Random Structures & Algorithms

WILEY

RESEARCH ARTICLE OPEN ACCESS

Uniform Temporal Trees

Caelan Atamanchuk[1] | Luc Devroye[2] | Gábor Lugosi[3,4,5]

[1]Department of Mathematics and Statistics, McGill University, Montreal, Canada | [2]School of Computer Science, McGill University, Montreal, Canada | [3]Department of Economics and Business, Pompeu Fabra University, Barcelona, Spain | [4]ICREA, Barcelona, Spain | [5]Barcelona School of Economics, Barcelona, Spain

Correspondence: Gábor Lugosi (gabor.lugosi@gmail.com)

Received: 22 January 2025 | **Revised:** 9 November 2025 | **Accepted:** 6 December 2025

Keywords: random temporal trees | random trees | temporal graphs

ABSTRACT

Motivated by the study of random temporal networks, we introduce a class of random trees that we coin *uniform temporal trees*. A uniform temporal tree is obtained by assigning independent uniform $[0,1]$ labels to the edges of a rooted complete infinite n-ary tree and keeping only those vertices for which the path from the root to the vertex has decreasing edge labels. The p-percolated uniform temporal tree, denoted by $\mathcal{T}_{n,p}$, is obtained similarly, with the additional constraint that the edge labels on each path are all below p. We study several properties of these trees, including their size, height, the typical depth of a vertex, and degree distribution. In particular, we establish a limit law for the size of $\mathcal{T}_{n,p}$ which states that $\frac{|\mathcal{T}_{n,p}|}{e^{np}}$ converges in distribution to an Exponential(1) random variable as $n \to \infty$. For the height $H_{n,p}$, we prove that $\frac{H_{n,p}}{np}$ converges to e in probability. Uniform temporal trees show some remarkable similarities to uniform random recursive trees.

1 | Introduction

In network science, the graph modeling the network is often equipped with edge labels representing time stamps. For example, in a network describing human interactions, the network's vertices represent individuals, edges stand for encounters, and the edges may be labeled by the time the encounter happens. Such *temporal* networks allow one to study the spread of an infection or information (see Holme and Saramäki (2012), Holme and Saramäki (2013), Holme and Saramäki (2019), Hosseinzadeh et al. (2022), Sanjay Kumar and Panda (2020)).

A simple mathematical model for temporal networks that has been gaining attention is *random temporal graphs*. In this model, the time stamps are obtained by assigning a uniform random permutation to the edge set. If one is only interested in the ordering of the edge labels, equivalently, every edge of a graph is assigned an independent random label, uniformly distributed in $[0,1]$. In particular, the *random simple temporal graph* model is obtained by adding such labels to the edges of an Erdős-Rényi random graph $G_{n,p}$. Random simple temporal graphs exhibit some remarkable phase transitions (see,

