J. Appl. Probab. **62,** 67–83 (2025)
*doi:*10.1017/jpr.2024.40

QUENCHED WORST-CASE SCENARIO FOR ROOT DELETION IN TARGETED CUTTING OF RANDOM RECURSIVE TREES

LAURA ESLAVA ,* **
SERGIO I. LÓPEZ ,* *** AND
MARCO L. ORTIZ ,* **** *Universidad Nacional Autónoma de México*

Abstract

We propose a method for cutting down a random recursive tree that focuses on its higher-degree vertices. Enumerate the vertices of a random recursive tree of size n according to the decreasing order of their degrees; namely, let $(v^{(i)})_{i=1}^{n}$ be such that $\deg(v^{(1)}) \geq \cdots \geq \deg(v^{(n)})$. The targeted vertex-cutting process is performed by sequentially removing vertices $v^{(1)}, v^{(2)}, \ldots, v^{(n)}$ and keeping only the subtree containing the root after each removal. The algorithm ends when the root is picked to be removed. The total number of steps for this procedure, K_n, is upper bounded by $Z_{\geq D}$, which denotes the number of vertices that have degree at least as large as the degree of the root. We prove that $\ln Z_{\geq D}$ grows as $\ln n$ asymptotically and obtain its limiting behavior in probability. Moreover, we obtain that the kth moment of $\ln Z_{\geq D}$ is proportional to $(\ln n)^k$. As a consequence, we obtain that the first-order growth of K_n is upper bounded by $n^{1-\ln 2}$, which is substantially smaller than the required number of removals if, instead, the vertices were selected uniformly at random.

Keywords: Random trees; cutting down process; random cuts; degree sequence

2020 Mathematics Subject Classification: Primary 05C80
Secondary 60F05

1. Introduction

The idea of cutting random labeled trees was first introduced by Meir and Moon in 1970 and, soon after, they considered random recursive trees (abbreviated as RRTs) [42, 43]. The original cutting procedure, for a given tree, is as follows. Start with a tree with n vertices. Choose an edge uniformly at random and remove it, along with the cut subtree that does not contain the root. Repeat until the remaining tree consists of only the root; at this point, we say that the tree has been *deleted*.

Random recursive trees are rooted trees, where each vertex has a unique label, obtained by the following procedure: Let T_1 be a single vertex labeled 1. For $n > 1$ the tree T_n is obtained

Received 7 December 2022; accepted 22 April 2024.
* Postal address: Instituto de investigación en Matemáticas Aplicadas y Sistemas, Universidad Nacional Autónoma de México, Av. Universidad 3000, Mexico City, C.P.04510.
** Email address: laura@sigma.iimas.unam.mx
*** Email address: silo@ciencias.unam.mx
**** Email address: malo@ciencias.unam.mx

from the tree T_{n-1} by adding an edge directed from a new vertex labeled n to a vertex with label in $\{1, \ldots, n-1\}$, chosen uniformly at random and independent for each n. We say that T_n has size n and that the degree of a vertex v is the number of edges directed towards v.

In this paper, we introduce a *targeted vertex-cutting procedure* for random recursive trees that focuses on high-degree vertices (defined in Section 1.3). Our interest in this procedure stems from the fact that it drastically reduces the number of cuts needed to destroy an RRT compared to random cuts (see Proposition 1.1), while this model may serve as a proxy for malicious attacks.

In the case of RRTs, the currently known properties of the structure of high-degree vertices do not hold to the distinct strategies previously used for studying the cutting of random trees. See [6] for a thorough survey on cutting procedures and other related processes on random recursive trees up to 2014. We give a brief overview of these and more recent models of cutting processes in Section 1.2.

To circumvent these difficulties, the core of the paper is drawn towards exploiting our knowledge on the number of high-degree vertices in RRTs. In Section 1.1, we present our main results, together with a brief description of their proof strategies. In Sections 1.2 and 1.3, we give an overview of previous cutting models and define the targeted cutting model; we also discuss in more detail the difficulties in studying the targeted cutting of RRTs and examine the implications of Theorem 1.1 on the targeted cutting process. In a nutshell, Proposition 1.1 supports the heuristic that malicious attacks may hold a strategy to take better advantage of some characteristics of the network. That is, the number of cuts required is significantly fewer than in a completely random attack. The proofs of the main results, Theorems 1.1 and 1.2, are presented in Sections 2 and 3, respectively. We close with simulations and further avenues of research in Sections 4 and 5, respectively.

1.1. Main results

For $d \in \mathbb{N}$, we let Z_d and $Z_{\geq d}$ be the number of vertices in T_n that have degree d and degree at least d, respectively (we omit the dependence on n for these variables throughout). The variables $(Z_d)_{d\geq 0}$ are referred to as the degree sequence of T_n; its asymptotic joint distribution is described by the limiting normal distribution, as $n \to \infty$, of certain urn models in [32].

The tails of the degree sequence are predominantly studied in two different ranges. The sequence $(Z_{d+\lfloor \log_2 n \rfloor})_{d\in\mathbb{Z}}$ represents degrees close to the maximum degree. Its limiting joint distribution is described (along suitable subsequences) by an explicit Poisson point process in $\mathbb{Z} \cup \{\infty\}$ in [2]; the lattice shift by $\log_2 n$ stems from the fact that $\Delta_n/\log_2 n$, the renormalized maximum degree in T_n, converges almost surely (a.s.) to 1 [20]. More generally, high-degree vertices are counted by $Z_{\geq c \ln n}$, for $c > 0$, and their moments are described in [2]. The techniques developed in [24] provide bounds on the total variation distance between $Z_{\geq c \ln n}$ and a Poisson random variable with mean n^{α} with a suitable $\alpha = \alpha(c)$ for $c \in (1, \log_2 n)$.

Our main results concern $Z_{\geq D}$, a tail of the degree sequence of T_n with random index. More precisely, let $D = D(n)$ denote the degree of the root in T_n and let $Z_{\geq D}$ denote the number of vertices with degree at least D. Although the root degree D is concentrated around $\ln n$ (see, e.g., (2.5)), its tails do not vanish sufficiently fast. This leads us to consider the variable $\ln Z_{\geq D}$ instead of directly analyzing $Z_{\geq D}$.

https://doi.org/10.1017/jpr.2024.40 Published online by Cambridge University Press

The first theorem provides convergence in probability, after a normalization by $\ln n$. Its proof is based on the concentration of D and the first- and second-moment methods for suitable tails $Z_{\geq d}$; see Section 2.

Theorem 1.1. *Let $\gamma := 1 - \ln 2$; then the following convergence in probability holds:*

$$\frac{\ln Z_{\geq D}}{\ln n} \xrightarrow{\mathbb{P}} \gamma.$$

We also obtain matching asymptotics for the moments of $\ln Z_{\geq D}/\ln n$.

Theorem 1.2. *Let $\gamma := 1 - \ln 2$. For any positive integer k,*

$$\mathbb{E}\left[(\ln Z_{\geq D})^k\right] = (\gamma \ln n)^k (1 + o(1)).$$

To establish Theorem 1.2, we resort to a coupling between a random recursive tree T_n of size n and a random recursive tree $T_n^{(\varepsilon)}$ of size n conditioned on D to take values in $((1-\varepsilon)\ln n, (1+\varepsilon)\ln n)$, for any given $\varepsilon \in (0, 1)$; see Section 3.

Briefly described, the coupling is the following. Let $n \in \mathbb{N}$. For the construction of both T_n and $T_n^{(\varepsilon)}$, it suffices to define the parent of vertex i, for each $1 < i \leq n$. For each tree, we break down this choice into two steps: First, sample a random Bernoulli to decide whether i attaches to the root or not. Then the parent of i is either the root or a uniformly sampled vertex among the rest of the possible parents. The Bernoulli variables corresponding to both T_n and $T_n^{(\varepsilon)}$ are coupled in such a way that the root degree has the desired distribution. Additionally, if a given vertex chooses a parent distinct from the root, then this choice is *precisely the same* for both the unconditioned and conditioned trees.

1.2. The random cutting process and related models

There is a wide variety of random trees for which the cutting process has been analyzed. The survey by Baur and Bertoin provides a detailed description of the advances from the introduction of cutting processes in the 1970s up to 2014 [6]. In this section we provide a panoramic overview of the references therein, recent progress on the k-cut model, and further literature on resilience of networks more broadly.

The random cutting process described in the first part of the introduction naturally extends to more complex processes on RRTs. For example, it leads to the Bolthausen–Sznitman coalescent in the case of RRTs [27]; the construction of both the cut-tree and the tree of component sizes [5, 6, 8, 9]. The last two processes are tightly related to percolation clusters of the trees.

