac{\alpha}{p}\right) \leq \mathbb{P}\left(N(n - m, s - m) > \frac{\beta}{p}\right).$$

Write $\xi = \lfloor \frac{\beta}{p} \rfloor$. We now need to verify that $q^{\xi+(s-m)-(n-m)} \leq \xi(1 - q)$, so that Proposition 4.5 applies. For this, since $\xi \sim \frac{\beta}{p}$ and $(s - m) - (n - m) \sim -\frac{\alpha}{p}$, and using that $\alpha = o(\beta)$ and that $q^{\frac{1}{p}} = (1 - p)^{\frac{1}{p}} \rightarrow e^{-1}$, we have

$$\frac{q^{\xi+(s-m)-(n-m)}}{\xi p} = (1 + o(1)) \frac{e^{-(\beta-\alpha)(1+o(1))}}{\beta} = o(1),$$

which proves that this ratio is less than 1 for n large enough. This means that we can indeed apply Proposition 4.5 and obtain

$$\begin{aligned} \mathbb{P}\left(N(n - m, s - m) > \frac{\beta}{p}\right) &\leq \frac{n - m}{\xi!} q^{\frac{\xi^2}{2}} \left(\frac{q^{s-n}}{p}\right)^\xi \\ &= \exp\left(\log(n - m) - \log \xi! + \xi \left[\frac{\xi}{2} + (s - n)\right] \log q - \xi \log p\right). \end{aligned}$$

By the definition of ξ and s , we know that

$$-\log \xi! - \xi \log p = -\xi \log \xi + O(\log \xi) - \xi \log p \sim -\xi \log \beta$$

and

$$\xi \left[\frac{\xi}{2} + (s - n)\right] \log q \sim \frac{\xi^2}{2} \log q \sim -\frac{\xi^2 p}{2} \sim -\frac{\xi \beta}{2}.$$

Since $\log(n - m) \leq \log n = o(\xi \beta)$ and $\xi \log \beta = o(\xi \beta)$, this implies that

$$\mathbb{P}\left(N(n - m, s - m) > \frac{\beta}{p}\right) \leq \exp\left(- (1 + o(1)) \frac{\xi \beta}{2}\right) = o(1).$$

Plugging this result back into (4.9) proves the upper tail bound of the proposition.

For the lower tail bound, we similarly divide the probability to obtain

$$\begin{aligned} \text{LB} &= \mathbb{P}\left(N(n - m, M_m - m) < \frac{e^{-\beta}}{p}, M_m \leq n + \frac{\alpha}{p}\right) \\ (4.10) \quad &+ \mathbb{P}\left(N(n - m, M_m - m) < \frac{e^{-\beta}}{p}, M_m > n + \frac{\alpha}{p}\right) \\ &\leq \mathbb{P}\left(N(n - m, s - m) < \frac{e^{-\beta}}{p}\right) + o(1), \end{aligned}$$

where $s = \lfloor n + \frac{\alpha}{p} \rfloor = n + \lfloor \frac{\alpha}{p} \rfloor$. Write $\xi = \lfloor \frac{e^{-\beta}}{p} \rfloor$. To verify that the requirement of Proposition 4.6 that $q^{\xi+(s-m)-(n-m)} \geq \xi(1 - q)^2$ is satisfied, note that

$$\frac{q^{\xi+(s-m)-(n-m)}}{\xi p^2} = (1 + o(1)) \frac{e^{(1+o(1))(\beta-\alpha)}}{p} \rightarrow \infty,$$

so is larger than 1 for n large enough. For such n , applying Proposition 4.6, we obtain

$$\begin{aligned} \mathbb{P}\left(N(n - m, s - m) < \frac{e^{-\beta}}{p}\right) &\leq \frac{2p^{n-s}}{((\xi - 1)_+)((s - n)_+)!} q^{\frac{\xi^2}{2}} \left(\frac{q^{s-n}}{p^2}\right)^\xi \\ &\leq 2 \frac{p^{n-s-2\xi}}{(s - n)!} \\ &= 2 \exp((n - s - 2\xi) \log p - \log(s - n)!). \end{aligned}$$

Using the definition of ξ and s , we know that

$$(n - s) \log p - \log(s - n)! = (n - s) \log(p(s - n)) + O(s - n) \sim -\frac{\alpha \log \alpha}{p}$$

and

$$2\xi \log p \sim -\frac{2e^{-\beta} \log n}{p} = o\left(\frac{1}{p}\right).$$

This implies that

$$\mathbb{P}\left(N(n - m, s - m) < \frac{e^{-\beta}}{p}\right) \leq 2 \exp\left(-\left(1 + o(1)\right) \frac{\alpha \log \alpha}{p}\right) = o(1),$$

which proves the desired lower tail bound by plugging this result back into (4.10). This concludes the proof of the proposition. \square

4.4. *Convergence in probability.* We conclude this section with the proof of Proposition 1.13 and then Theorem 1.1. We start with two straightforward lemmas.

LEMMA 4.8. *Let $(q_n)_{n \geq 0}$ be such that $\log(n(1 - q_n)) = O(\sqrt{\log n})$ and $n(1 - q_n) = \omega(\sqrt{\log n})$. For $n \geq 0$, let $m = m(n) = \min\{\ell \geq 0 : \ell(1 - q_n) + \log \ell \geq n(1 - q_n)\}$. Then we have*

$$\frac{R_{m(n)}^B}{n(1 - q_n)} \longrightarrow 1$$

in probability as $n \rightarrow \infty$.

PROOF. By applying Lemma 2.4, and since $c \log(\frac{c}{c-1}) < 1$ for all $c \neq 1$, we have

$$\frac{R_{m(n)}^B}{\mu_1(m(n), q_n)} \xrightarrow{\mathbb{P}} 1.$$

Moreover, since $n(1 - q_n) = \omega(\sqrt{\log n})$, by (4.7), we have that $\mu_1(m(n), q_n) \sim n(1 - q_n)$, which proves the desired result. \square

LEMMA 4.9. *Let $n \geq 0$ and $q \in (0, 1)$. Then, for all integers $d \geq 0$ and $\ell \geq 0$, we have*

$$|T_{n,q}(\bar{1}^{d+\ell})| \geq \frac{\log(1 - P_\ell(1 - q^{|T_{n,q}(\bar{1}^d)|}))}{\log q} - d,$$

where P_ℓ is distributed as a product of ℓ independent $\text{UNIFORM}([0, 1])$, and is independent of $|T_{n,q}(\bar{1}^d)|$.

PROOF. This is simply a restatement of the lower bound from Lemma 2.8, when applied to the tree $T_{n,q}(\bar{1}^d)$. By Proposition 1.5, $T_{n,q}(\bar{1}^d)$ is Mallows tree once conditioned on its size, so this application of Lemma 2.8 is indeed valid. \square

The proof of Proposition 1.13 now follows from combining Lemmas 4.8 and 4.9 with Proposition 4.1.

PROOF OF PROPOSITION 1.13. Since T_n^B has the same distribution as $T_{n,q}$, we can prove the proposition by showing that

$$h(T_n^B(\bar{1}^{\lfloor n(1-q_n) \rfloor})) = (c^* + o_{\mathbb{P}}(1)) \log n.$$

In order to prove this asymptotic result, we will show that $N = |T_n^B(\bar{1}^{\lfloor n(1-q_n) \rfloor})|$ satisfies that $N(1 - q_n) = o_{\mathbb{P}}(\log N)$ and that $\log N = (1 + o_{\mathbb{P}}(1)) \log n$. This will allow us to apply Proposition 2.1 to conclude the proof. For the remainder of the proof, we drop the subscript n on q_n and write $m = m(n) = \min\{\ell \geq 0 : \ell(1 - q) + \log \ell \geq n(1 - q)\}$.

We start with the upper bound on the size of the tree. In the case where $R_m^B > \lfloor n(1 - q) \rfloor$, we have that

$$\begin{aligned} T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor}) &= T_n^B(\bar{1}^{R_m^B}) \cup \{\bar{1}^k : \lfloor n(1 - q) \rfloor \leq k < R_m^B\} \cup \bigcup_{\lfloor n(1-q) \rfloor \leq k < R_m^B} T_n^B(\bar{1}^k \bar{0}) \\ &\subseteq T_n^B(\bar{1}^{R_m^B}) \cup \{\bar{1}^k : \lfloor n(1 - q) \rfloor \leq k < R_m^B\} \cup \bigcup_{\lfloor n(1-q) \rfloor \leq k < R_m^B} T^B(\bar{1}^k \bar{0}). \end{aligned}$$

Moreover, this inclusion remains true when $R_m^B \leq \lfloor n(1 - q) \rfloor$, since

$$T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor}) \subseteq T_n^B(\bar{1}^{R_m^B}).$$

Overall, we obtain that

$$|T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})| \leq |T_n^B(\bar{1}^{R_m^B})| + (R_m^B - \lfloor n(1 - q) \rfloor)_+ + \sum_{\lfloor n(1-q) \rfloor \leq k < R_m^B} |T^B(\bar{1}^k \bar{0})|.$$