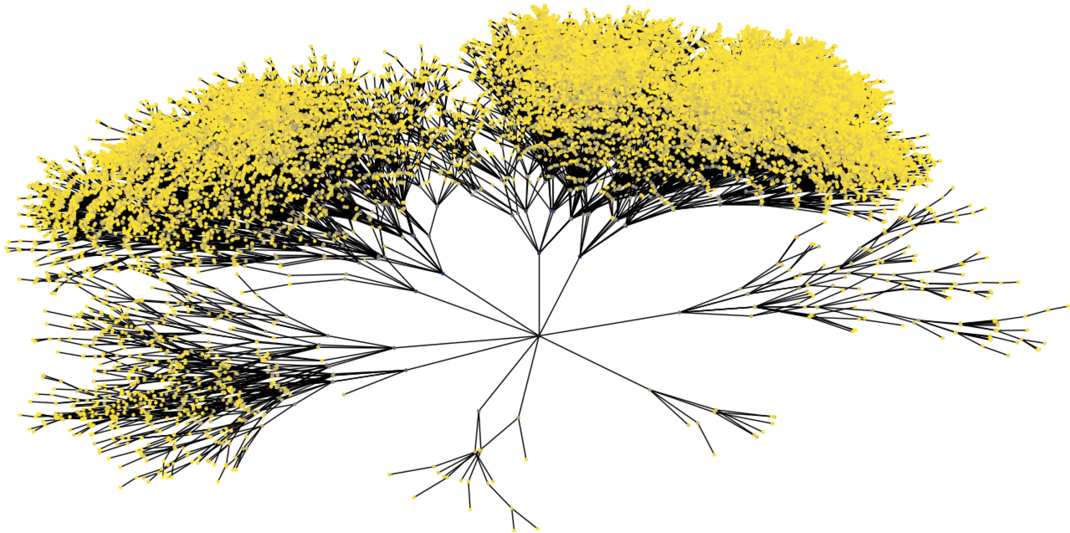

FIGURE 1 | A uniform temporal tree with $n = 10$.

e.g., Angel et al. (2020); Mertzios et al. (2024); Becker et al. (2022); Broutin et al. (2023); Casteigts et al. (2024); Atamanchuk et al. (2024)).

Since (sparse) Erdős-Rényi random graphs are locally tree-like, it is natural to study analogous random trees. This motivates our definition of a *uniform temporal tree*, specified below. The paper's main goal is to study the basic properties of such random trees, including their size, height, and degree distribution.

The definition is the following. For a positive integer n, let T_n be a rooted infinite complete n-ary tree (i.e., the root vertex has degree n and every other vertex has degree $n+1$). To each edge e of T_n, assign an independent random variable U_e, uniformly distributed in $[0,1]$. U_e is called the *label* of the edge e. A path between the root and a vertex v is called *decreasing* if the edge labels in the path appear in decreasing order. Sometimes it is more convenient to assign labels to vertices. The label ℓ_v of a vertex $v \in T_n$ is the label U_e of the edge e connecting v to its parent (the parent of a vertex v is the vertex adjacent to v that is on the path between the root and v). In some cases, we focus on the vertex labels, calling a path decreasing if the vertex labelling is decreasing, though this is an equivalent definition. Note that the root vertex does not have a parent or a label. The (p-percolated) uniform temporal tree, $\mathcal{T}_{n,p}$, is a random tree obtained from T_n by assigning the root the label p and deleting all vertices whose path from the root is not decreasing with respect to the vertex labelling. Note that every vertex in a $\mathcal{T}_{n,p}$ has a label at most p. When $p = 1$, we simplify the notation and just write $\mathcal{T}_n$ for $\mathcal{T}_{n,1}$ (Figure 1).

It is clear that $\mathcal{T}_n$ (and therefore $\mathcal{T}_{n,p}$) is almost surely finite. Indeed, there are n^k vertices in the k-th generation of T_n and each vertex v in the k-th generation exists in $\mathcal{T}_n$ with probability $1/k!$ as the edges on the path from the root to v must be in decreasing order. Thus the probability that $\mathcal{T}_n$ has a vertex at depth k is at most $n^k/k!$ which goes to zero as $k \to \infty$.

This is not the first work to consider increasing paths on trees. A model that is closely related to temporal graphs, called *accessibility percolation*, has been studied in recent years (Berestycki et al. (2014); Martinsson (2015); Berestycki et al. (2013)). Accessibility percolation on n-ary trees has been studied before, and its study is closely related to the heights of uniform temporal trees (Nowak and Krug (2013); Roberts and Zhao (2013)).

In the next section, we present the main results of the paper concerning the distribution of the size of $\mathcal{T}_{n,p}$, the typical depth of a vertex, the height (i.e., the depth of the deepest vertex), and the degree distribution. The proofs of these results are given in subsequent sections.

2 | Results

Before presenting the main results, we fix some terminology and notation. Let T be a rooted tree. $P(v)$ for any $v \in T$ (except for the root) denotes the unique path between the root and v. The set $P^-(v)$ is $P(v)$ with the root removed. The *depth* of v, $|v|$, is the number of edges in $P(v)$. The *parent* of v, $p(v)$ is the single neighbor of v in $P(v)$, and the set of *children* of v, $C(v)$, contains all vertices at depth $|v|+1$ that are adjacent to v. The *out-degree* of a vertex v is the number $|C(v)|$ of its children. Two vertices u and v are *siblings* if $p(u) = p(v)$. For a vertex $v \in T$, $T(v)$ is the subtree of T rooted at v containing all *descendants* of v in T, that is, the tree containing v, its children, grandchildren, and so on. Convergence in distribution for a sequence of random variables is denoted by $\xrightarrow{\mathcal{L}}$, and $\overset{\mathcal{L}}{=}$ is used for equality in distribution. Finally, we let $\xrightarrow{\mathbb{P}}$ represent convergence in probability for a sequence of random variables.