In this section we only provide details on the results for X_n, the number of cuts needed to isolate the root of an RRT of size n, under random edge cuts. Using the splitting property, [43] proved that, as $n \to \infty$,

$$\frac{\mathbb{E}[X_n] \ln n}{n} \to 1. \tag{1.1}$$

The splitting property refers to the fact that if we remove an edge from an RRT, then the two subtrees resulting from the split are in turn, conditionally on their sizes, independent RRTs. This property has been also used in [23, 30, 45].

In particular, [45] improved (1.1) to convergence in probability by extending the analysis to the *cost* of the random cutting procedure; that is, each cut has a cost in a function of the size of the removed tree. The random variable X_n is recovered when the cost is a constant function.

Finally, the limiting distribution of X_n, after a suitable rescaling, was obtained in [23] via singularity analysis and in [30] using probabilistic and recursive methods. Namely,

$$Y_n := \frac{(\ln n)^2}{n} X_n - \ln n - \ln \ln n$$

converges weakly to a complete asymmetric Cauchy variable Y with characteristic function given by

$$\varphi_Y(\lambda) := \exp\left\{ \mathrm{i}\lambda \ln |\lambda| - \frac{\pi |\lambda|}{2} \right\}.$$

Other methods of selecting edges in an RRT include using the order of the vertex labels. This directly means that the number of removals to isolate the root is precisely the number of vertices attached to it; however, more can be said about the cluster evolution if we do not discard the cut trees during the process [6]. [35] modified the process with the objective of isolating the last vertex added to the network. [36–38] focused specifically on isolating a distinguished vertex; for example, the isolation of a leaf or isolating multiple distinguished vertices.

More recently, [14, 15] introduced the k-cut model, in which edges are selected uniformly at random but these are removed only after their kth selection. After the initial results for paths and some trees, this analysis has been applied to complete binary trees [16], Galton–Watson trees [11], and the Brownian continuum random tree [46].

Broadly speaking, results on the k-cut process exploit a strategy based on counting records [31, 33]. This strategy consists of noting that the number of cuts needed to destroy a tree is equivalent to the number of records that arise from a random labeling of the edges. This method has also been applied for split trees and binary trees [28, 29].

Cutting processes of random trees and graphs can serve as a proxy for the resilience of a particular network to breakdowns, either from intentional attacks or fortuitous events. What is considered a breakdown and resilience may differ depending on the context. In their work on RRTs, Meir and Moon consider contamination from a certain source within an organism, and we may think of the number of cuts as the necessary steps to isolate the source of the contamination [43]. On the other hand, malicious attacks may refer to failures that arise from a hacker or enemy trying to disconnect some given network of servers. Thus, malicious attacks can be mathematically described by targeted cutting toward highly connected servers. These ideas on resilience were posed in [17, 18] for scale-free graphs. Later, [12] compared random versus malicious attacks on scale-free graphs; more precisely, the comparison is on the cluster sizes in percolation versus cluster sizes after removing a constant number of the smallest-labeled vertices.

As first described in [44], the Albert–Barabási random tree is part of the large class of scale-free graphs that exhibit *persistence* of node centrality. That is, for any fixed $k \in \mathbb{N}$, there is a finite time (a.s.) in the evolution of the Albert–Barabási random tree such that the identity of the k largest degree vertices does not change. Interestingly for our purposes, RRTs do not exhibit such persistence of node centrality; evidence of this fact can be found in [2]. [4] studied the persistence of a large class of evolving networks.

Finally, we name a few similar procedures in more general graphs. [7] studied public transport networks; [47] studied a broad range of dynamical processes, including birth–death processes, regulatory dynamics, and epidemic processes on scale-free graphs and Erdös–Renyi networks under two different targeted attacks, with both high-degree and low-degree vertices; [3] compared different graph metrics, such as node betweenness, node centrality, and node degree, to measure network resilience to random and directed attacks. In the context of Galton–Watson trees, cutting down trees has been predominantly studied from the perspective of vertex cutting; that is, vertices are selected to be removed; see, e.g., [1, 10, 21].

1.3. Targeted cutting for RRTs

Let T_n be an RRT of size n and enumerate its vertices as $(v^{(i)})_{i=1}^n$ according to the decreasing order of their degrees; that is, $\deg(v^{(1)}) \geq \cdots \geq \deg(v^{(n)})$, breaking ties uniformly at random. The targeted vertex-cutting process is performed by sequentially removing vertices $v^{(1)}, v^{(2)}, \ldots, v^{(n)}$ and keeping only the subtree containing the root after each removal (skip a step if the chosen vertex has previously been removed). The procedure ends when the root is picked to be removed.

Let K_n denote the number of vertex deletions before we select the root to be removed. Similarly to the random cutting of trees, K_n corresponds to the number of (largest) records on vertex degrees along paths from the root towards the leaves. However, there is a lack of understanding of the structure of vertices with high degree. The study of the location of high-degree vertices was initiated by the first author by establishing asymptotic results on the depth of vertices conditioned to have high degrees [25]. The same approach was later extended to include the label of such vertices [39]. Both the maximum degree and the label of a vertex attaining such a maximum have been asymptotically described for a more general class of weighted RRTs [26, 40, 41].

Also, the splitting property cannot be exploited in the targeted cutting procedure. In this case, the first vertex that is deleted, $v^{(1)}$, is conditioned on having the largest degree. To the best of our knowledge, the distribution of the size of the subtree rooted at $v^{(1)}$ remains unknown.

Now, recall that $Z_{\geq D}$ denotes the number of vertices in T_n which have degree at least the degree of the root. This random variable corresponds to the *quenched worst-case scenario* for the deletion time of the targeted cutting. That is, once the tree has been sampled and all the degrees are known, $Z_{\geq D}$ is a natural upper bound for K_n: when all the vertices in the list which have degree at least D have been deleted, then the root (and so the complete tree) has been deleted.

Theorem 1.1 provides a stochastic upper bound for K_n, the deletion time of the targeted cutting process in RRTs.

Proposition 1.1. *The random number of cuts K_n in the targeted cutting of T_n satisfies, for any $\varepsilon > 0$, $K_n = O_p(n^{\gamma+\varepsilon})$, where $\gamma := 1 - \ln 2$. Namely, for each $\delta > 0$, there are $M = M(\varepsilon, \delta) > 0$ and $N = N(\varepsilon, \delta) \in \mathbb{N}$ such that, for all $n \geq N$, $\mathbb{P}(K_n > Mn^{\gamma+\varepsilon}) < \delta$.*

Proof. Theorem 1.1 implies that the result is satisfied for $\varepsilon, \delta > 0$. Indeed, let N be large enough that $\mathbb{P}(Z_{\geq D} \geq n^{\gamma+\varepsilon}) < \delta$ for $n \geq N$. The result follows, setting $M(\varepsilon, \delta) = 1$, since $K_n \leq Z_{\geq D}$ a.s. □

Note that removing a uniformly random edge is equivalent to removing a uniformly chosen vertex other than the root. This difference may be neglected in the case of RRTs, so we can compare (1.1), in its probability version, with $K_n = O_p(n^{\gamma+\varepsilon})$. While Proposition 1.1 provides only a stochastic upper bound for K_n, it proves that the targeted cutting deletion time is significantly smaller compared to the deletion time for uniformly random removals (note that $1 - \ln 2 \approx 0.306\,85$).

The upper bound $K_n \leq Z_{\geq D}$ may not be tight for two reasons. First, some of the vertices in the list may not be deleted, but rather removed when some ancestor is deleted, and second, the

https://doi.org/10.1017/jpr.2024.40 Published online by Cambridge University Press

root may not be the last vertex, among those with the same degree, to be deleted. We perform simulations that suggest K_n and $Z_{\geq D}$ have the same growth order; see Section 4.

1.4. Notation

We use $|A|$ to denote the cardinality of a set A. For $n \in \mathbb{N}$ we write $[n] := \{1, 2, \ldots, n\}$ and, for $m < n$, $[m, n] = \{m, m+1, \ldots, n\}$. For $a, b, c, x \in \mathbb{R}$ with $b, c > 0$, we use $x \in (a \pm b)c$ as an abbreviation for $(a-b)c \leq x \leq (a+b)c$. In what follows we denote natural and base 2 logarithms by ln and log, respectively. We often use the identity $\ln a = \ln 2 \cdot \log a$ for $a > 0$.