Lemma 4.8 tells us that $R_m^B - \lfloor n(1 - q) \rfloor = o_{\mathbb{P}}(\log n)$. Moreover, by Lemma 1.10, we know that the entries of the sequence $(|T^B(\bar{1}^k \bar{0})|)_{k \geq 0}$ are independent GEOMETRIC(1 - q) random variables, which gives us that

$$\sum_{\lfloor n(1-q) \rfloor \leq k < R_m^B} |T^B(\bar{1}^k \bar{0})| = o_{\mathbb{P}}\left(\frac{\log n}{1 - q}\right).$$

Finally, by Proposition 4.1, we know that $|T_n^B(\bar{1}^{R_m^B})| = o_{\mathbb{P}}(\frac{\log n}{1 - q})$. Combining all those results, we obtain that

$$(4.11) \quad |T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})| = o_{\mathbb{P}}\left(\frac{\log n}{1 - q}\right).$$

We focus now on bounding the size of the tree from below. Write $E^+ = \{R_m^B > \lfloor n(1 - q) \rfloor\}$ and $E^- = \{R_m^B \leq \lfloor n(1 - q) \rfloor\}$. Since on E^+ , we have

$$T_n^B(\bar{1}^{R_m^B}) \subseteq T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor}),$$

it follows that

$$|T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})| \mathbb{1}_{E^+} \geq |T_n^B(\bar{1}^{R_m^B})| \mathbb{1}_{E^+}.$$

Moreover, by Proposition 4.1, we know that

$$|T_n^B(\bar{1}^{R_m^B})| \geq (1 + o_{\mathbb{P}}(1)) \frac{e^{-(\log n)^{\frac{3}{4}}}}{1 - q},$$

which implies that

$$|T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})|_{\mathbb{1}_{E^+}} \geq (1 + o_{\mathbb{P}}(1)) \frac{e^{-(\log n)^{\frac{3}{4}}}}{1 - q} \mathbb{1}_{E^+}.$$

Using the fact that $-\log(1 - q) \sim \log n$, we obtain

$$(4.12) \quad \mathbb{1}_{E^+} \log |T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})| \geq \mathbb{1}_{E^+} (1 + o_{\mathbb{P}}(1)) \log n.$$

For the second part of the lower bound, letting $D = \lfloor n(1 - q) \rfloor - R_m^B$ and applying Lemma 4.9, we have that

$$|T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})|_{\mathbb{1}_{E^-}} \geq \left[\frac{\log(1 - P_D(1 - q^{|T_n^B(\bar{1}^{R_m^B})|}))}{\log q} - D \right] \mathbb{1}_{E^-}.$$

Using the lower bound of Proposition 4.1, we obtain

$$\begin{aligned} \frac{\log(1 - P_D(1 - q^{|T_n^B(\bar{1}^{R_m^B})|}))}{\log q} \mathbb{1}_{E^-} &\geq \frac{\log(1 - P_D(1 - q^{(1+o_{\mathbb{P}}(1)) \frac{e^{-(\log n)^{\frac{3}{4}}}}{1-q}}))}{\log q} \mathbb{1}_{E^-} \\ &= (1 + o_{\mathbb{P}}(1)) \frac{P_D e^{-(\log n)^{\frac{3}{4}}}}{1 - q} \mathbb{1}_{E^-}. \end{aligned}$$

Recalling that P_D is a product of D independent uniforms, it is immediate that $\log P_D = \Theta_{\mathbb{P}}(D)$. Moreover, by Lemma 4.8, we know that $D = \lfloor n(1 - q) \rfloor - R_m^B = o_{\mathbb{P}}(\log n)$. Combining these results, we obtain that

$$\left(\frac{P_D e^{-(\log n)^{\frac{3}{4}}}}{1 - q} - D \right) \mathbb{1}_{E^-} = \left(\frac{1}{1 - q} e^{o_{\mathbb{P}}(\log n)} + o_{\mathbb{P}}(\log n) \right) \mathbb{1}_{E^-} = n^{1-o_{\mathbb{P}}(1)} \mathbb{1}_{E^-},$$

hence

$$\mathbb{1}_{E^-} \log |T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})| \geq \mathbb{1}_{E^-} (1 + o_{\mathbb{P}}(1)) \log n.$$

Combined with (4.12), this implies that

$$(4.13) \quad \log |T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})| \geq (1 + o_{\mathbb{P}}(1)) \log n.$$

On the other hand, taking the logarithm in (4.11), we have that

$$\log |T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})| \leq (1 + o_{\mathbb{P}}(1)) \log n,$$

and combining this upper bound with (4.13), it follows that

$$\log |T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})| = (1 + o_{\mathbb{P}}(1)) \log n.$$

Plugging this back into (4.11) shows that

$$|T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})|(1 - q) = o_{\mathbb{P}}(\log |T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})|).$$

This bound implies that we can apply Proposition 2.1 to this subtree and obtain that

$$h(T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})) = (c^* + o_{\mathbb{P}}(1)) \log |T_n^B(\bar{1}^{\lfloor n(1-q) \rfloor})| = (c^* + o_{\mathbb{P}}(1)) \log n.$$

Since T_n^B and $T_{n,q}$ are identically distributed, this concludes the proof of the proposition. \square

We conclude this section with the proof of Theorem 1.1.

PROOF OF THEOREM 1.1. Let us first assume that $(q_n)_{n \geq 0}$ is such that $n(1 - q_n) = \Theta(\log n)$, and prove that

$$h(T_{n,q_n}) = n(1 - q_n) + c^* \log n + o_{\mathbb{P}}(\log n).$$

Using Proposition 1.13 and the fact that

$$h(T_n^B) \geq \lfloor n(1 - q_n) \rfloor + h(T_n^B(\bar{1}^{\lfloor n(1-q_n) \rfloor})),$$

it follows that

$$h(T_{n,q_n}) \stackrel{d}{=} h(T_n^B) \geq n(1 - q_n) + c^* \log n + o_{\mathbb{P}}(\log n).$$

For the upper bound, recall the stochastic inequality from (1.3) and apply it with $d = \lfloor n(1 - q_n) \rfloor$:

$$h(T_{n,q}) \leq \lfloor n(1 - q_n) \rfloor + \max \left\{ \sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\}, h(T_n^B(\bar{1}^{\lfloor n(1-q_n) \rfloor})) \right\}.$$

Now, Proposition 1.13 tells us that

$$h(T_n^B(\bar{1}^{\lfloor n(1-q_n) \rfloor})) = (c^* + o_{\mathbb{P}}(1)) \log n,$$

and from Proposition 1.12, we know that

$$\sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} \leq c^* \log \left(\frac{1}{1 - q_n} \right) + O_{\mathbb{P}} \left(\sqrt{\log \left(\frac{1}{1 - q_n} \right)} \right).$$

Since $\log \frac{1}{1 - q_n} \sim \log n$, it follows that

$$\sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} \leq (c^* + o_{\mathbb{P}}(1)) \log n,$$

and this proves that

$$h(T_{n,q_n}) \leq n(1 - q_n) + c^* \log n + o_{\mathbb{P}}(\log n).$$

This concludes the proof of the fact

$$h(T_{n,q_n}) = n(1 - q_n) + c^* \log n + o_{\mathbb{P}}(\log n)$$

whenever $n(1 - q_n) = \Theta(\log n)$.

Now, in order to prove Theorem 1.1, let $(q_n)_{n \geq 0}$ be any sequence taking values in $[0, 1]$. By Proposition 2.7, it suffices to prove that

$$\left(\frac{h(T_{n,q_n})}{n(1 - q_n) + c^* \log n} \right)_{n \geq 1}$$

converges in probability to 1. By considering subsequences if necessary, we can assume that $(q_n)_{n \geq 0}$ falls into one of the following regimes:

- $n(1 - q_n) = \Theta(\log n)$.

- $n(1 - q_n) = \omega(\log n)$.
- $n(1 - q_n) = o(\log n)$ and $q_n \neq 1$ for all $n \geq 0$.
- $q_n = 1$ for all $n \geq 0$.

The case $n(1 - q_n) = \Theta(\log n)$ was handled in the first part of the proof. The case $n(1 - q_n) = \omega(\log n)$ follows from Theorem 1.2. When $n(1 - q_n) = o(\log n)$, the results follow from Proposition 2.1, with the bounding sequence $(\gamma_n)_{n \geq 0}$ in that proposition chosen so that $\sqrt{\log n} \vee n(1 - q_n) \ll \gamma_n \ll \log n$. Finally, the case when $q_n = 1$ for all n is simply that of binary search trees, in which case the result was proved by Devroye [11]. This concludes the proof of Theorem 1.1. \square

5. Distributional limits. In this last section, we prove Theorem 1.3 and Theorem 1.4. The bulk of this section is devoted to proving the central limit theorem for the right depth R_n^B ; this was stated as Proposition 1.14 above. We then prove the central limit theorem for the height by combining this proposition with Proposition 1.15 and 1.12. We conclude this section with the proof of Theorem 1.4.

5.1. *Central limit theorem for the right depth.* Before proving the central limit theorem for $h(T_{n,q_n})$, we prove Proposition 1.14, which corresponds to a central limit theorem for the right depth R_n^B .

PROOF OF PROPOSITION 1.14. Let $(q_n)_{n \geq 0}$ be such that $nq_n = \omega(1)$. We use the Lindeberg–Feller theorem [15], Theorem 3.4.10, to prove this results.