Many features of uniform temporal trees are quite similar to those of uniform random recursive trees, so we define this model before presenting our results. The uniform random recursive tree on n vertices is a random rooted tree with vertices labeled in $\{1, \dots, n\}$. The root has label 1. Vertices $i \in \{2, \dots, n\}$ are attached recursively such that vertex i is attached to a vertex in $\{1, \dots, i-1\}$ selected uniformly at random. The uniform random recursive tree is one of the most ubiquitous trees in computer science and has been thoroughly studied (see Meir and Moon (1978); Devroye (1988); Pittel (1994); Janson (2005); Drmota (2009); Addario-Berry and Eslava (2018)).

The first result concerns the size of $\mathcal{T}_{n,p}$. As is shown as the start of Section 4, a quick computation yields that the expected size of $\mathcal{T}_{n,p}$ is e^{np}. However, the size does not concentrate around the mean. We show that it admits a limit law: the size divided by its expectation converges, in distribution, to an exponential random variable. Moreover, we establish a joint limit law for the "distribution of mass" at the root, that is, for the sizes of the subtrees of children of the root with the largest labels.

Theorem 2.1. *Let $p \in (0,1]$ and consider a percolated uniform temporal tree $\mathcal{T}_{n,p}$. Then*

$$\mathbf{E}\,|\mathcal{T}_{n,p}| = e^{np}$$

and

$$\frac{|\mathcal{T}_{n,p}|}{e^{np}} \xrightarrow{>\mathcal{L}} E \quad \text{as} \quad n \to \infty,$$

where E is an exponential(1) random variable.

Moreover, for $1 \le i \le n$, let v_i be the child of the root with the i-th largest label. Then for any fixed $m \ge 1$,

$$\left(\frac{|\mathcal{T}_{n,p}(v_1)|}{e^{np}}, \dots, \frac{|\mathcal{T}_{n,p}(v_m)|}{e^{np}} \right) \xrightarrow{>\mathcal{L}} \left(E_1 U_1, E_2 U_1 U_2, \dots, E_m U_1 \cdots U_m\right) \quad \text{as} \quad n \to \infty,$$

where $(E_k)_{k \ge 0}$ is a sequence of independent Exponential(1) *random variables and $(U_k)_{k \ge 0}$ is an independent sequence of independent uniform random variables on* $[0,1]$.

Remark. It follows from Theorem 2.1 that if $U_1, U_2, \dots$ are independent uniform $[0,1]$ and $E_1, E_2, \dots$ are independent exponential(1) random variables, then

$$\sum_{i=1}^{\infty} E_i \prod_{j=1}^{i} U_j$$

is an exponential(1) random variable. This identity may also be checked directly.

Theorem 2.1 is the technically most involved result of the paper. Its proof is given in Section 4. The key idea behind the proof is a representation of the labels in T_n as a sum of uniform spacings; this representation is given in Section 3. Using this representation, we identify a connection between the evolution of the labels in T_n and the evolution of values in a branching random walk on an infinite binary tree, which we briefly describe now.

Let $\mathcal{T}^*$ be a rooted infinite complete binary tree (the root has one child, and every other vertex has precisely two children). We consider the root to be in generation -1. Let $(X_v : v \in \mathcal{T}^*)$ be a branching random walk on $\mathcal{T}^*$ with step size Exponential(1)$/n$ (see e.g., Shi (2015) for a monograph on branching random walks). It is worth pointing out that this step size is an asymptotic approximation for one uniform spacing. To each vertex, associate a tree $\mathcal{T}_v$ which is distributed like $\mathcal{T}_{n,p-X_v}$. Assume that the collection $\left(\mathcal{T}_v : |v| = L\right)$ is conditionally independent given the vector $\left(X_v : |v| = L\right)$ for any $L \geq -1$. The proof in Section 4 is mostly focused on arguing, for sufficiently large L, that