For $r \in \mathbb{R}$ and $a \in \mathbb{N}$ define the *falling factorial* $(r)_a := r(r-1)\cdots(r-a+1)$ and $(r)_0 = 1$. For real functions f, g we write $f(x) = o(g(x))$ when $\lim_{x\to\infty} f(x)/g(x) = 0$ and $f(x) = O(g(x))$ when $|f(x)/g(x)| \leq C$ for some $C > 0$. Convergence in probability will be written as $\xrightarrow{\mathbb{P}}$. H_n denotes the nth *harmonic number* $H_n = \sum_{i=1}^{n} 1/i$. A rooted tree is a tree with a distinguished vertex, which we call the root. We always consider the edges to be directed towards the root. A directed edge $e = uv$ is directed from u to v and, in this case, we say that v is the parent of u. Given a rooted tree T and one of its vertices v, the degree of v in T, denoted by $\deg_T(v) = d_T(v)$, is the number of edges directed towards v. We say that T has size n if it has n vertices, and $[n]$ will denote the set of vertices of a tree of size n.

2. Deterministic tails of the degree sequence

Given a random recursive tree T_n, let $Z_{\geq d}$ denote the number of vertices with degree at least d, that is

$$Z_{\geq d} \equiv Z_{\geq d}(n) := |\{v \in [n] : d_{T_n}(v) \geq d\}|. \tag{2.1}$$

Recall that $Z_{\geq D}$ has random index $D = D(n)$ which corresponds to the degree of the root in T_n. Our theorems build upon results on the convergence of the variables $Z_{\geq d}$; note that these variables are non-increasing on d. The following two propositions are simplified versions of [2, Proposition 2.1] and [24, Theorem 1.8], respectively.

Proposition 2.1. (Moments of $Z_{\geq d}$.) *For any $d \in \mathbb{N}$, $\mathbb{E}[Z_{\geq d}] \leq 2^{\log n - d}$. Moreover, there exists $\alpha > 0$ such that, if $d < \frac{3}{2}\ln n$ and $k \in \{1, 2\}$ then*

$$\begin{aligned} \mathbb{E}[(Z_{\geq d})_k] &= 2^{(\log n - d)k}(1 + o(n^{-\alpha})), \\ \mathbb{E}\big[Z_{\geq d}^2\big] &= \mathbb{E}[Z_{\geq d}]^2(1 + o(n^{-\alpha})). \end{aligned} \tag{2.2}$$

Proposition 2.2. (Total variation distance.) *Let $0 < \varepsilon < \frac{1}{3}$. Then for $(1+\varepsilon)\ln n < d < (1+3\varepsilon)\ln n$ there exists $\alpha' = \alpha'(\varepsilon) > 0$ such that* $\mathrm{d_{TV}}(Z_{\geq d}, \mathrm{Poi}(\mathbb{E}[Z_{\geq d}])) \leq O(n^{-\alpha'})$.

As we mentioned before, the error bounds in the previous propositions are not strong enough to estimate the moments of $Z_{\geq D}$. Instead, we focus on the variable $\ln Z_{\geq D}$. Furthermore, we transfer the task of moment estimation, for Theorem 1.2, to $\mathbb{E}[(\ln X)^\ell]$ where instead of having $X = Z_{\geq D}$ we consider either $1 + Z_{\geq m_-}$ or $1 + Z_{\geq m_+}$ for suitable values $m_\pm = c_\pm \ln n$ with $c_- < 1 < c_+$.

The upper bound in the following proposition follows from a straightforward application of Jensen's inequality, together with Proposition 2.1, while the lower bound uses the refined bounds given in Proposition 2.2.

Proposition 2.3. *Let $\varepsilon \in (0, \frac{1}{3})$ and $\ell \in \mathbb{N}$. Let*

$$m_- := \left\lfloor \left(1 - \frac{\varepsilon}{2\ln 2}\right) \ln n \right\rfloor, \qquad m_+ := \left\lceil \left(1 + \frac{\varepsilon}{2\ln 2}\right) \ln n + 1 \right\rceil.$$

There exist $\alpha' = \alpha'(\varepsilon) > 0$ and $C_\ell = C_\ell(\varepsilon) > 0$ such that

$$\mathbb{E}[(\ln(1 + Z_{\geq m_-}))^\ell] \leq [(1 - \ln 2 + \varepsilon) \ln n]^\ell + C_\ell, \tag{2.3}$$

$$\mathbb{E}[(\ln(1 + Z_{\geq m_+}))^\ell] \geq [(1 - \ln 2 - \varepsilon) \ln n]^\ell (1 - O(n^{-\alpha'})). \tag{2.4}$$

Proof. First, let $f(x) = (\ln(x))^\ell$ and note that $f''(x) \leq 0$ for $x > \mathrm{e}^{\ell-1}$. Hence, by Jensen's inequality, for any positive integer-valued random variable X,

$$\mathbb{E}[(\ln X)^\ell] = \mathbb{E}\big[(\ln X)^\ell \mathbf{1}_{\{X > \mathrm{e}^{\ell-1}\}}\big] + \mathbb{E}\big[(\ln X)^\ell \mathbf{1}_{\{X \leq \mathrm{e}^{\ell-1}\}}\big] \leq (\ln \mathbb{E}[X])^\ell + (\ell - 1)^\ell.$$

Next, we use the upper bound in Proposition 2.1 for $\mathbb{E}[Z_{\geq m_-}]$. Note that $\log n - m_- \leq (1 - \ln 2 + \varepsilon/2) \log n$. Thus, $\mathbb{E}[1 + Z_{\geq m_-}] \leq 1 + 2n^{1-\ln 2+\varepsilon/2} \leq n^{1-\ln 2+\varepsilon}$, where the second inequality holds for $n \geq n_0 = n_0(\varepsilon)$ large enough. Then

$$\mathbb{E}[(\ln(1 + Z_{\geq m_-}))^\ell] \leq [\,\ln \mathbb{E}[1 + Z_{\geq m_-}]]^\ell + (\ell - 1)^\ell \leq [(1 - \ln 2 + \varepsilon) \ln n]^\ell + C_\ell,$$

where $C_\ell = \sup_{n \leq n_0}\{(\ln(1 + 2n^{1-\ln 2+\varepsilon/2}))^\ell\} + (\ell - 1)^\ell$.

Next, let $\mu = \mathbb{E}[Z_{\geq m_+}]$ and let (X, X') be random variables coupled so that $X \overset{\mathrm{D}}{=} Z_{\geq m_+}$, $X' \overset{\mathrm{D}}{=} \mathrm{Poi}(\mu)$, and $\mathbb{P}(X = X')$ is maximized; that is, $\mathbb{P}(X \neq X') = \mathrm{d_{TV}}(X, X')$. Note that

$$\mathbb{E}[(\ln(1 + X))^\ell] \geq \mathbb{E}\big[(\ln(1 + X))^\ell \mathbf{1}_{[X > n^{\gamma-\varepsilon}]}\big] \geq (\ln(n^{\gamma-\varepsilon}))^\ell \mathbb{P}(X > n^{\gamma-\varepsilon});$$

then (2.4) boils down to lower bounding $\mathbb{P}(X > n^{\gamma-\varepsilon})$. Since $X < n$, by the coupling assumption, we have

$$\begin{aligned}\mathbb{P}(X > n^{\gamma-\varepsilon}) &\geq \mathbb{P}(n^{\gamma-\varepsilon} < X' < n) - \mathbb{P}(X \neq X') \\ &= 1 - \mathbb{P}(X' \geq n) - \mathbb{P}(X' \leq n^{\gamma-\varepsilon}) - \mathrm{d_{TV}}(X, X').\end{aligned}$$

By Proposition 2.2, $\mathrm{d_{TV}}(X, X') = O(n^{-\alpha'})$ for $\alpha'(\varepsilon) > 0$. Using the Chernoff bounds for the tails of a Poisson variable (see, e.g., [13, Section 2.2]) and that $\mu = n^{\gamma-\varepsilon/2}(1 + o(1))$ we have

$$\begin{aligned}\mathbb{P}(X' \geq n) &\leq \left(\frac{\mathrm{e}n^{\gamma-\varepsilon/2}}{n}\right)^n \mathrm{e}^{-n^{\gamma-\varepsilon/2}} \leq \left(\frac{\mathrm{e}n^{\ln 2}}{\mathrm{e}n^{\ln 2}}\right)^{-n}, \\ \mathbb{P}(X' \leq n^{\gamma-\varepsilon}) &\leq \left(\frac{\mathrm{e}n^{\gamma-\varepsilon/2}}{n^{\gamma-\varepsilon}}\right)^{n^{\gamma-\varepsilon}} \mathrm{e}^{-n^{\gamma-\varepsilon/2}} \leq \left(\frac{\mathrm{e}n^{\varepsilon/2}}{\mathrm{e}n^{\varepsilon/2}}\right)^{-n^{\gamma-\varepsilon}};\end{aligned}$$

both bounds are $o(n^{-\alpha'})$, so the proof is completed. □

https://doi.org/10.1017/jpr.2024.40 Published online by Cambridge University Press

2.1. Proof of Theorem 1.1

We use the first- and second-moment methods, together with Proposition 2.1, the concentration of D, and the fact that, for each n, $Z_{\geq m}$ is non-increasing in m.