First, using the characteristic of R_n^B from Proposition 1.11, it follows that R_n^B can be seen as an independent sum of Bernoulli random variables:

$$R_n^B \stackrel{d}{=} \sum_{1 < k \leq n} X_k^{(n)},$$

where $X_k^{(n)} \sim \text{BERNOULLI}(\frac{1 - q_n}{1 - q_n^k})$. Define now $s_n^2 = \sum_{1 < k \leq n} \text{Var}(X_k^{(n)})$ and note that

$$s_n^2 = \sum_{1 < k \leq n} \left[\frac{1 - q_n}{1 - q_n^k} - \left(\frac{1 - q_n}{1 - q_n^k} \right)^2 \right] = \mu_1(n, q_n) - \mu_2(n, q_n).$$

Using the results of Propositions 2.5 and 2.6, we have

$$\begin{aligned} s_n^2 &= n(1 - q_n) + \log\left(n \wedge \frac{1}{1 - q_n}\right) + O\left(\sqrt{\log\left(n \wedge \frac{1}{1 - q_n}\right)}\right) - n(1 - q_n)^2 + O(1) \\ &= (1 + o(1))\left(n(1 - q_n)q_n + \log\left(n \wedge \frac{1}{1 - q_n}\right)\right). \end{aligned}$$

Note that s_n is asymptotically equivalent to the denominator in the first convergence result asserted by the proposition. Note also that the above asymptotic implies that $s_n \rightarrow \infty$.

Now fix $\varepsilon > 0$. Since the $(X_k^{(n)})_{1 < k \leq n}$ are Bernoulli random variables, it is immediate that $|X_k^{(n)} - \mathbb{E}[X_k^{(n)}]| \leq 1$, which implies that, for n large enough,

$$\frac{1}{s_n^2} \sum_{1 < k \leq n} \mathbb{E}[(X_k^{(n)} - \mathbb{E}[X_k^{(n)}])^2 \mathbb{1}_{\{|X_k^{(n)} - \mathbb{E}[X_k^{(n)}]| > \varepsilon s_n\}}] = 0.$$

This proves that Linderberg condition holds for $(X_k^{(n)})_{1 < k \leq n}$, from which it follows that

$$\frac{1}{s_n} \sum_{1 < k \leq n} (X_k^{(n)} - \mathbb{E}[X_k^{(n)}]) \xrightarrow{d} \text{NORMAL}(0, 1).$$

The first statement of the proposition follows since $\sum_{1 < k \leq n} \mathbb{E}[X_k^{(n)}] = \sum_{1 < k \leq n} \frac{1 - q_n}{1 - q_n^k}$. Assuming furthermore that $n(1 - q_n) = \omega(\log n)$, the second statement becomes a consequence of Propositions 2.5 and 2.6 since we have that

$$\sum_{1 < k \leq n} \mathbb{E}[X_k^{(n)}] = \mu_1(n, q_n) = n(1 - q_n) + \log((1 - q_n)^{-1}) + o(\sqrt{n(1 - q_n)q_n})$$

and that $s_n^2 \sim n(1 - q_n)q_n$. \square

5.2. *Central limit theorem for the height of Mallows trees.* Before proving Theorem 1.3, we use results from previous sections to prove Proposition 1.15.

PROOF OF PROPOSITION 1.15. Let $(q_n)_{n \geq 0}$ and $m = m(n)$ be defined as in the statement of the proposition. We want to prove that the sequence of random variables

$$\left(\frac{h(T_n^B(\bar{1}^{R_m^B+1})) - c^* \log n}{\sqrt{\log n}} \right)_{n \geq 2}$$

is tight. In order to do so, we will prove that, for any sequence $(\gamma_n)_{n \geq 0}$ such that $\gamma_n = \omega(\sqrt{\log n})$, we have

$$(5.1) \quad \mathbb{P}(|h(T_n^B(\bar{1}^{R_m^B+1})) - c^* \log n| \geq \gamma_n) \longrightarrow 0.$$

Let $(\gamma_n)_{n \geq 0}$ be a sequence such that $\gamma_n = \omega(\sqrt{\log n})$, and assume without loss of generality that $\gamma_n = o(\log n)$.

Consider a sequence $(\beta_n)_{n \geq 0}$ such that $\beta_n = \omega(\sqrt{\log n})$, and that $\beta_n = o(\gamma_n)$; in other words, $\sqrt{\log n} \ll \beta_n \ll \gamma_n \ll \log n$. By applying Proposition 4.1, we know that

$$\mathbb{P}(e^{-\beta_n} \leq |T_n^B(\bar{1}^{R_m^B+1})|(1 - q_n) \leq \beta_n) \longrightarrow 1.$$

This implies that

$$\begin{aligned} &\mathbb{P}(|h(T_n^B(\bar{1}^{R_m^B+1})) - c^* \log n| \geq \gamma_n) \\ &= \mathbb{P}(|h(T_n^B(\bar{1}^{R_m^B+1})) - c^* \log n| \geq \gamma_n \mid e^{-\beta_n} \leq |T_n^B(\bar{1}^{R_m^B+1})|(1 - q_n) \leq \beta_n) + o(1). \end{aligned}$$

Recall that $\log(n(1 - q_n)) = O(\sqrt{\log n})$. Since $\sqrt{\log n} \ll \beta_n \ll \gamma_n \ll \log n$, this implies that, for any sequence $(s_n)_{n \geq 0}$ such that $e^{-\beta_n} \leq s_n(1 - q_n) \leq \beta_n$, we have $\log s_n = \log n + O(\beta_n) \sim \log n$. It follows that

$$s_n(1 - q_n) \leq \beta_n = o(\log s_n)$$

and that

$$\frac{\gamma_n}{s_n(1 - q_n) \vee \sqrt{\log n}} \geq \frac{\gamma_n}{\beta_n \vee \sqrt{\log n}} = \omega(1).$$

This corresponds to the assumptions of Proposition 2.1, and we henceforth know that

$$\mathbb{P}(|h(T_{s_n, q_n}) - c^* \log s_n| \geq \gamma_n) \longrightarrow 0.$$

Moreover, since $\log s_n = \log n + O(\beta_n) = \log n + o(\gamma_n)$, it follows that

$$\mathbb{P}(|h(T_{s_n, q_n}) - c^* \log n| \geq \gamma_n) \longrightarrow 0.$$

Since, conditioned on having size s_n , $T_n^B(\bar{1}^{R_m^B+1})$ is distributed as T_{s_n, q_n} , this implies that

$$\mathbb{P}(|h(T_n^B(\bar{1}^{R_m^B+1})) - c^* \log n| \geq \gamma_n \mid e^{-\beta_n} \leq |T_n^B(\bar{1}^{R_m^B+1})|(1 - q_n) \leq \beta_n) \longrightarrow 0,$$

which proves that (5.1) holds, and concludes the proof of the proposition. \square

With the previous results, we can now prove Theorem 1.3.

PROOF OF THEOREM 1.3. We will prove that, for all $t \in \mathbb{R}$, we have

$$(5.2) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{h(T_n^B) - n(1 - q_n) - c^* \log((1 - q_n)^{-1})}{\sqrt{n(1 - q_n)q_n}} \leq t \right) = \Phi(t),$$

where Φ is the cumulative density function of the NORMAL(0, 1) distribution. By considering subsequences if necessary, we can assume either that $n(1 - q_n) = \omega(\log^2 n)$, or that $n(1 - q_n) = O(\log^2 n)$. We now fix $t \in \mathbb{R}$ and prove that (5.2) holds by dividing the proof into the two previous cases.

Assume first that $n(1 - q_n) = \omega(\log^2 n)$. Since $nq_n = \omega(1)$, this implies that

$$\frac{\sqrt{n(1 - q_n)q_n}}{\log((1 - q_n)^{-1})} \rightarrow \infty.$$

Let $(\gamma_n)_{n \geq 0}$ be any sequence that converges to infinity such that $\gamma_n = \omega(\log((1 - q_n)^{-1}))$ and $\gamma_n = o(\sqrt{n(1 - q_n)q_n})$, and define

$$E_n = \{h(T_n^B) - R_n^B \geq \gamma_n\}.$$

Recall the upper bound from (1.2):

$$h(T_n^B) - R_n^B \leq 1 + \max_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\}.$$

Using this bound, we have

$$\mathbb{P}(E_n) \leq \mathbb{P} \left(1 + \max_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} \geq \gamma_n \right),$$

and the right-hand side converges to 0 when $n \rightarrow \infty$, thanks to Proposition 1.12 applied with $\xi_n = \gamma_n - 1 - c^* \log((1 - q_n)^{-1}) - M\sqrt{\log(1 - q_n)^{-1}}$; this tends to infinite since $\gamma_n = \omega(\log((1 - q_n)^{-1}))$. It follows that

$$(5.3) \quad \mathbb{P} \left(\frac{h(T_n^B) - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t \right) = \mathbb{P} \left(\frac{h(T_n^B) - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t, E_n^c \right) + o(1).$$