$$\left| \frac{|\mathcal{T}_{n,p}|}{e^{np}} - \frac{\sum_{|v|=L} |\mathcal{T}_v|}{e^{np}} \right|$$

is small as $n \to \infty$. This approximation is explored in Lemma 4.2. The labels of the roots in $\left(\mathcal{T}_v : |v| = L\right)$ are exactly described by a branching walk, and so we can complete precise computations concerning them. The final limit law comes from a martingale analysis of $\mathbf{E}\left[\frac{\sum_{|v|=L} |\mathcal{T}_v|}{e^{np}} \;|\left(X_v : |v| = L\right)\right]$. This analysis is summarized in Lemma 4.4.

The next result concerns the *height* $H_{n,p}$ of a uniform temporal tree, that is, the maximum vertex depth in $\mathcal{T}_{n,p}$. The following theorem, proved in Section 5, states that the height is about e times the logarithm of the size of the tree. This property is reminiscent of uniform random recursive trees (see Devroye (1987) and Pittel (1994)). The result can be seen as a generalization of results for accessibility percolation for n-ary trees found in (Roberts and Zhao (2013); Nowak and Krug (2013)). It is worth noting that the techniques used in our proof differ from those used to derive similar results in previous works.

Theorem 2.2. *Fix $p \in (0, 1]$, and let $H_{n,p}$ denote the height of a percolated uniform temporal tree $\mathcal{T}_{n,p}$. Then*

$$\frac{H_{n,p}}{np} \xrightarrow{\mathbb{P}} e \quad \textit{and} \quad \frac{H_{n,p}}{\log |\mathcal{T}_{n,p}|} \xrightarrow{\mathbb{P}} e \text{ as } n \to \infty.$$

Note that the second statement follows from the first since Theorem 2.1 implies that $\left(\log |\mathcal{T}_{n,p}|\right)/(np) \to 1$ in probability. The proof of this theorem is based on a connection between uniform temporal trees and branching random walks that is similar to the one used for Theorem 2.1.

The next property establishes the typical depth of a vertex in a uniform temporal tree. Just like in a uniform random recursive tree, the depth is concentrated around the natural logarithm of the size of the tree (see Devroye (1998) for the corresponding results on depths in other recursive trees). The proof of the next theorem is provided in Section 6.

Theorem 2.3. *Let $p \in (0, 1]$, and let $D_{n,p}$ denote the depth of a uniformly chosen vertex in a percolated uniform temporal tree $\mathcal{T}_{n,p}$. Then*

$$\frac{D_{n,p}}{np} \xrightarrow{\mathbb{P}} 1 \quad \textit{and} \quad \frac{D_{n,p}}{\log |\mathcal{T}_{n,p}|} \xrightarrow{\mathbb{P}} 1 \quad \textit{as } n \to \infty.$$

Finally, the next theorem establishes the asymptotic expected degree distribution of a uniform temporal tree. In particular, the expected number of leaves is about half of the expected number of vertices, the expected number of vertices with one child is a quarter of the expected size, etc. Once again, this property is similar to the corresponding asymptotic degree distribution of a uniform random recursive tree. The following theorem is proved in Section 7.

Theorem 2.4. *Let $p \in (0, 1]$, and, for $k \geq 0$, let $L_{n,k}$ denote the number of vertices of out-degree k in a percolated uniform temporal tree $\mathcal{T}_{n,p}$. Then*

$$\frac{\mathbf{E} L_{n,k}}{e^{np}} \to 2^{-(k+1)} \text{ as } n \to \infty.$$

As has been highlighted above, uniform temporal trees share several of their basic characteristics with random recursive trees. One may wonder whether a random recursive tree is equivalent to a uniform temporal tree $\mathcal{T}_n$ when conditioning on the size $|\mathcal{T}_n|$. However, while in a uniform temporal tree, the root vertex always has maximal degree, in a uniform random recursive tree, there are vertices with much higher degree Devroye and Lu (1995); Addario-Berry and Eslava (2018); Eslava (2022). Moreover, the distribution of mass at the root established in Theorem 2.1 is different from the "stick-breaking" distribution of the uniform random recursive tree.