Proof of Theorem 1.1. For completeness, we show that D is concentrated around $\ln n$ (in fact, it is asymptotically normal; see, e.g., [22, Section 3.1.1, (6.3)] or [19, Section 3]). Indeed, D is a sum of independent Bernoulli random variables $(B_i)_{1<i\leq n}$, each with mean $1/(i-1)$, and so $\mathbb{E}[D] = H_{n-1} > \ln n$. From the fact that $H_n - \ln n$ is a decreasing sequence we infer that, for any $0 < \varepsilon < \frac{3}{2}$ and n sufficiently large, $|D - H_{n-1}| \leq (\varepsilon/2)H_{n-1}$ implies $|D - \ln n| \leq \varepsilon \ln n$. Using the contrapositive of this statement and Bernstein's inequality (see, e.g., [34, Theorem 2.8]) we obtain, for n large enough,

$$\mathbb{P}(D \notin (1 \pm \varepsilon) \ln n) \leq \mathbb{P}(|D - H_{n-1}| > (\varepsilon/2)H_{n-1}) \leq 2 \exp\left\{-\frac{\varepsilon^2}{12} H_{n-1}\right\} \leq 2n^{-\varepsilon^2/12}. \quad (2.5)$$

Recall that $\gamma = 1 - \ln 2$. It suffices to prove that, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(Z_{\geq D} \notin (n^{\gamma-\varepsilon}, n^{\gamma+\varepsilon})) = 0.$$

We infer from (2.5) that $\mathbb{P}(Z_{\geq D} \notin (n^{\gamma-\varepsilon}, n^{\gamma+\varepsilon}), D \notin (1 \pm \varepsilon) \ln n)$ vanishes as n grows.

Let $m_- := m_-(\varepsilon) = \lfloor (1-\varepsilon) \ln n \rfloor$ and $m_+ := m_+(\varepsilon) = \lceil (1+\varepsilon) \ln n \rceil$. Using the monotonicity of $Z_{\geq m}$ on m, we have

$$\mathbb{P}(Z_{\geq D} \notin (n^{\gamma-\varepsilon}, n^{\gamma+\varepsilon}), D \in (1 \pm \varepsilon) \ln n) \leq \mathbb{P}(Z_{\geq m_-} \geq n^{\gamma+\varepsilon}) + \mathbb{P}(Z_{\geq m_+} \leq n^{\gamma-\varepsilon}), \quad (2.6)$$

so it remains to show that both terms on the right-hand side of (2.6) vanish. First, using that $\ln n = \ln 2 \cdot \log n$ and so $\log n - (1 \pm \varepsilon) \ln n = (1 - \ln 2 \mp \varepsilon \ln 2) \log n$, we infer from Proposition 2.1 that

$$\frac{1}{2} n^{\gamma - \varepsilon \ln 2}(1 - o(n^{-\alpha})) \leq \mathbb{E}[Z_{\geq m_+}] \leq \mathbb{E}[Z_{\geq m_-}] \leq 2n^{\gamma + \varepsilon \ln 2}.$$

Markov's inequality then gives $\mathbb{P}(Z_{\geq m_-} \geq n^{\gamma+\varepsilon}) \leq 2n^{-\varepsilon\gamma} \to 0$. Next, let θ be defined so that $n^{\gamma-\varepsilon} = \theta \mathbb{E}[Z_{\geq m_+}]$; in particular, $\theta \leq 2n^{-\varepsilon\gamma}$. The Paley–Zygmund inequality gives

$$\mathbb{P}(Z_{\geq m_+} > n^{\gamma-\varepsilon}) \geq \mathbb{P}(Z_{\geq m_+} > \theta \mathbb{E}[Z_{\geq m_+}]) \geq (1-\theta)^2 \frac{\mathbb{E}[Z_{\geq m_+}]^2}{\mathbb{E}[Z_{\geq m_+}^2]},$$

which tends to 1 as $n \to \infty$ by the upper bound for θ and (2.2). This implies that $\mathbb{P}(Z_{\geq m_+} \leq n^{\gamma-\varepsilon})$ vanishes, as desired. □

3. Control of *D* through a coupling

Let $n \in \mathbb{N}$. Write $\mathcal{I}_n$ for the set of increasing trees of size n; namely, labeled rooted trees with label set $[n]$ such that the vertex labels are increasing along any path starting from the root. It is straightforward to verify that the law of T_n is precisely the uniform distribution on $\mathcal{I}_n$.

Consider the following construction of an increasing tree of size n. Let $(b_i)_{1<i\leq n}$ and $(y_i)_{1<i\leq n}$ be integer-valued sequences such that $b_2 = y_2 = 1$, $b_i \in \{0, 1\}$, and $2 \leq y_i \leq i-1$ for $3 \leq i \leq n$. Let the vertex labeled 1 be the root and, for each $1 < i \leq n$, let vertex i be connected to vertex 1 if $b_i = 1$; otherwise, let vertex i be connected to vertex y_i.

https://doi.org/10.1017/jpr.2024.40 Published online by Cambridge University Press

The following coupling is exploited in the proof of Theorem 1.2. Define random vectors $(B_i)_{1<i\le n}$, $\big(B_i^{(\varepsilon)}\big)_{1<i\le n}$, and $(Y_i)_{1<i\le n}$ as follows. Let $(B_i)_{1<i\le n}$ be independent Bernoulli$(1/(i-1))$ random variables, let $\big(B_i^{(\varepsilon)}\big)_{1<i\le n}$ have the law of $(B_i)_{1<i\le n}$ conditioned on $\sum_{i=2}^n B_i \in (1\pm\varepsilon)\ln(n)$, and let $(Y_i)_{1<i\le n}$ be independent random variables such that $Y_2=1$ a.s. and Y_i is uniform over $\{2, \ldots, i-1\}$ for $3\le i\le n$. We assume that the vector $(Y_i)_{1<i\le n}$ is independent from the rest, while the coupling of $(B_i)_{1<i\le n}$ and $\big(B_i^{(\varepsilon)}\big)_{1<i\le n}$ is arbitrary.

The tree obtained from $(B_i)_{1<i\le n}$, $(Y_i)_{1<i\le n}$, and the construction above has the distribution of an RRT. To see this, write v_i for the parent of vertex i; then $1\le v_i<i$ for each $1<i\le n$. First, note that each v_i is independent from the rest since $(B_i)_{1<i\le n}$ and $(Y_i)_{1<i\le n}$ are independent. Next, we show that v_i is chosen uniformly at random from $\{1, 2, \ldots, i-1\}$. First, we have $v_2=1$ almost surely. For $2\le \ell<i\le n$, by the independence of B_i and Y_i, we have

$$\mathbb{P}(v_i=1)=\mathbb{P}(B_i=1)=\frac{1}{i-1}=\mathbb{P}(B_i=0,\, Y_i=\ell)=\mathbb{P}(v_i=\ell);$$

therefore, the tree obtained has the law of an RRT and so we denote it by T_n. Analogously, write $T_n^{(\varepsilon)}$ for the tree obtained from $\big(B_i^{(\varepsilon)}\big)_{1<i\le n}$ and $(Y_i)_{1<i\le n}$, and let $D^{(\varepsilon)}$ be the degree of its root.

By the definition of D and the construction above we have $D=\sum_{i=2}^n B_i$. Thus, conditioning on $\sum_{i=2}^n B_i$ means, under this construction, conditioning T_n on the root degree D. In particular, the distribution of $\big(B_i^{(\varepsilon)}\big)_{1<i\le n}$ is defined so that $T_n^{(\varepsilon)}$ has the distribution of an RRT of size n conditioned on $D^{(\varepsilon)}\in(1\pm\varepsilon)\ln(n)$.

The condition $D^{(\varepsilon)}\in(1\pm\varepsilon)\ln n$ corresponds to an event of probability close to 1 and so the degree sequences of T_n and $T_n^{(\varepsilon)}$ do not differ by much. Hence, the proof strategy of Theorem 1.2 is to estimate the moments $\mathbb{E}[(\ln Z_{\ge D})^k]$ using the monotonicity of $Z_{\ge d}$, by conditioning on $D\in(1\pm\varepsilon)\ln n$ while retaining $Z_{\ge d}$ instead of $Z_{\ge d}^{(\varepsilon)}$; see (3.9)–(3.11).