To bound the right-hand side, on one hand, using that $R_n^B \leq h(T_n^B)$, we have

$$\mathbb{P} \left(\frac{h(T_n^B) - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t, E_n^c \right) \leq \mathbb{P} \left(\frac{R_n^B - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t \right),$$

and by applying Proposition 1.14, the upper bound converges to $\Phi(t)$; we are using here that $\sqrt{n(1 - q_n)q_n} = \omega(\log((1 - q_n)^{-1}))$, so the $\log((1 - q_n)^{-1})$ term in the numerator of Proposition 1.14 is asymptotically negligible and can be dropped. On the other hand, by the definition of E_n , we have $h(T_n^B) < R_n^B + \gamma_n$ on E_n^c , and it follows that

$$\begin{aligned} \mathbb{P} \left(\frac{h(T_n^B) - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t, E_n^c \right) &\geq \mathbb{P} \left(\frac{R_n^B + \gamma_n - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t, E_n^c \right) \\ &\geq \mathbb{P} \left(\frac{R_n^B - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t - \frac{\gamma_n}{\sqrt{n(1 - q_n)q_n}} \right) - \mathbb{P}(E_n). \end{aligned}$$

Since $\frac{\gamma_n}{\sqrt{n(1 - q_n)q_n}}$ and $\mathbb{P}(E_n)$ converge to 0, this lower bound also converges to $\Phi(t)$, again thanks to Proposition 1.14. Combing the last two results with (5.3) and again using that

$\log((1 - q_n)^{-1}) = o(\sqrt{n(1 - q_n)q_n})$, it follows that

$$\begin{aligned} &\mathbb{P}\left(\frac{h(T_n^B) - n(1 - q_n) - c^* \log((1 - q_n)^{-1})}{\sqrt{n(1 - q_n)q_n}} \leq t, E_n^c\right) \\ &= \mathbb{P}\left(\frac{h(T_n^B) - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t, E_n^c\right) + o(1) \\ &= \Phi(t) + o(1), \end{aligned}$$

which concludes the proof of Theorem 1.3 in the case where $n(1 - q_n) = \omega(\log^2 n)$.

Assume now that $n(1 - q_n) = O(\log^2 n)$ and that $n(1 - q_n) = \omega(\log n)$. Note that this implies that $\log(n(1 - q_n)) = O(\log \log n) = O(\sqrt{\log n})$ and then $\log n = \log((1 - q_n)^{-1}) + O(\sqrt{\log n})$. Let $m = m(n) = \min\{\ell \geq 0 : \ell(1 - q_n) + \log \ell \geq n(1 - q_n)\}$ and $(\gamma_n)_{n \geq 0}$ be a sequence such that $\gamma_n = \omega(\sqrt{\log n})$ and $\gamma_n = o(\sqrt{n(1 - q_n)q_n})$, which is possible since $n(1 - q_n)q_n = \omega(\log n)$. Define the event

$$F_n = \{|h(T_n^B) - R_m^B - 1 - c^* \log((1 - q_n)^{-1})| \geq \gamma_n\}.$$

Using both bounds of (1.4), we have that

$$\begin{aligned} &\mathbb{P}(F_n) \\ &= \mathbb{P}(h(T_n^B) - R_m^B - 1 \geq c^* \log((1 - q_n)^{-1}) + \gamma_n) \\ &\quad + \mathbb{P}(h(T_n^B) - R_m^B - 1 \leq c^* \log((1 - q_n)^{-1}) - \gamma_n) \\ &\leq \mathbb{P}\left(\max_{k \geq 0} \left\{ \sup\{h(T^B(\bar{1}^k \bar{0})) - k\}, h(T_n^B(\bar{1}^{R_m^B+1})) \right\} \geq c^* \log((1 - q_n)^{-1}) + \gamma_n\right) \\ &\quad + \mathbb{P}(h(T_n^B(\bar{1}^{R_m^B+1})) \leq c^* \log((1 - q_n)^{-1}) - \gamma_n). \end{aligned}$$

By applying Proposition 1.15, which states that $h(T_n^B(\bar{1}^{R_m^B+1})) = c^* \log n + O_{\mathbb{P}}(\sqrt{\log n})$, and Proposition 1.12, which states that $\sup_{k \geq 0} \{h(T^B(\bar{1}^k \bar{0})) - k\} \leq c^* \log n + O_{\mathbb{P}}(\sqrt{\log n})$, we obtain that $\mathbb{P}(F_n) \rightarrow 0$.

Separating (5.2) according to F_n and F_n^c as previously, we now obtain that

$$\begin{aligned} &\mathbb{P}\left(\frac{h(T_n^B) - n(1 - q_n) - c^* \log((1 - q_n)^{-1})}{\sqrt{n(1 - q_n)q_n}} \leq t\right) \\ &= \mathbb{P}\left(\frac{h(T_n^B) - n(1 - q_n) - c^* \log((1 - q_n)^{-1})}{\sqrt{n(1 - q_n)q_n}} \leq t, F_n^c\right) + o(1). \end{aligned}$$

Note that, by definition of m , we have $m(1 - q_n) + \log m = n(1 - q_n) + O(1)$. Since $n(1 - q_n) = \omega(\log n)$, it follows that $m \sim n$ and so $\log m = \log n + O(1)$. This implies that $m(1 - q_n) = \omega(\log m)$ and that $m q_n = \omega(1)$, and by Proposition 1.14, we obtain

$$\mathbb{P}\left(\frac{R_m^B - m(1 - q_n) - \log((1 - q_n)^{-1})}{\sqrt{m(1 - q_n)q_n}} \leq t\right) \rightarrow \Phi(t).$$

The previous identities also imply that $m(1 - q_n)q_n \sim n(1 - q_n)q_n$ and, since $\log(n(1 - q_n)) = O(\log \log n)$, that

$$m(1 - q_n) + \log((1 - q_n)^{-1}) = n(1 - q_n) + O(\log \log n) = n(1 - q_n) + o(\sqrt{n(1 - q_n)q_n}).$$

It follows that

$$\mathbb{P}\left(\frac{R_m^B - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t\right) \rightarrow \Phi(t).$$

Now, use the definition of F_n and the previous asymptotic result to obtain that, on one hand

$$\begin{aligned} & \mathbb{P}\left(\frac{h(T_n^B) - n(1 - q_n) - c^* \log((1 - q_n)^{-1})}{\sqrt{n(1 - q_n)q_n}} \leq t, F_n^c\right) \\ & \leq \mathbb{P}\left(\frac{R_n^B + 1 - n(1 - q_n) - \gamma_n}{\sqrt{n(1 - q_n)q_n}} \leq t, F_n^c\right) \\ & \leq \mathbb{P}\left(\frac{R_n^B - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t + \frac{\gamma_n - 1}{\sqrt{n(1 - q_n)q_n}}\right) \\ & = \Phi(t) + o(1), \end{aligned}$$

and on the other hand

$$\begin{aligned} & \mathbb{P}\left(\frac{h(T_n^B) - n(1 - q_n) - c^* \log((1 - q_n)^{-1})}{\sqrt{n(1 - q_n)q_n}} \leq t, F_n^c\right) \\ & \geq \mathbb{P}\left(\frac{R_n^B + 1 - n(1 - q_n) + \gamma_n}{\sqrt{n(1 - q_n)q_n}} \leq t, F_n^c\right) \\ & \geq \mathbb{P}\left(\frac{R_n^B - n(1 - q_n)}{\sqrt{n(1 - q_n)q_n}} \leq t - \frac{\gamma_n + 1}{\sqrt{n(1 - q_n)q_n}}\right) - \mathbb{P}(F_n) \\ & = \Phi(t) + o(1). \end{aligned}$$

This proves that

$$\mathbb{P}\left(\frac{h(T_n^B) - n(1 - q_n) - c^* \log((1 - q_n)^{-1})}{\sqrt{n(1 - q_n)q_n}} \leq t, F_n^c\right) = \Phi(t) + o(1),$$

which concludes the proof of Theorem 1.3 in the second and last case. \square

5.3. *Poisson fluctuations for the height.* We conclude this section with the proof of Theorem 1.4.

PROOF OF THEOREM 1.4. We will prove that, when $nq_n \rightarrow \lambda \in [0, \infty)$, then

$$n - 1 - h(T_n^B) \xrightarrow{d} \text{POISSON}(\lambda).$$

Start by considering the event

$$E_n = \{h(T_n^B) > R_n^B\}.$$

Since $R_n^B \leq h(T_n^B)$, this means that $E_n^c = \{R_n^B = h(T_n^B)\}$. Now, for the height of the whole tree to be larger than the right depth, there must be a nonempty left subtree $T_n^B(\bar{1}^k \bar{0})$ for some $0 \leq k \leq R_n^B$. Moreover, if this nonempty left subtree is not $T_n^B(\bar{1}^{R_n^B} \bar{0})$ or $T_n^B(\bar{1}^{R_n^B - 1} \bar{0})$, then its size has to be larger than 2. This implies that

$$E_n \subseteq \{ |T_n^B(\bar{1}^{R_n^B} \bar{0})| \geq 1 \} \cup \{ |T_n^B(\bar{1}^{R_n^B - 1} \bar{0})| \geq 1 \} \cup \bigcup_{0 \leq k \leq R_n^B - 2} \{ |T_n^B(\bar{1}^k \bar{0})| \geq 2 \}.$$

Recall from Lemma 1.10 that the sizes of the trees $(|T^B(\bar{1}^k \bar{0})|)_{k \geq 0}$ are all independent, GEOMETRIC($1 - q_n$) distributed random variable. Using that $R_n^B \leq n - 1$ and that $T_n^B(\bar{1}^k \bar{0}) \subseteq T^B(\bar{1}^k \bar{0})$, this implies that

$$\mathbb{P}(E_n) \leq 2\mathbb{P}(|T^B(\bar{0})| \geq 1) + (n - 2)\mathbb{P}(|T^B(\bar{0})| \geq 2) = 2q_n + (n - 2)q_n^2 = o(1).$$

This proves that $\mathbb{P}(h(T_n^B) = R_n^B) = 1 - o(1)$, when $nq_n = O(1)$. We will now prove that

$$n - 1 - R_n^B \xrightarrow{d} \text{POISSON}(\lambda),$$

by showing that the characteristic function of $n - 1 - R_n^B$ converges to that of a $\text{POISSON}(\lambda)$ random variable.