Broutin et al. (2023) utilizes direct couplings between neighborhoods of vertices in sparse random simple temporal graphs and uniform random recursive trees to prove statements about connectivity. In sparse random graphs, neighborhoods around vertices are tree-like and so match the structure of the $\mathcal{T}_{n,p}$ closely.

3 | The Uniform Spacings Coupling

In the proof of Theorems 2.1 and 2.2 we use representations of the labels in a uniform temporal tree as a sum of uniform spacings. We explore the connection in this section.

Let $U_1, \dots, U_n$ be a collection of independent Uniform[0, 1] random variables and let $V_1 \geq \cdots \geq V_n$ be the corresponding order statistics. We also set $V_0 = 1$, $V_{n+1} = 0$. Writing $S_i = V_{i-1} - V_i$ ($i \in \{1, \dots, n+1\}$) for the induced spacings, we use the representation

$$\left(S_1, \dots, S_{n+1}\right) := \left(V_0 - V_1, \dots, V_n - V_{n+1}\right) \overset{\mathcal{L}}{=} \left(\frac{E_1}{\sum_{i=1}^{n+1} E_i}, \dots, \frac{E_{n+1}}{\sum_{i=1}^{n+1} E_i} \right),$$

where $E_1, \dots, E_{n+1}$ are independent Exponential(1) random variables (see, e.g., Devroye (1986)). We record a key observation about uniform random variables needed to incorporate the uniform spacings.

Lemma 3.1. *Let $U_1, \dots, U_n$ be a collection of independent* Uniform[0, 1] *random variables, let $x \in [0, 1]$, and let $I = \{1 \leq i \leq n : U_i \geq x\}$. Define a collection of random variables $\left(U_1^*, \dots, U_n^*\right)$, where*

$$U_i^* = \begin{cases} U_i - 1 & i \in I \\ U_i & \textit{otherwise} \end{cases}$$

Then, $\left(U_1^*, \dots, U_n^*\right)$ is distributed like a vector of independent uniform random variables on the interval $[x-1, x]$. In particular, if $V_1^*, \dots, V_n^*$ are the order statistics of $\left(U_1^*, \dots, U_n^*\right)$ and $\left(S_1, \dots, S_{n+1}\right)$ is a vector of uniform spacings, then

$$\left(V_1^*, \dots, V_n^*\right) \overset{\mathcal{L}}{=} \left(x - S_1, \dots, x - \left(S_1 + \cdots + S_n\right)\right).$$

Proof. The transformation described in the lemma is equivalent to moving the section of the interval [0, 1] above x (and the corresponding points in $\{U_1, \dots, U_n\}$ that lie above x) to be below zero. The points in $[0, x]$ and $[1-x, 0]$ are still uniformly distributed over these intervals, and thus together are uniform over $[1-x, x]$ (see Figure 2). □

The uniform spacings representation then yields an equivalent description of how labels evolve in uniform temporal trees. Let $\left(\left(S_{v,1}, \dots, S_{v,n+1}\right) : v \in T_n\right)$ be a collection of independent uniform spacings. We define labels recursively for all of the vertices in T_n. First, the root ρ is given the label $\ell_\rho = p$. Then, for a vertex $v \in T_n$ with label ℓ_v, we define the labels of its children $v_1, \dots, v_n$ as

$$\ell_{v_j} = \ell_v - \sum_{i=1}^{j} S_{v,i} \quad \text{for} \quad j \in \{1, \dots, n\}.$$

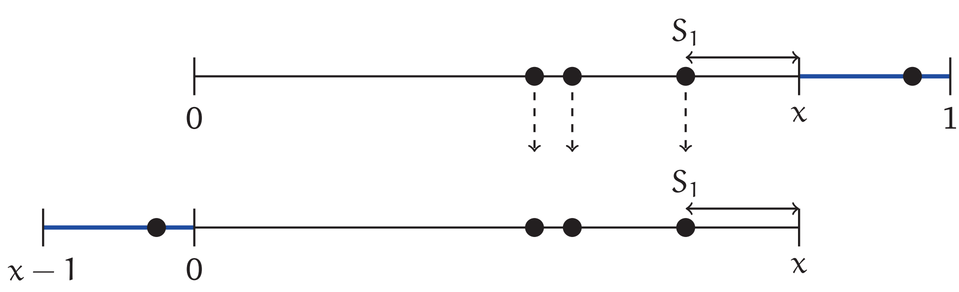

FIGURE 2 | The rotation of the uniform spacings around a vertex x. The blue section above x is moved from above x to below 0. After the segment is moved, the points are distributed uniformly over $[x-1, x]$.