The following two propositions make this idea rigorous. For $d\ge 0$, let

$$W_d=\frac{1+Z_{\ge d}^{(\varepsilon)}}{1+Z_{\ge d}},$$

where $Z_{\ge d}$ is defined as in (2.1) and, similarly, let $Z_{\ge d}^{(\varepsilon)} := \Big|\Big\{v\in T_n^{(\varepsilon)} : d_{T_n^{(\varepsilon)}}(v)\ge d\Big\}\Big|$.

The key in the proof of Proposition 3.1 lies in (3.3), which yields an upper bound on the number of vertices that have differing degrees in T_n and $T_n^{(\varepsilon)}$ under the coupling. In turn, this allows us to infer bounds on the ratio W_d that hold with high probability and are uniform in d.

Proposition 3.1. *Let $\varepsilon, \delta\in(0, 1)$ and $0\le d\le(1+\varepsilon)\ln n$. There are $C>0$, $\beta=\beta(\varepsilon)>0$, and $n_0=n_0(\delta)$ such that, for $n\ge n_0$, under the coupling described above, $\mathbb{P}(W_d\in(1\pm\delta))\ge 1-Cn^{-\beta}$.*

Proof. Let $m_+=\lceil(1+\varepsilon)\ln n\rceil$, so that $Z_{\ge d}\ge Z_{\ge m_+}$. By Chebyshev's inequality and (2.2), for $c\in(0, 1)$,

$$\mathbb{P}(Z_{\ge m_+}\le c\mathbb{E}[Z_{\ge m_+}])\le\mathbb{P}(|Z_{\ge m_+}-\mathbb{E}[Z_{\ge m_+}]|\ge(1-c)\mathbb{E}[Z_{\ge m_+}])=o(n^{-\alpha}). \tag{3.1}$$

Rewrite

$$W_d=1+\frac{Z_{\ge d}^{(\varepsilon)}-Z_{\ge d}}{1+Z_{\ge d}}$$

https://doi.org/10.1017/jpr.2024.40 Published online by Cambridge University Press

to see that $W_d \in (1 \pm \delta)$ is equivalent to $\left|Z^{(\varepsilon)}_{\geq d} - Z_{\geq d}\right| \in [0, \delta(1 + Z_{\geq d}))$. Hence, it suffices to show that there is $n_0(\delta) \in \mathbb{N}$ such that, for $n \geq n_0$,

$$\{Z_{\geq m_+} \geq c\mathbb{E}[Z_{\geq m_+}],\, D \in (1 \pm \varepsilon) \ln n\} \subset \left\{\left|Z^{(\varepsilon)}_{\geq d} - Z_{\geq d}\right| \leq \delta(1 + Z_{\geq d})\right\}, \tag{3.2}$$

which, by a contrapositive argument, together with (2.5) and (3.1), yields, for $\beta = \min\{\alpha, \varepsilon^2/12\} > 0$,

$$\mathbb{P}\left(\left|Z^{(\varepsilon)}_{\geq d} - Z_{\geq d}\right| > \delta(1 + Z_{\geq d})\right) \leq \mathbb{P}(D \notin (1 \pm \varepsilon) \ln n) + \mathbb{P}(Z_{\geq m_+} \leq c\mathbb{E}[Z_{\geq m_+}]) = O(n^{-\beta}).$$

We extend the notation introduced for the coupling. Let $\mathcal{S} := \{i \in [2, n] \colon v_i = v^{(\varepsilon)}_i\}$ be the set of vertices that have the same parent in T_n and $T^{(\varepsilon)}_n$. For $i \in [n]$, denote the set of children of i in T_n and $T^{(\varepsilon)}_n$, respectively, by

$$\mathcal{C}(i) := \{j \in [2, n] \colon v_j = i\}, \qquad \mathcal{C}^{(\varepsilon)}(i) := \{j \in [2, n] \colon v^{(\varepsilon)}_j = i\}.$$

By the coupling construction, $\mathcal{S} \cup (\mathcal{C}(1) \triangle \mathcal{C}^{(\varepsilon)}(1))$ is a partition of $[2,n]$; that is, whenever the parent of a vertex $i \in [2, n]$ differs in T_n and $T^{(\varepsilon)}_n$ we infer that $B_i \neq B^{(\varepsilon)}_i$ and so either $i \in \mathcal{C}(1) \setminus \mathcal{C}^{(\varepsilon)}(1)$ or $i \in \mathcal{C}(1)^{(\varepsilon)} \setminus \mathcal{C}(1)$. The consequences of this observation are two-fold: First, for any $i \in [2, n]$, a necessary condition for $d_{T_n}(i) \neq d_{T^{(\varepsilon)}_n}(i)$ is that $\mathcal{C}(i) \neq \mathcal{C}^{(\varepsilon)}(i)$; second, the function $j \mapsto Y_j$ that maps $\mathcal{C}(1) \triangle \mathcal{C}^{(\varepsilon)}(1)$ to $\{i \in [2, n] \colon \mathcal{C}(i) \neq \mathcal{C}^{(\varepsilon)}(i)\}$ is surjective. Indeed, if $i \neq 1$ and $\mathcal{C}(i) \neq \mathcal{C}^{(\varepsilon)}(i)$ then there exists $j \in \mathcal{C}(1) \triangle \mathcal{C}^{(\varepsilon)}(1)$ such that $Y_j = i$. Together they imply the following chain of inequalities:

$$\left|\left\{i \in [2, n] \colon d_{T_n}(i) \neq d_{T^{(\varepsilon)}_n}(i)\right\}\right| \leq |\{i \in [2, n] \colon \mathcal{C}(i) \neq \mathcal{C}^{(\varepsilon)}(i)\}| \leq |\mathcal{C}(1) \triangle \mathcal{C}(1)^{(\varepsilon)}|; \tag{3.3}$$

the first inequality holds by containment of the corresponding sets, while the second one follows from the surjective function described above. On the other hand, $|Z_{\geq d} - Z^{(\varepsilon)}_{\geq d}|$ equals

$$\begin{aligned}\left|\sum_{i=1}^{n} \mathbf{1}_{\{d_{T_n}(i) \geq d\}}(i) - \sum_{i=1}^{n} \mathbf{1}_{\{d_{T^{(\varepsilon)}_n}(i) \geq d\}}(i)\right| &\leq \sum_{i=1}^{n} \left|\mathbf{1}_{\{d_{T_n}(i) \geq d\}}(i) - \mathbf{1}_{\{d_{T^{(\varepsilon)}_n}(i) \geq d\}}(i)\right| \\ &\leq 1 + \left|\left\{i \in \{2, \ldots, n\} \colon d_{T_n}(i) \neq d_{T^{(\varepsilon)}_n}(i)\right\}\right|.\end{aligned}$$

Therefore,

$$\left|Z_{\geq d} - Z^{(\varepsilon)}_{\geq d}\right| \leq 1 + |\mathcal{C}(1) \triangle \mathcal{C}^{(\varepsilon)}(1)| \leq 1 + 2\max\{|\mathcal{C}(1)|, |\mathcal{C}^{(\varepsilon)}(1)|\}. \tag{3.4}$$

We are now ready to prove (3.2). Fix, e.g., $c = \frac{1}{2}$ and let $n_0 = n_0(\delta)$ be large enough that $(1 + 4\ln n)/\delta < c\mathbb{E}[Z_{\geq m_+}]$; this is possible since $\mathbb{E}[Z_{\geq m_+}]$ grows polynomially in n, by Proposition 2.1. In particular, recalling that $Z_{\geq d} \geq Z_{\geq m_+}$ for $n \geq n_0$ and any $\varepsilon \in (0, 1)$, we have

$$\{Z_{\geq m_+} \geq c\mathbb{E}[Z_{\geq m_+}]\} \subset \left\{Z_{\geq d} \geq \frac{1 + 2(1 + \varepsilon) \ln n}{\delta}\right\}. \tag{3.5}$$

Moreover, by the construction of $T^{(\varepsilon)}_n$, $|\mathcal{C}(1)^{(\varepsilon)}| = D^{(\varepsilon)} \leq (1 + \varepsilon) \ln n$, so that (3.4) implies

$$\left\{Z_{\geq d} \geq \frac{1 + 2(1 + \varepsilon) \ln n}{\delta},\, D \in (1 \pm \varepsilon) \ln n\right\} \subset \left\{\left|Z^{(\varepsilon)}_{\geq d} - Z_{\geq d}\right| \leq \delta(1 + Z_{\geq d})\right\};$$

note that this holds for all $n \in \mathbb{N}$. Together with (3.5), this implies (3.2) for $n \geq n_0$, as desired. □

Proposition 3.2. *Let $\varepsilon \in (0, 1)$, $\ell \in \mathbb{N}$. For $0 \leq d \leq (1+\varepsilon)\ln n$, there is a $\beta = \beta(\varepsilon) > 0$ such that, under the coupling described above,* $|\mathbb{E}[(\ln W_d)^\ell]| \leq O((\ln n)^\ell n^{-\beta}) + 1$.