Since $q_n = o(1)$, we have that $\frac{1-q_n}{1-q_n^k} = 1 - q_n + o(q_n)$ for all $k \geq 2$, where the small- o term can be chosen to be independent of k . Consider now the characteristic function for R_n^B from Proposition 1.11 to obtain

$$\begin{aligned} \mathbb{E}[e^{it(n-1-R_n^B)}] &= e^{it(n-1)} \prod_{1 < k \leq n} \left(1 + (e^{-it} - 1) \frac{1 - q_n}{1 - q_n^k}\right) \\ &= e^{it(n-1)} (1 + (e^{-it} - 1)(1 - q_n + o(q_n)))^{n-1} \\ &= (1 + (e^{it} - 1 + o(1))q_n)^{n-1}. \end{aligned}$$

Since $nq_n \rightarrow \lambda$, it follows that

$$\mathbb{E}[e^{it(n-1-R_n^B)}] \longrightarrow e^{\lambda(e^{it}-1)};$$

this concludes the proof of the theorem. \square

6. Further questions. This paper studies Mallows trees and proved some of the properties of its height. However, several related questions remain open. We discuss some of the possible further studies below:

- Between Theorem 1.3 and Theorem 1.4, we have a good understand of the distributional limit of $h(T_{n,q_n})$ when $n(1 - q_n)/\log n \rightarrow \infty$. Moreover, we know from the results of [13, 30] that when $q_n \equiv 1$, the central limit theorem does not hold anymore, and the variance of the height is $\Theta(1)$. This means that there exists a transition between the regime $n(1 - q_n)/\log n \rightarrow \infty$ and $q_n = 1$ where the central limit theorem of Theorem 1.3 stops holding and moves to a more concentrated process with finite variance. It is natural to ask where this transition occurs; it is not clear to us whether the condition $n(1 - q_n)/\log n \rightarrow \infty$ is necessary in order for a Gaussian central limit theorem to hold, or what other distributional limits are possible for sequences $(q_n)_{n \geq 0}$ with $\limsup_{n \rightarrow \infty} n(1 - q_n)/\log n < \infty$.
- In this paper, we studied the height of T_{n,q_n} in large part by relating it to the length of the rightmost path in T_{n,q_n} ; we did so by bounding from above the height of the left subtrees $T_{n,q_n}(\bar{1}^k \bar{0})$, for $k \geq 0$. The intrinsic properties of the left subtrees both for finite n and in the $n \rightarrow \infty$ limit, deserve further exploration in our view, and we next list a couple of specific questions of interest.

We know that for any fixed $q \in [0, 1]$ and $k \in \mathbb{N}$, the trees $(T_{n,q}(\bar{1}^k \bar{0}))_{n \geq 0}$ are stochastically increasing in n . Moreover, working in the infinite b -model, we have $T_n^{B(q)}(\bar{1}^k \bar{0}) = T^{B(q)}(\bar{1}^k \bar{0})$ for all n sufficiently large (recall that, by definition, $T^{B(q)} = \lim_{n \rightarrow \infty} T_n^{B(q)}$). It would be interesting to understand this filling process, that is, to study the behaviour of

$$D_n^{B(q)} = \max\{0 \leq k \leq R_n^{B(q)} : T_n^{B(q)}(\bar{1}^k \bar{0}) = T^{B(q)}(\bar{1}^k \bar{0})\},$$

which corresponds to the depth until which all trees are filled, as both n and q vary, and of

$$S_{n,k}^{B(q)} = \frac{|T_n^{B(q)}(\bar{1}^k \bar{0})|}{|T^{B(q)}(\bar{1}^k \bar{0})|},$$

which corresponds to the proportion of the subtree $T^{B(q)}(\bar{1}^k\bar{0})$ already present at time n ; here both q and k may depend on n .

- Another direction of studies regarding the left subtrees is to consider the structure of the tree $T^{B(q)}(\bar{0})$, especially as $q \rightarrow 1$. With the results from this paper, it is fairly straightforward to verify that, in the case when $q \rightarrow 1$, the height of $T^{B(q)}(\bar{0})$ is $(c^* + o_{\mathbb{P}}(1)) \log(1/(1 - q))$. It would be interesting to understand the lower order corrections to this height. We expect the height of $T^{B(q)}(\bar{0})$ to have bounded variance as $q \rightarrow 1$, and to converge in distribution after re-centering, at least along subsequences. It could also be interesting to characterize the filling levels (as in [11]) or the total path length of this tree.
- Corollary 1.9 says that $T_{n,q}$ is stochastically increasing in n when q is fixed. Computations for small values of n suggest that $T_{n,q}$ is also stochastically decreasing in q . This would be interesting if true and would also provide a useful comparison tool, which would simplify some of the arguments of the current work (in particular Proposition 1.7, in which we could simply chose $m = n$).

APPENDIX A: MOMENTS OF THE RIGHT DEPTH

In this Appendix, we prove Proposition 2.3 and deduce Fact 2.2 from it. Recall the definition of μ_α from Section 2.1:

$$\mu_\alpha(n, q) = \sum_{1 < k \leq n} \left(\frac{1 - q}{1 - q^k} \right)^\alpha.$$

PROOF OF PROPOSITION 2.3. For $\beta \geq 0$, let e_β be the elementary symmetric polynomial,

$$e_\beta(x_1, \dots, x_n) = \sum_{\{1 \leq k_1 \neq \dots \neq k_\beta \leq n\}} \prod_{\{1 \leq i \leq \beta\}} x_{k_i},$$

with the convention that $e_0(x_1, \dots, x_n) = 1$. By using that

$$\prod_{1 \leq k \leq n} (1 + x_k) = \sum_{0 \leq \beta \leq n} e_\beta(x_1, \dots, x_n)$$

along with the formula for the moment generating function of R_n^B from Proposition 1.11, it follows that

$$\begin{aligned} \mathbb{E}[e^{tR_n^B}] &= \prod_{1 < k \leq n} \left(1 + (e^t - 1) \frac{1 - q}{1 - q^k} \right) \\ &= \sum_{0 \leq \beta \leq n-1} (e^t - 1)^\beta e_\beta \left(\frac{1 - q}{1 - q^2}, \dots, \frac{1 - q}{1 - q^n} \right). \end{aligned}$$

Pitman [28] shows that

$$(e^t - 1)^\beta = \beta! \sum_{\alpha \geq \beta} \begin{Bmatrix} \alpha \\ \beta \end{Bmatrix} \frac{t^\alpha}{\alpha!},$$

which implies that

$$\mathbb{E}[e^{tR_n^B}] = \sum_{\alpha \geq 0} \left(\sum_{0 \leq \beta \leq \alpha \wedge (n-1)} \beta! \begin{Bmatrix} \alpha \\ \beta \end{Bmatrix} e_\beta \left(\frac{1 - q}{1 - q^2}, \dots, \frac{1 - q}{1 - q^n} \right) \right) \frac{t^\alpha}{\alpha!}.$$

It follows that the moments of R_n^B are given by

$$\mathbb{E}[(R_n^B)^\alpha] = \sum_{0 \leq \beta \leq \alpha \wedge (n-1)} \beta! \begin{Bmatrix} \alpha \\ \beta \end{Bmatrix} e_\beta \left(\frac{1-q}{1-q^2}, \dots, \frac{1-q}{1-q^n} \right).$$

Now apply Newton’s identity [24] recursively to obtain that

$$e_\beta(x_1, \dots, x_n) = (-1)^\beta \sum_{s \in S_\beta} \prod_{1 \leq i \leq \beta} \frac{(-1)^{s_i}}{s_i! i^{s_i}} \left(\sum_{1 \leq k \leq n} x_k^i \right)^{s_i}.$$

Combining the last two identities, the proposition then follows from the definition of μ_α . \square

From this proposition, or by direct computation using the moment generating function of R_n^B from Proposition 1.11, we obtain that

$$\mathbb{E}[R_n^B] = \mu_1(n, q)$$

and

$$\mathbb{E}[(R_n^B)^2] = \mu_1(n, q) + \mu_1(n, q)^2 - \mu_2(n, q),$$

which proves Fact 2.2.

APPENDIX B: ASYMPTOTICS OF μ_α

In this Appendix, we prove Proposition 2.5 and 2.6. Both these proofs will be based on the following bounds for μ_α .