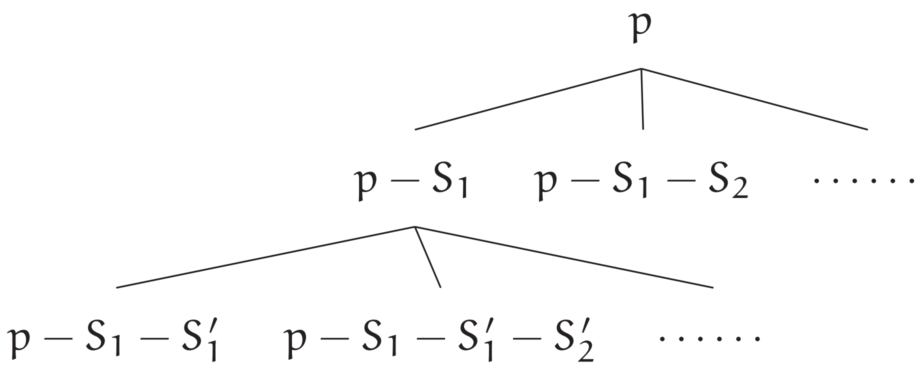

FIGURE 3 | The evolution of labels in a $\mathcal{T}_{n,p}$ according to the spacings coupling. The random variables S_1, S_2, S_1', S_2' are all uniform spacings. The label of a vertex is the label of its next lower-rank sibling (or parent if its rank is 1) minus a spacing.

Deleting the vertices with negative labels in T_n is equivalent to deleting the vertices that have labels above p and also generates $\mathcal{T}_{n,p}$. Moreover, applying the reverse of the transformation from Lemma 3.1 reveals that this labelling is equivalent to the original from the introduction. We refer to generating the labels of T_n in this manner as the *uniform spacings coupling*. The following notation is useful when discussing the evolution of labels under the uniform spacings coupling.

i. Define the *rank* of a vertex in a $\mathcal{T}_{n,p}$ as the placement of its label among its siblings. That is, the sibling with the largest label gets rank 1, the sibling with the second largest label gets 2, and so forth. We denote the rank of a vertex by $r(v)$, and by convention, the rank of the root is 0.
ii. The index of a vertex, which we denote by $\imath(v)$, is defined to be the sum of all the ranks of vertices in $P(v)$, $\imath(v) = \sum_{u \in P(v)} r(u)$.
iii. The set $I(j)$ for any $j \geq 0$ is the subset of all vertices in a $\mathcal{T}_{n,p}$ that have index j, $I(j) = \left\{ v \in \mathcal{T}_{n,p} : \imath(v) = j \right\}$. Note that if $v_1, \dots, v_n$ are the children of some vertex v in T_n listed in decreasing order of label, then $r\left(v_i\right) = i$.

Using this notation, we may describe the label of any vertex v that exists in a $\mathcal{T}_{n,p}$ in terms of ranks via the uniform spacings coupling:

$$\ell_v = p - \sum_{u \in P^-(v)} \sum_{i=1}^{r(u)} S_{p(u),i}.$$

See Figure 3 for an illustration. We use this representation in the proofs of Theorems 2.1 and 2.2.