Proof. We first simplify to $|\mathbb{E}[(\ln W_d)^\ell]| \leq \mathbb{E}[|(\ln W_d)^\ell|] \leq \mathbb{E}[|\ln W_d|^\ell]$. Note that $1/n \leq W_d \leq n$ for every $0 \leq d \leq (1 \pm \varepsilon)\ln n$. Indeed, it is straightforward that $W_0 \equiv 1$ and, for $d \geq 1$, the bounds follow since any tree has at least one vertex of degree zero and so $\max\left\{Z_{\geq d}, Z^{(\varepsilon)}_{\geq d}\right\} < n$. Then, for any $\delta \in (0, 1)$,

$$\mathbb{E}[|\ln W_d|^\ell \mathbf{1}_{[W_d<1-\delta]}] \leq (\ln n)^\ell \mathbb{P}(W_d < 1-\delta),$$
$$\mathbb{E}[|\ln W_d|^\ell \mathbf{1}_{[W_d>1+\delta]}] \leq (\ln n)^\ell \mathbb{P}(W_d > 1+\delta).$$

Proposition 3.1 implies that these two terms are $O((\ln n)^\ell n^{-\beta})$, where the implicit constant depends on the choice of δ. With foresight, fix δ to satisfy $(\delta/(1-\delta))^\ell + \delta^\ell = 1$. Using that $1-(1/x) \leq \ln x \leq x-1$ for all $x \geq 0$,

$$\begin{aligned}\mathbb{E}[|\ln W_d|^\ell \mathbf{1}_{[W_d\in(1\pm\delta)]}] &\leq \mathbb{E}\left[\left|1-\frac{1}{W_d}\right|^\ell \mathbf{1}_{[W_d\in(1-\delta,1)]}\right] + \mathbb{E}[|W_d-1|^\ell \mathbf{1}_{[W_d\in[1,1+\delta)]}] \\ &\leq \left(\frac{\delta}{1-\delta}\right)^\ell + \delta^\ell = 1,\end{aligned}$$

and so the result follows. □

3.1. Proof of Theorem 1.2

Proof of Theorem 1.2. Fix $k \in \mathbb{N}$ and recall that $\gamma := 1 - \ln(2)$. Suppose that, for any $\varepsilon \in \left(0, \frac{1}{3}\right)$, there exist $c = c(\varepsilon) > 0$ and $C = C(k, \varepsilon) > 0$ such that

$$((\gamma-\varepsilon)\ln n)^k(1+O(n^{-c})) - C \leq \mathbb{E}[(\ln Z_{\geq D})^k] \leq ((\gamma+\varepsilon)\ln n)^k + C. \tag{3.6}$$

It is straightforward to verify that (3.6) establishes Theorem 1.2. So, it remains to prove (3.6).

Let $\varepsilon' = \varepsilon/(2\ln(2))$ and write $m_- = \lfloor(1-\varepsilon')\ln n\rfloor$, $m_+ = \lceil(1+\varepsilon')\ln n + 1\rceil$. We focus on the term $\mathbb{E}[(\ln Z_{\geq D})^k \mathbf{1}_{\{D\in(1\pm\varepsilon')\ln n\}}]$, as (2.5) and $1 \leq Z_{\geq D} \leq n$ imply

$$0 \leq \mathbb{E}[(\ln Z_{\geq D})^k \mathbf{1}_{\{D\notin(1\pm\varepsilon')\ln n\}}] \leq (\ln n)^k \mathbb{P}(D \notin (1\pm\varepsilon')\ln n) = (\ln n)^k O(n^{-\varepsilon'/12}). \tag{3.7}$$

Using the monotonicity of $Z_{\geq d}$, we have

$$\mathbb{E}[(\ln Z_{\geq D})^k \mathbf{1}_{\{D\in(1\pm\varepsilon')\ln n\}}] \leq \mathbb{E}[(\ln Z_{\geq m_-})^k \mathbf{1}_{\{D\in(1\pm\varepsilon')\ln n\}}] \leq \mathbb{E}[(\ln Z_{\geq m_-})^k],$$

which, by (2.3), yields

$$\mathbb{E}[(\ln Z_{\geq D})^k \mathbf{1}_{\{D\in(1\pm\varepsilon')\ln n\}}] \leq ((\gamma+\varepsilon)\ln n)^k + C_k. \tag{3.8}$$

For the lower bound, recall the variable $Z^{(\varepsilon')}_{\geq d}$ defined in the previous section with $d = m_+$. Observe that, if $D \in (1\pm\varepsilon')\ln n$, then $Z_{\geq D} \geq 1 + Z_{m_+}$, and we thus obtain

$$\begin{aligned}\mathbb{E}[(\ln Z_{\geq D})^k \mathbf{1}_{\{D\in(1\pm\varepsilon')\ln n\}}] &\geq \mathbb{E}[(\ln(1+Z_{\geq m_+}))^k \mathbf{1}_{\{D\in(1\pm\varepsilon')\ln n\}}] \\ &\geq \mathbb{E}\left[\left(\ln\left(1+Z^{(\varepsilon')}_{\geq m_+}\right)\right)^k\right]\mathbb{P}(D\in(1\pm\varepsilon')\ln n).\end{aligned} \tag{3.9}$$

https://doi.org/10.1017/jpr.2024.40 Published online by Cambridge University Press

The definition of W_d gives

$$\ln\left(1+Z^{(\varepsilon')}_{\geq m_+}\right)=\ln\left(\left(1+Z^{(\varepsilon')}_{\geq m_+}\right)\frac{1+Z_{\geq m_+}}{1+Z_{\geq m_+}}\right)=\ln(1+Z_{\geq m_+})+\ln W_{m_+}; \tag{3.10}$$

similarly, for $k\geq 2$, the binomial expansion implies

$$\left(\ln\left(1+Z^{(\varepsilon')}_{\geq m_+}\right)\right)^k=(\ln(1+Z_{\geq m_+}))^k+\sum_{\ell=1}^{k}\binom{k}{\ell}\ln(1+Z_{\geq m_+})^{k-\ell}(\ln W_{m_+})^{\ell}. \tag{3.11}$$

We use (2.4) for a lower bound on the expectation of the main term in these last two decompositions. For the error terms involving W_{m_+} we use Proposition 3.2. If $k=1$, we directly get

$$\begin{aligned}\mathbb{E}\left[\ln\left(1+Z^{(\varepsilon')}_{m_+}\right)\right]&=\mathbb{E}[\ln(1+Z_{\geq m_+})]+\mathbb{E}[\ln W_{m_+}]\\&\geq(\gamma-\varepsilon)\ln n(1+O(n^{-\alpha'}))+O((\ln n)n^{-\beta})-1.\end{aligned}$$

If $k\geq 2$, we control each of the terms in the sum of (3.11). For $1\leq\ell\leq k$, the Cauchy–Schwarz inequality gives

$$|\mathbb{E}[(\ln(1+Z_{\geq m_+}))^{k-\ell}(\ln W_{m_+})^{\ell}]|\leq\mathbb{E}[(\ln(1+Z_{\geq m_+}))^{2(k-\ell)}]^{1/2}\mathbb{E}[(\ln W_{m_+})^{2\ell}]^{1/2}.$$

The deterministic bound $Z_{\geq m_+}<n$ implies $\mathbb{E}[(\ln(1+Z_{\geq m_-}))^{2(k-\ell)}]^{1/2}\leq(\ln n)^{k-\ell}$. On the other hand, Proposition 3.2 yields

$$\mathbb{E}[(\ln W_{m_+})^{2\ell}]^{1/2}\leq(O((\ln n)^{2\ell}n^{-\beta})+1)^{1/2}\leq O((\ln n)^{\ell}n^{-\beta})+1.$$

Thus, after taking expectations in (3.11), we get

$$\mathbb{E}\left[\left(\ln\left(1+Z^{(\varepsilon')}_{\geq m_+}\right)\right)^k\right]\geq((\gamma-\varepsilon)\ln n)^k(1-o(n^{-c}))-C, \tag{3.12}$$

where $c=\min\{\alpha',\beta\}$ and $C=\max\{C_k,2^k\}$. Then, (3.7), (3.8), and (3.12) together with (2.5) imply (3.6), completing the proof. □

4. Simulations

The approach of counting the number of records, although difficult in directly studying K_n, provides us with a way to test our results via simulations in Python. We implement an algorithm that identifies the number of (largest) vertex-degree records; in the case of ties, we keep track of a set of uniform random variables that determines the order of deletion of vertices with the same degree. See Appendix A.