PROPOSITION B.1. *Let $\alpha \geq 1, n \geq 1$, and $q \in [0, 1)$. Then, for all $m \in [n]$, we have*

$$\mu_\alpha(n, q) \geq (n - m)(1 - q)^\alpha + \sum_{1 < k \leq m} \frac{1}{k^\alpha}$$

and

$$\mu_\alpha(n, q) \leq n(1 - q)^\alpha + \alpha m(1 - q)^\alpha q^m \frac{1 - q^n}{(1 - q^m)^{\alpha+1}} + \left(\frac{m(1 - q)}{1 - q^m} \right)^\alpha \sum_{1 < k \leq m} \frac{1}{k^\alpha}.$$

PROOF. First, write

$$(B.1) \quad \mu_\alpha(n, q) = \sum_{1 < k \leq m} \left(\frac{1 - q}{1 - q^k} \right)^\alpha + \sum_{m < k \leq n} \left(\frac{1 - q}{1 - q^k} \right)^\alpha.$$

For the lower bound use in the first sum that $1 - q^k \leq k(1 - q)$ and in the second sum that $1 - q^k \leq 1$ to obtain

$$\mu_\alpha(n, q) \geq \sum_{1 < k \leq m} \frac{1}{k^\alpha} + \sum_{m < k \leq n} (1 - q)^\alpha = \sum_{1 < k \leq m} \frac{1}{k^\alpha} + (n - m)(1 - q)^\alpha,$$

which is the desired bound.

For the upper bound, we start with the first term in (B.1). Define the function $\phi(x) = \frac{x}{1 - e^{-x}}$, so that

$$\sum_{1 < k \leq m} \left(\frac{1 - q}{1 - q^k} \right)^\alpha = \sum_{1 < k \leq m} \left(\frac{1 - q}{1 - e^{-k|\log q|}} \right)^\alpha = \sum_{1 < k \leq m} \left(\frac{1 - q}{|k|\log q|} \phi(k|\log q|) \right)^\alpha.$$

Note that ϕ is increasing, from which we deduce the following bound:

$$\begin{aligned} \sum_{1 < k \leq m} \left(\frac{1-q}{1-q^k} \right)^\alpha &\leq \sum_{1 < k \leq m} \left(\frac{1-q}{k|\log q|} \phi(m|\log q|) \right)^\alpha \\ &= \left(\frac{1-q}{|\log q|} \phi(m|\log q|) \right)^\alpha \sum_{1 < k \leq m} \frac{1}{k^\alpha} \\ &= \left(\frac{m(1-q)}{1-q^m} \right)^\alpha \sum_{1 < k \leq m} \frac{1}{k^\alpha}. \end{aligned}$$

which is the last term in the desired upper bound.

Consider now the second term of (B.1). Since $k \mapsto \frac{1-q}{1-q^k}$ is decreasing in k , we have $\frac{1-q}{1-q^k} \leq \frac{1-q}{1-q^{\ell m}}$ for $\ell m \leq k \leq (\ell + 1)m$, so

$$\sum_{m < k \leq n} \left(\frac{1-q}{1-q^k} \right)^\alpha \leq m \sum_{1 \leq \ell \leq \frac{n}{m}} \left(\frac{1-q}{1-q^{\ell m}} \right)^\alpha$$

Now, rewrite and bound the last sum as follows:

$$\begin{aligned} \sum_{1 \leq \ell \leq \frac{n}{m}} \left(\frac{1-q}{1-q^{\ell m}} \right)^\alpha &= (1-q)^\alpha \left\lfloor \frac{n}{m} \right\rfloor + \sum_{1 \leq \ell \leq \frac{n}{m}} \left[\left(\frac{1-q}{1-q^{\ell m}} \right)^\alpha - (1-q)^\alpha \right] \\ &\leq \frac{n(1-q)^\alpha}{m} + (1-q)^\alpha \sum_{1 \leq \ell \leq \frac{n}{m}} \frac{1 - (1-q^{\ell m})^\alpha}{(1-q^{\ell m})^\alpha}, \end{aligned}$$

to obtain that

$$\begin{aligned} \sum_{m < k \leq n} \left(\frac{1-q}{1-q^k} \right)^\alpha &\leq m \left[\frac{n(1-q)^\alpha}{m} + (1-q)^\alpha \sum_{1 \leq \ell \leq \frac{n}{m}} \frac{1 - (1-q^{\ell m})^\alpha}{(1-q^{\ell m})^\alpha} \right] \\ &= n(1-q)^\alpha + m(1-q)^\alpha \sum_{1 \leq \ell \leq \frac{n}{m}} \frac{1 - (1-q^{\ell m})^\alpha}{(1-q^{\ell m})^\alpha}. \end{aligned}$$

Simplifying the sum on the right using that $\frac{1}{1-q^{\ell m}} \leq \frac{1}{1-q^m}$ when $\ell \geq 1$, and that $1 - (1 - q^{\ell m})^\alpha \leq \alpha(1 - (1 - q^{\ell m})) = \alpha q^{\ell m}$, we have

$$\begin{aligned} \sum_{m < k \leq n} \left(\frac{1-q}{1-q^k} \right)^\alpha &\leq n(1-q)^\alpha + m(1-q)^\alpha \sum_{1 \leq \ell \leq \frac{n}{m}} \frac{\alpha q^{\ell m}}{(1-q^m)^\alpha} \\ &\leq n(1-q)^\alpha + \alpha m(1-q)^\alpha \frac{q^m}{(1-q^m)^\alpha} \frac{1-q^n}{1-q^m}. \end{aligned}$$

This corresponds to the first two terms in the desired upper bound and the proposition follows. □

In order to apply the two bounds from this proposition, we now choose the right sequence $(m_n)_{n \geq 0}$ corresponding to $(q_n)_{n \geq 0}$, so that the first terms in the asymptotic behaviour of $\mu_\alpha(n, q_n)$ correspond to $n(1 - q_n)^\alpha$ and $\sum_{1 < k \leq m_n} \frac{1}{k^\alpha}$.

PROOF OF PROPOSITION 2.5. We want to prove that

$$\mu_1(n, q_n) = n(1 - q_n) + \log \left(n \wedge \frac{1}{1 - q_n} \right) + O \left(\sqrt{\log \left(n \wedge \frac{1}{1 - q_n} \right)} \right).$$

First, in the case where $(q_n)_{n \geq 0}$ converges to 0, since $\frac{1}{1-q_n^k} = 1 + O(q_n^k)$ uniformly over k , it follows from the definition of $\mu_1(n, q_n)$ that

$$\mu_1(n, q_n) = \sum_{1 < k \leq n} (1 - q_n)(1 + O(q_n^k)) = n(1 - q_n) + O(q_n^2).$$

Since $\log(n \wedge \frac{1}{1-q_n}) = -\log(1 - q_n) \sim q_n$, the desired asymptotic behaviour follows.

We assume now that $(q_n)_{n \geq 0}$ is bounded away from 0. In this case, we prove the lower and the upper bounds separately.

For the lower bound, define $m_n = \lfloor n \wedge \frac{1}{1-q_n} \rfloor$. Using the lower bound in Proposition B.1 and since $m_n(1 - q_n) \leq 1$, we have

$$\mu_1(n, q_n) \geq (n - m_n)(1 - q_n) + \sum_{1 < k \leq m_n} \frac{1}{k} = n(1 - q_n) + \log m_n + O(1).$$

This result is actually stronger than what we aim to prove, nevertheless the desired lower bound follows.

For the upper bound, define $m_n = \lfloor \frac{n \wedge \frac{1}{1-q_n}}{\sqrt{\log(n \wedge \frac{1}{1-q_n})}} \rfloor$ and note that, since q_n is bounded away from 0, we have $m_n(1 - q_n) = O(1)$. Using the upper bound of Proposition B.1, we have

$$\mu_1(n, q_n) \leq n(1 - q_n) + m_n \frac{(1 - q_n)q_n^{m_n}}{1 - q_n^{m_n}} \frac{1 - q_n^n}{1 - q_n^{m_n}} + \frac{(1 - q_n)m_n}{1 - q_n^{m_n}} \sum_{1 < k \leq m_n} \frac{1}{k}.$$

By studying the variations of the function $\phi : x \mapsto \frac{(1-x)x^m}{1-x^m}$, one can see that it is increasing and smaller than $\frac{1}{m}$ for $x \in [0, 1)$, and it follows that

$$m_n \frac{(1 - q_n)q_n^{m_n}}{1 - q_n^{m_n}} \leq 1.$$

Moreover, since $m_n(1 - q_n) = O(1)$, we have

$$q_n^{m_n} = e^{m_n \log q_n} = e^{-m_n(1-q_n) + O(m_n(1-q_n)^2)} = 1 - m_n(1 - q_n) + O((m_n(1 - q_n))^2),$$

and so

$$\frac{(1 - q_n)m_n}{1 - q_n^{m_n}} = \frac{m_n(1 - q_n)}{m_n(1 - q_n) + O((m_n(1 - q_n))^2)} = 1 + O(m_n(1 - q_n)).$$

Combining the last results with the fact that $\sum_{1 < k \leq m_n} \frac{1}{k} = \log m_n + O(1)$ and that $1 - q_n^n \leq 1 \wedge (n(1 - q_n))$, we obtain

$$\begin{aligned} \mu_1(n, q_n) &\leq n(1 - q_n) + O\left(\frac{1 \wedge (n(1 - q_n))}{m_n(1 - q_n)}\right) + (1 + O(m_n(1 - q_n)))(\log m_n + O(1)) \\ &= n(1 - q_n) + \log m_n + O\left(\frac{1 \wedge (n(1 - q_n))}{m_n(1 - q_n)}\right) + O(m_n(1 - q_n) \log m_n) + O(1). \end{aligned}$$

The desired upper bound follows from this formula since $m_n = \lfloor \frac{n \wedge \frac{1}{1-q_n}}{\sqrt{\log(n \wedge \frac{1}{1-q_n})}} \rfloor$, which implies that

$$\frac{1 \wedge (n(1 - q_n))}{m_n(1 - q_n)} = \Theta\left(\sqrt{\log\left(n \wedge \frac{1}{1 - q_n}\right)}\right),$$

that

$$\log m_n = \log\left(n \wedge \frac{1}{1 - q_n}\right) + O\left(\log \log\left(n \wedge \frac{1}{1 - q_n}\right)\right),$$

and that

$$m_n(1 - q_n) \log m_n = O\left(\sqrt{\log\left(n \wedge \frac{1}{1 - q_n}\right)}\right). \quad \square$$

PROOF OF PROPOSITION 2.6. We use a similar technique, by proving an upper and a lower bound separately, to show that

$$\mu_\alpha(n, q_n) = n(1 - q_n)^\alpha + \zeta(\alpha) - 1 + O\left(\left((1 - q_n) \vee \frac{1}{n}\right)^{\frac{\alpha-1}{\alpha+1}}\right)$$

holds for all $\alpha > 1$.

Let $m_n = \lfloor n \wedge \frac{1}{1 - q_n} \rfloor$. Using the lower bound of Proposition B.1, we have

$$\mu_\alpha(n, q_n) \geq (n - m_n)(1 - q_n)^\alpha + \sum_{1 < k \leq m_n} \frac{1}{k^\alpha}.$$

Since $\sum_{1 < k \leq m_n} \frac{1}{k^\alpha} = \zeta(\alpha) - 1 + O\left(\frac{1}{(m_n)^{\alpha-1}}\right)$, we obtain

$$\mu_\alpha(n, q_n) \geq n(1 - q_n)^\alpha + \zeta(\alpha) - 1 + O(m_n(1 - q_n)^\alpha) + O\left(\frac{1}{(m_n)^{\alpha-1}}\right)$$

and the lower bound now follows from the fact that $m_n(1 - q_n)^\alpha \leq \frac{1}{(m_n)^{\alpha-1}} \leq \left(\frac{1}{m_n}\right)^{\frac{\alpha-1}{\alpha+1}}$.

For the upper bound, let

$$m_n = \left\lfloor \left(\frac{1}{(1 - q_n)^2} \wedge \frac{n}{1 - q_n} \wedge n^{\alpha+1} \right)^{\frac{1}{\alpha+1}} \right\rfloor.$$

This definition just encodes that there are three cases to study: $\frac{1}{1 - q_n} \leq n$, $n < \frac{1}{1 - q_n} \leq n^\alpha$, and $n^\alpha < \frac{1}{1 - q_n}$. In all cases, we have $1 \leq m_n \leq n$. Applying the upper bound of Proposition B.1, it follows that

$$(B.2) \quad \mu_\alpha(n, q_n) \leq n(1 - q_n)^\alpha + \alpha m_n \frac{(1 - q_n)^\alpha q_n^{m_n}}{(1 - q_n^{m_n})^\alpha} \frac{1 - q_n^n}{1 - q_n^{m_n}} + \left(\frac{m_n(1 - q_n)}{1 - q_n^{m_n}}\right)^\alpha \sum_{1 < k \leq m_n} \frac{1}{k^\alpha}.$$

On one hand, by the definition of m_n , we have $m_n(1 - q_n) \leq \frac{1 - q_n}{(1 - q_n)^{\frac{\alpha+1}{2}}} = (1 - q_n)^{\frac{\alpha-1}{\alpha+1}} \leq 1$.

It follows that $1 - q_n^{m_n} = \Omega(m_n(1 - q_n))$, and since $1 - q_n^n \leq 1 \wedge n(1 - q_n)$, we have

$$\begin{aligned} \alpha m_n \frac{(1 - q_n)^\alpha q_n^{m_n}}{(1 - q_n^{m_n})^\alpha} \frac{1 - q_n^n}{1 - q_n^{m_n}} &= \alpha m_n (1 - q_n)^\alpha q_n^{m_n} (1 - q_n^n) \cdot O\left(\frac{1}{(m_n(1 - q_n))^{\alpha+1}}\right) \\ &= O\left(\frac{1 \wedge (n(1 - q_n))}{m_n^\alpha (1 - q_n)}\right). \end{aligned}$$

On the other hand, we claim that

$$\left(\frac{m_n(1 - q_n)}{1 - q_n^{m_n}}\right)^\alpha = 1 + O(m_n(1 - q_n)).$$

This holds since if $q_n < 1 - \sqrt{2}$ then $m_n = 1$, and the claim simply asserts that $1 = 1 + O(1 - q_n)$; and if q_n is bounded away from zero then, using the fact that $m_n(1 - q_n) \leq 1$ again, we have

$$\left(\frac{m_n(1 - q_n)}{1 - q_n^{m_n}}\right)^\alpha = \left(\frac{m_n(1 - q_n)}{m_n(1 - q_n) + O((m_n(1 - q_n))^2)}\right)^\alpha = 1 + O(m_n(1 - q_n)).$$

We also know that $\sum_{1 < k \leq m_n} \frac{1}{k^\alpha} = \zeta(\alpha) - 1 + O(\frac{1}{m_n^{\alpha-1}})$, which yields the following bound:

$$\begin{aligned} \left(\frac{m_n(1 - q_n)}{1 - q_n^{m_n}}\right)^\alpha \sum_{1 < k \leq m_n} \frac{1}{k^\alpha} &= (1 + O(m_n(1 - q_n))) \left(\zeta(\alpha) - 1 + O\left(\frac{1}{m_n^{\alpha-1}}\right)\right) \\ &= \zeta(\alpha) - 1 + O\left(\frac{1}{m_n^{\alpha-1}}\right) + O(m_n(1 - q_n)). \end{aligned}$$

Plugging these results back into (B.2), we obtain

$$\begin{aligned} \mu_\alpha(n, q_n) &\leq n(1 - q_n)^\alpha + \zeta(\alpha) - 1 \\ &\quad + O\left(\frac{1 \wedge (n(1 - q_n))}{m_n^\alpha(1 - q_n)}\right) + O\left(\frac{1}{m_n^{\alpha-1}}\right) + O(m_n(1 - q_n)). \end{aligned}$$

It thus suffices to show that

$$(B.3) \quad O\left(\frac{1 \wedge (n(1 - q_n))}{m_n^\alpha(1 - q_n)}\right) + O\left(\frac{1}{m_n^{\alpha-1}}\right) + O(m_n(1 - q_n)) = O\left(\left((1 - q_n) \vee \frac{1}{n}\right)^{\frac{\alpha-1}{\alpha+1}}\right).$$

We divide the proof into three cases according to the value of q_n .

The first case is $\frac{1}{1 - q_n} \leq n$. In this case, we have

$$m_n = \left\lfloor \frac{1}{(1 - q_n)^{\frac{2}{\alpha+1}}} \right\rfloor = \Theta\left(\frac{1}{(1 - q_n)^{\frac{2}{\alpha+1}}}\right),$$

and (B.3) holds since

$$\frac{1 \wedge (n(1 - q_n))}{(m_n)^\alpha(1 - q_n)} = \Theta\left(\frac{(1 - q_n)^{\frac{2\alpha}{\alpha+1}}}{(1 - q_n)}\right) = \Theta((1 - q_n)^{\frac{\alpha-1}{\alpha+1}}),$$

and

$$\frac{1}{(m_n)^{\alpha-1}} = \Theta((1 - q_n)^{\frac{2(\alpha-1)}{\alpha+1}}) = O((1 - q_n)^{\frac{\alpha-1}{\alpha+1}}),$$

and

$$m_n(1 - q_n) = \Theta\left(\frac{1 - q_n}{(1 - q_n)^{\frac{2}{\alpha+1}}}\right) = \Theta((1 - q_n)^{\frac{\alpha-1}{\alpha+1}}).$$

The second case is $n < \frac{1}{1 - q_n} \leq n^\alpha$. In this case, note that $(1 - q_n) \leq \frac{1}{n}$. We also have

$$m_n = \left\lfloor \left(\frac{n}{1 - q_n}\right)^{\frac{1}{\alpha+1}} \right\rfloor = \Theta\left(\left(\frac{n}{1 - q_n}\right)^{\frac{1}{\alpha+1}}\right),$$

and (B.3) holds since

$$\frac{1 \wedge (n(1 - q_n))}{(m_n)^\alpha(1 - q_n)} = \Theta\left(n\left(\frac{1 - q_n}{n}\right)^{\frac{\alpha}{\alpha+1}}\right) = \Theta(n^{\frac{1}{\alpha+1}}(1 - q_n)^{\frac{\alpha}{\alpha+1}}) = O\left(\frac{1}{n^{\frac{\alpha-1}{\alpha+1}}}\right),$$

and

$$\frac{1}{(m_n)^{\alpha-1}} = \Theta\left(\left(\frac{(1-q_n)}{n}\right)^{\frac{\alpha-1}{\alpha+1}}\right) = O\left(\frac{1}{n^{\frac{2(\alpha-1)}{\alpha+1}}}\right) = O\left(\frac{1}{n^{\frac{\alpha-1}{\alpha+1}}}\right),$$

and

$$m_n(1-q_n) = \Theta\left(\left(\frac{n}{1-q_n}\right)^{\frac{1}{\alpha+1}}(1-q_n)\right) = \Theta\left(n^{\frac{1}{\alpha+1}}(1-q_n)^{\frac{\alpha}{\alpha+1}}\right) = O\left(\frac{1}{n^{\frac{\alpha-1}{\alpha+1}}}\right).$$