We obtain 10 000 samples of RRTs of size 100 000 and use the data to graph Figures 1 and 2. Figure 1 shows the empirical distribution of both K_n and $Z_{\geq D}$. The distributions appear to follow a power law. However, due to the nature of the variables under study, this computational probe is far from providing a better insight. We compare the empirical distributions of K_n and $Z_{\geq D}$ using a Q–Q plot, shown in Figure 2. From this plot, it seems that $Z_{\geq D}$ and K_n share similar behavior.

To further explore the growth behavior of $\mathbb{E}[K_n]$ and $\mathbb{E}[Z_{\geq D}]$, we generate 10 000 samples of RRTs of 50 different tree sizes, equally spaced on a logarithmic scale. Figure 3 shows the

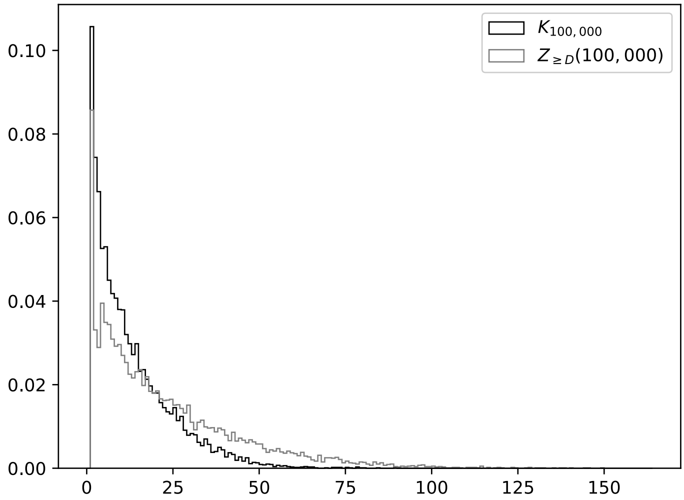

FIGURE 1. Empirical distributions of $K_{100\,000}$ and $Z_{\geq D}(100\,000)$ using a sample of size 10 000.

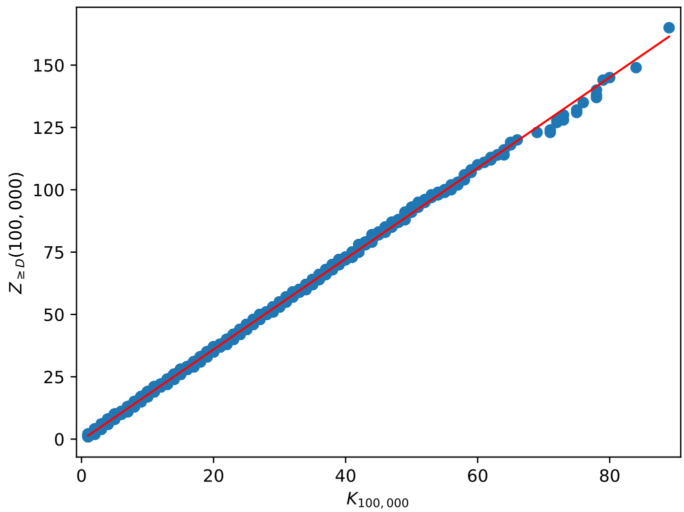

FIGURE 2. Q–Q plot for $Z_{\geq D}(100\,000)$ and $K_{100\,000}$ using a sample size of 10 000.

mean for each size, together with n^{γ} as a comparison. A shortcoming we found was that since $\ln n$ varies slowly, it was difficult to obtain a deeper insight into this resemblance.

Summing up, the simulations provide limited yet useful information. The Q–Q plot in Figure 2 contributes to our understanding of the process since it may suggest that K_n and $Z_{\geq D}$ have the same growth order.

https://doi.org/10.1017/jpr.2024.40 Published online by Cambridge University Press

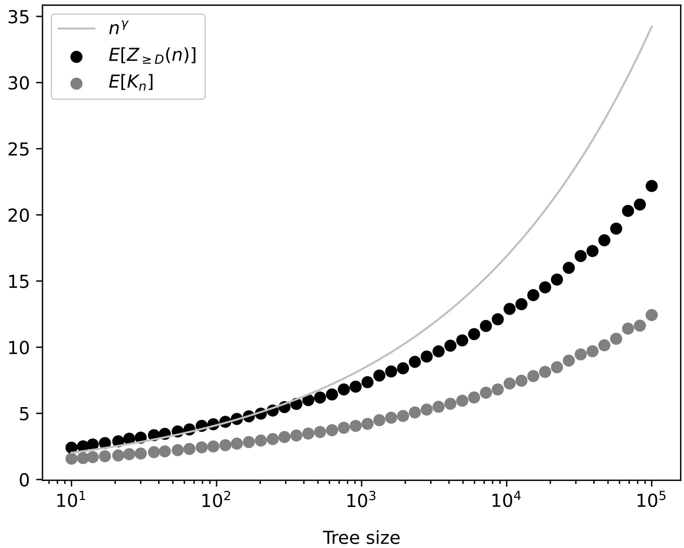

FIGURE 3. Estimates for $\mathbb{E}[Z_D]$ and $\mathbb{E}[K_n]$ in logarithmic scale, given a sample size of 10 000.

5. Conclusion and open problems

In this paper we introduced the notion of targeted cutting of a random recursive tree. This procedure serves as a proxy for the resilience of a particular network to malicious attacks from an enemy. To a certain extent, this problem has been tackled for scale-free random networks, where the malicious strategy boils down to removing a given proportion of the smallest-labeled vertices, regardless of the structure of the graph. The reason for this is mainly because preferential attachment models have the property of *persistence*, which refers to the fact that high-degree vertices are fixed early in the process. In contrast, RRTs lack persistence and the picture for high-degree vertices in RRTs is more nuanced and far less complete.

Our results provide the first quantitative estimate for the deletion time K_n of the targeted cutting process for random recursive trees. To do so, we shifted the focus to $Z_{\geq D}$, the number of vertices with degree at least as large as the degree of the root; this random variable corresponds to the quenched worst-case scenario for K_n.

In our main theorems, we found that $\ln Z_{\geq D}$ grows as $\ln n$ asymptotically and obtains its limiting behavior in probability. Moreover, we obtained the growth order of the kth moment of $\ln Z_{\geq D}$ for $k \geq 1$.

Proposition 1.1, confirms the intuition that the targeted procedure requires substantially fewer cuts than the random edge deletion procedure. This result was obtained by observing that, once the tree has been sampled and all the degrees are known, $Z_{\geq D}$ is an upper bound for K_n. Unfortunately, the sensitivity that $Z_{\geq D}$ may provide for the targeted cutting procedure is limited since it fails to consider the structure of T_n, in particular high-degree vertices.

It remains open to understand the structure of the tree induced by, say, the first $k = k(n)$ vertices with highest degree. In turn, this could shed light on a theoretical analysis of vertex-degree records. Similarly, we may study the sizes of the subtrees of maximal degree vertices, in order to exploit the splitting property for cutting processes.

https://doi.org/10.1017/jpr.2024.40 Published online by Cambridge University Press

On the other hand, analyzing $Z_{\geq D}$ in more detail is not a straightforward task. It would be interesting to sharpen the results on $Z_{\geq c \ln n}$ for a wider range of $c \geq 0$ for RRTs and, more generally, weighted RRTs.

Appendix A. The targeted cutting time algorithm

We use a depth-first search algorithm to simulate K_n, as this variable corresponds to the number of (largest) records of vertex degrees in paths from the root to the leaves, which is denoted by C below.

Given $n \in \mathbb{N}$, generate an RRT T_n where, initially, each vertex v stores its degree $d_{T_n}(v)$ and its parent P_v, and set $R_v = \deg_{T_n}(v)$, $L_v = v$. Let C be a counting variable initialized to 1 (because the root itself is a record). Using a depth-first search exploration, each vertex v compares its degree, $d_{T_n}(v)$, with the current record of its parent, R_{P_v}. Additionally, L_v keeps track of the largest degree vertex encountered in its path to the root.

We encounter three cases:

1. If $d_{T_n}(v) < R_{P_v}$, set $R_v = R_{P_v}$ and $L_v = L_{P_v}$. In the cutting process, v is removed because one of its ancestors is deleted.
2. If $d_{T_n}(v) > R_{P_v}$, update $C = C + 1$. That is, v is removed before any of its ancestors and so v counts as a record.
3. If $d_{T_n}(v) = R_{P_v}$, then let $U(v)$ and $U(L_{P_v})$ be independent Uniform(0, 1) random variables (note that $U(L_{P_v})$ may already have been generated). If $U(v) < U(L_{P_v})$, proceed as in step 1, otherwise proceed as in step 2.

Return the variable C.

Funding information

Sergio I. López was supported by DGAPA-PAPIIT-UNAM grant IN-102822, and Marco L. Ortiz was supported by CONACYT grant 955991.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ADDARIO-BERRY, L., BROUTIN, N. AND HOLMGREN, C. (2014). Cutting down trees with a Markov chainsaw. *Ann. Appl. Prob.* **24**, 2297–2339.

[2] ADDARIO-BERRY, L. AND ESLAVA, L.(2017). High degrees in random recursive trees. *Random Structures Algorithms* **52**, 560–575.

[3] ALENAZI, M. AND STERBENZ, J. (2015). Comprehensive comparison and accuracy of graph metrics in predicting network resilience. In *Proc. 11th Int. Conf. Design of Reliable Communication Networks*, pp. 157–164.

[4] BANERJEE, S. AND BHAMIDI, S. (2021). Persistence of hubs in growing random networks. *Prob. Theory Relat. Fields* **180**, 891–953.

[5] BAUR, E. (2016). Percolation on random recursive trees. *Random Structures Algorithms* **48**, 655–680.

[6] BAUR, E. AND BERTOIN, J. (2014). Cutting edges at random in large recursive trees. In *Stochastic Analysis and Applications 2014* (Springer Proc. Math. Statist. 100), eds D. Crisan, B. Hambly and T. Zariphopolou. Springer, New York, pp. 51–76.

[7] BERCHE, B., VON FERBER, C., HOLOVATCH, T. AND HOLOVATCH, Y. (2009). Resilience of public transport networks against attacks. *Eur. Phys. J. B* **71**, 125–137.
[8] BERTOIN, J. (2014). Sizes of the largest clusters for supercritical percolation on random recursive trees. *Random Structures Algorithms* **44** 29–44.
[9] BERTOIN, J. (2015). The cut-tree of large recursive trees. *Ann. Inst. Henri Poincaré Prob. Statist.* **51**, 478–488.
[10] BERTOIN, J. AND MIERMONT, G. (2013). The cut-tree of large Galton–Watson trees and the Brownian CRT. *Ann. Appl. Prob.* **23**, 1469–1493.
[11] BERZUNZA, G., CAI, X. S. AND HOLMGREN, C. (2020). The *k*-cut model in conditioned Galton–Watson trees. *Leibniz Int. Proc. Informatics* **159**, 5:1–5:10.
[12] BOLLOBÁS, B. AND RIORDAN, O. (2004). Robustness and vulnerability of scale-free random graphs. *Internet Math.* **1**, 1–35.
[13] BOUCHERON, S., LUGOSI, G. AND MASSART, P. (2013). *Concentration Inequalities*. Oxford University Press.
[14] CAI, X. S., DEVROYE, L., HOLMGREN, C. AND SKERMAN, F. (2019). *k*-cuts on a path. In *Algorithms and Complexity* (Lect. Notes Comp. Sci. **11485**), ed. P. Heggernes. Springer, Cham.
[15] CAI, X. S., DEVROYE, L., HOLMGREN, C. AND SKERMAN, F. (2019). *k*-cut on paths and some trees. *Electron. J. Prob.* **25**, 22.
[16] CAI, X. S. AND HOLMGREN, C. (2019). Cutting resilient networks – Complete binary trees. *Electron. J. Combinatorics* **26**, P4.43.
[17] COHEN, R., EREZ, K., ben Avraham, D. and Havlin, S. (2000). Resilience of the internet to random breakdowns. *Phys. Rev. Lett.* **85**, 4626–4628.
[18] COHEN, R., EREZ, K., ben Avraham, D. and Havlin, S. (2001). Breakdown of the internet under intentional attack. *Phys. Rev. Lett.* **86**, 3682–3685.
[19] DEVROYE, L. (1988). Applications of the theory of records in the study of random trees. *Acta Informatica* **26**, 123–130.
[20] DEVROYE, L. AND LU, J. (1995). The strong convergence of maximal degrees in uniform random recursive trees and dags. *Random Structures Algorithms* **7**, 1–14.
[21] DIEULEVEUT, D. (2015). The vertex-cut-tree of Galton–Watson trees converging to a stable tree. *Ann. Appl. Prob.* **25**, 2215–2262.
[22] DRMOTA, M. (2009). *Random Trees: An Interplay between Combinatorics and Probability*. Springer, Vienna.
[23] DRMOTA, M., IKSANOV, A., MÖHLE, M. AND ROESLER, U. (2009). A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree. *Random Structures Algorithms* **34**, 319–336.
[24] ESLAVA, L. (2020). A non-increasing tree growth process for recursive trees and applications. *Combinatorics Prob. Comput.* **30**, 79–104.
[25] ESLAVA, L. (2022). Depth of vertices with high degree in random recursive trees. *ALEA, Lat. Am. J. Prob. Math. Statist.* **19**, 839–857.
[26] ESLAVA, L., LODEWIJKS, B. AND ORTGIESE, M. (2023). Fine asymptotics for the maximum degree in weighted recursive trees with bounded random weights. *Stoch. Process. Appl.* **158**, 505–569.
[27] GOLDSCHMIDT, C. AND MARTIN, J. (2005). Random recursive trees and the Bolthausen–Sznitman coalescent. *Electron. J. Prob.* **10**, 718–745.
[28] HOLMGREN, C. (2008). Random records and cuttings in split trees: Extended abstract. In *Proc. 5th Colloq. Math. Theoret. Comp. Sci.*, pp. 269–282.
[29] HOLMGREN, C. (2010). Random records and cuttings in binary search trees. *Combinatorics Prob. Comput.* **19**, 391–424.
[30] IKSANOV, A. AND MÖHLE, M. (2007). A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Electron. Commun. Prob.* **12**, 28–35.
[31] JANSON, S. (2004). Random records and cuttings in complete binary trees. In *Mathematics and Computer Science III*, eds M. Drmota *et al.* Birkhäuser, Basel, pp. 241–253.
[32] JANSON, S. (2005). Asymptotic degree distribution in random recursive trees. *Random Structures Algorithms* **26**, 69–83.
[33] JANSON, S. (2006). Random cutting and records in deterministic and random trees. *Random Structures Algorithms* **29**, 139–179.
[34] JANSON, S., ŁUCZAK, T. AND RUCINSKI, A. (2000). *Random Graphs*. John Wiley, Chichester.
[35] JAVANIAN, M AND VAHIDI-AS, ML. (2004). Isolating the most recent entry in a random recursive tree by random cuts. *J. Appl. Math. Comput.* **16**, 115–123.
[36] KUBA, M. AND PANHOLZER, A. (2008). Isolating a leaf in rooted trees via random cuttings. *Ann. Combinatorics* **12**, 81–99.
[37] KUBA, M. AND PANHOLZER, A. (2008). Isolating nodes in recursive trees. *Aequat. Math.* **76**, 258–280.
[38] KUBA, M. AND PANHOLZER, A. (2013). Multiple isolation of nodes in recursive trees. *Online J. Anal. Combinatorics*, **9**.

[39] LODEWIJKS, B. (2022). On joint properties of vertices with a given degree or label in the random recursive tree. *Electron. J. Prob.* **27**, 1–45.
[40] LODEWIJKS, B. (2023) The location of high-degree vertices in weighted recursive graphs with bounded random weights. Preprint, arXiv:2110.05438.
[41] LODEWIJKS, B. AND ORTGIESE, M. (2020) The maximal degree in random recursive graphs with random weights. Preprint, arXiv:2007.05438.
[42] MEIR, A. AND MOON, J. (1970). Cutting down random trees. *J. Austral. Math. Soc.* **11**, 313–324.
[43] MEIR, A. AND MOON, J. (1974). Cutting down recursive trees. *Math. Biosci.* **21**, 173–181.
[44] MÓRI, T. F. (2005). The maximum degree of the Barabási–Albert random tree. *Combinatorics Prob. Comput.* **14**, 339–348.
[45] Panholzer, A. (2004). Destruction of recursive trees. In *Mathematics and Computer Science III*, eds M. Drmota *et al.* Birkhäuser, Basel, pp. 267–280.
[46] WANG, M. (2021). k-cut model for the Brownian continuum random tree. *Electron. Commun. Prob.* **26**, 46.
[47] XU, F., SI, S., DUAN, D., LV, C. AND XIE, J. (2019). Dynamical resilience of networks against targeted attack. *Physica A* **528**, 121329.