The third case is $\frac{1}{1-q_n} > n^\alpha$. In this case, we have $m_n = n$ and (B.3) holds since

$$\frac{1 \wedge (n(1-q_n))}{(m_n)^\alpha(1-q_n)} = \frac{n}{n^\alpha} = O\left(\frac{1}{n^{\frac{\alpha-1}{\alpha+1}}}\right),$$

and

$$\frac{1}{(m_n)^{\alpha-1}} = \frac{1}{n^{\alpha-1}} = O\left(\frac{1}{n^{\frac{\alpha-1}{\alpha+1}}}\right),$$

and

$$m_n(1-q_n) = n(1-q_n) < n \cdot \frac{1}{n^\alpha} = O\left(\frac{1}{n^{\frac{\alpha-1}{\alpha+1}}}\right).$$

This concludes the proof of the upper bound and of the proposition. \square

Acknowledgements. Both authors would like to thank an anonymous referee whose feedback both greatly improved the presentation of the paper and allowed us to shorten several of the proofs. BC also wishes to thank Ms. Legrand for supporting and encouraging his interest in mathematics.

Funding. During the preparation of this research, LAB was supported by an NSERC Discovery Grant and an FRQNT Team Grant, and BC was supported by an ISM Graduate Scholarship.

REFERENCES

- [1] ADDARIO-BERRY, L. and REED, B. (2009). Minima in branching random walks. *Ann. Probab.* **37** 1044–1079. MR2537549 <https://doi.org/10.1214/08-AOP428>
- [2] AÏDÉKON, E. (2013). Convergence in law of the minimum of a branching random walk. *Ann. Probab.* **41** 1362–1426. MR3098680 <https://doi.org/10.1214/12-AOP750>
- [3] ANGEL, O., HOLROYD, A. E., HUTCHCROFT, T. and LEVY, A. (2018). Mallows permutations as stable matchings. arXiv preprint [arXiv:1802.07142](https://arxiv.org/abs/1802.07142).
- [4] BASU, R. and BHATNAGAR, N. (2017). Limit theorems for longest monotone subsequences in random Mallows permutations. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 1934–1951. MR3729641 <https://doi.org/10.1214/16-AIHP777>
- [5] BENJAMINI, I., BERGER, N., HOFFMAN, C. and MOSSEL, E. (2005). Mixing times of the biased card shuffling and the asymmetric exclusion process. *Trans. Amer. Math. Soc.* **357** 3013–3029. MR2135733 <https://doi.org/10.1090/S0002-9947-05-03610-X>
- [6] BHATNAGAR, N. and PELED, R. (2015). Lengths of monotone subsequences in a Mallows permutation. *Probab. Theory Related Fields* **161** 719–780. MR3334280 <https://doi.org/10.1007/s00440-014-0559-7>
- [7] BIGGINS, J. D. (1976). The first- and last-birth problems for a multitype age-dependent branching process. *Adv. in Appl. Probab.* **8** 446–459. MR0420890 <https://doi.org/10.2307/1426138>
- [8] BROUTIN, N., DEVROYE, L., MCLEISH, E. and DE LA SALLE, M. (2008). The height of increasing trees. *Random Structures Algorithms* **32** 494–518. MR2422392 <https://doi.org/10.1002/rsa.20202>
- [9] CHAUVIN, B., KLEIN, T., MARCKERT, J.-F. and ROUAULT, A. (2005). Martingales and profile of binary search trees. *Electron. J. Probab.* **10** 420–435. MR2147314 <https://doi.org/10.1214/EJP.v10-257>

- [10] CRANE, H., DESALVO, S. and ELIZALDE, S. (2018). The probability of avoiding consecutive patterns in the Mallows distribution. *Random Structures Algorithms* **53** 417–447. MR3854041 <https://doi.org/10.1002/rsa.20776>
- [11] DEVROYE, L. (1986). A note on the height of binary search trees. *J. Assoc. Comput. Mach.* **33** 489–498. MR0849025 <https://doi.org/10.1145/5925.5930>
- [12] DIACONIS, P. and RAM, A. (2000). Analysis of systematic scan Metropolis algorithms using Iwahori–Hecke algebra techniques. *Michigan Math. J.* **48** 157–190. Dedicated to William Fulton on the occasion of his 60th birthday. MR1786485 <https://doi.org/10.1307/mmj/1030132713>
- [13] DRMOTA, M. (2003). An analytic approach to the height of binary search trees. II. *J. ACM* **50** 333–374. MR2146358 <https://doi.org/10.1145/765568.765572>
- [14] DRMOTA, M. (2009). The height of increasing trees. *Ann. Comb.* **12** 373–402. MR2496124 <https://doi.org/10.1007/s00026-009-0009-x>
- [15] DURRETT, R. (2019). *Probability—Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics **49**. Cambridge Univ. Press, Cambridge. Fifth edition of [MR1068527]. MR3930614 <https://doi.org/10.1017/9781108591034>
- [16] EVANS, S. N., GRÜBEL, R. and WAKOLBINGER, A. (2012). Trickle-down processes and their boundaries. *Electron. J. Probab.* **17** no. 1, 58. MR2869248 <https://doi.org/10.1214/EJP.v17-1698>
- [17] GLADKICH, A. and PELED, R. (2018). On the cycle structure of Mallows permutations. *Ann. Probab.* **46** 1114–1169. MR3773382 <https://doi.org/10.1214/17-AOP1202>
- [18] GNEDIN, A. and OLSHANSKI, G. (2010). q -exchangeability via quasi-invariance. *Ann. Probab.* **38** 2103–2135. MR2683626 <https://doi.org/10.1214/10-AOP536>
- [19] GNEDIN, A. and OLSHANSKI, G. (2012). The two-sided infinite extension of the Mallows model for random permutations. *Adv. in Appl. Math.* **48** 615–639. MR2920835 <https://doi.org/10.1016/j.aam.2012.01.001>
- [20] HAMMERSLEY, J. M. (1974). Postulates for subadditive processes. *Ann. Probab.* **2** 652–680. MR0370721 <https://doi.org/10.1214/aop/1176996611>
- [21] HE, J. (2020). A central limit theorem for descents of a Mallows permutation and its inverse. arXiv preprint [arXiv:2005.09802](https://arxiv.org/abs/2005.09802).
- [22] KINGMAN, J. F. C. (1975). The first birth problem for an age-dependent branching process. *Ann. Probab.* **3** 790–801. MR0400438 <https://doi.org/10.1214/aop/1176996266>
- [23] MALLOWS, C. L. (1957). Non-null ranking models. I. *Biometrika* **44** 114–130. MR0087267 <https://doi.org/10.1093/biomet/44.1-2.114>
- [24] MEAD, D. G. (1992). Newton’s identities. *Amer. Math. Monthly* **99** 749–751. MR1186460 <https://doi.org/10.2307/2324242>
- [25] MUELLER, C. and STARR, S. (2013). The length of the longest increasing subsequence of a random Mallows permutation. *J. Theoret. Probab.* **26** 514–540. MR3055815 <https://doi.org/10.1007/s10959-011-0364-5>
- [26] MUKHERJEE, S. (2016). Fixed points and cycle structure of random permutations. *Electron. J. Probab.* **21** Paper No. 40, 18. MR3515570 <https://doi.org/10.1214/16-EJP4622>
- [27] PINSKY, R. (2019). Permutations avoiding a pattern of length three under Mallows distributions. Preprint.
- [28] PITMAN, J. (1997). Some probabilistic aspects of set partitions. *Amer. Math. Monthly* **104** 201–209. MR1436042 <https://doi.org/10.2307/2974785>
- [29] PITTEL, B. (1984). On growing random binary trees. *J. Math. Anal. Appl.* **103** 461–480. MR0762569 [https://doi.org/10.1016/0022-247X\(84\)90141-0](https://doi.org/10.1016/0022-247X(84)90141-0)
- [30] REED, B. (2003). The height of a random binary search tree. *J. ACM* **50** 306–332. MR2146357 <https://doi.org/10.1145/765568.765571>
- [31] SHI, Z. (2015). *Branching Random Walks. Lecture Notes in Math.* **2151**. Springer, Cham. Lecture notes from the 42nd Probability Summer School held in Saint Flour, 2012, École d’Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. MR3444654 <https://doi.org/10.1007/978-3-319-25372-5>
- [32] STANLEY, R. P. (2012). *Enumerative Combinatorics. Volume 1*, 2nd ed. Cambridge Studies in Advanced Mathematics **49**. Cambridge Univ. Press, Cambridge. MR2868112
- [33] STARR, S. (2009). Thermodynamic limit for the Mallows model on S_n . *J. Math. Phys.* **50** 095208, 15. MR2566888 <https://doi.org/10.1063/1.3156746>
- [34] STARR, S. and WALTERS, M. (2018). Phase uniqueness for the Mallows measure on permutations. *J. Math. Phys.* **59** 063301, 28. MR3817550 <https://doi.org/10.1063/1.5017924>