4 | The Size

Throughout this section, p is always a fixed parameter in $(0, 1]$. First, we examine the first two moments of the size of a $\mathcal{T}_{n,p}$. There are n^k vertices at depth k in T_n, and each $v \in T_n$ with $|v| = k$ exists in $\mathcal{T}_{n,p}$ with probability $p^k/k!$ as the edges in $P(v)$ must all have label below p and be monotone decreasing. Hence, for any $n \geq 1$,

$$\mathbf{E} \mid \mathcal{T}_{n,p} \mid = \sum_{k=0}^{\infty} \frac{(np)^k}{k!} = e^{np},$$

establishing the first statement of Theorem 2.1. In the next lemma, we provide an upper bound for the second moment. The upper bound yields a concentration inequality for the average size of large collections of uniform temporal trees.

Lemma 4.1.

(i) *For all n, $\mathbf{E}\left|\mathcal{T}_{n,p}\right|^2 \leq 5\left(\mathbf{E}|\mathcal{T}_{n,p}|\right)^2$.*

(ii) *Let m be a fixed positive integer and let $q_1, \dots, q_m \in (0, p)$. For $i \in [m]$, define $p_i = p - q_i$. Let $\mathcal{T}_{n,p_i}$, $i \in [m]$ be independent uniform temporal trees and let $\mu = \sum_{i=1}^m \mathbf{E} \mid \mathcal{T}_{n,p_i} \mid = \sum_{i=1}^m e^{np_i}$. For all $\epsilon > 0$ and n,*

$$\mathbf{P}\left(\left|\sum_{i=1}^{m}|\mathcal{T}_{n,p_i}|-\mu\right|>\epsilon\mu\right)\leq\frac{5}{\epsilon^2}\left(\frac{\max_{1\leq i\leq m}e^{-nq_i}}{\sum_{i=1}^{m}e^{-nq_i}}\right).$$

The proof of the lemma is given in the appendix. As mentioned in the introduction, aside from the above concentration inequality, the other main ingredient in the proof of Theorem 2.1 is a connection between the labels of $\mathcal{T}_{n,p}$ and branching random walks. The representation in terms of uniform spacings described in Section 3 plays a key role in the construction.

First, let us briefly recall the definition of a branching random walk. Given a random variable X (called the *step size*) and a locally finite rooted tree T, label each edge in T with an independent copy of X, $\left(X_e\right)_{e\in T}$. A branching random walk on T is a collection of vertex-indexed random variables $\left(Y_v\right)_{v\in V}$, where Y_v is the sum of the labels of the edges in $P(v)$. The random variable Y_v is called the *value* of the vertex v in the branching random walk.

We define a branching random walk on a rooted infinite complete binary tree $\mathcal{T}^*$. The root has one child, and every other vertex has precisely two children. To match the standard notation for branching random walks, the root is in generation -1, its child in generation 0, and so forth. For any generation $L\geq 0$, there are 2^L vertices.

Lemma 4.2. *Let $L\geq 0$ be an integer, let $\delta\in(0,1)$, and let $\epsilon,x>0$. Let $p-p_1,\dots,p-p_{2^L}$ be the values of the vertices in the L-th generation of a branching random walk on the infinite complete binary tree $\mathcal{T}^*$ with step size $X\overset{\mathcal{L}}{=}\frac{E}{n}$, where E is an exponential(1) random variable. Define $p_i^+(\epsilon)$ and $p_i^-(\epsilon)$ such that $p-p_i^+(\epsilon)=(1+\epsilon)\left(p-p_i\right)$ and $p-p_i^-(\epsilon)=(1-\epsilon)\left(p-p_i\right)$ for all $1\leq i\leq 2^L$. Then,*

$$\mathbf{P}\left(\frac{1}{e^{np}}|\mathcal{T}_{n,p}|>x\right)\leq\mathbf{P}\left(\frac{1}{e^{np}}\left(\sum_{i=1}^{2^L}|\mathcal{T}_{n,p_i^-(\epsilon)}|+|\mathcal{T}'_{n,p_i^-(\epsilon)}|\right)>x(1-\delta)\right)+o_n(1)$$

and

$$\mathbf{P}\left(\frac{1}{e^{np}}|\mathcal{T}_{n,p}|>x\right)\geq\mathbf{P}\left(\frac{1}{e^{np}}\left(\sum_{i=1}^{2^L}|\mathcal{T}_{n,p_i^+(\epsilon)}|+|\mathcal{T}'_{n,p_i^+(\epsilon)}|\right)>x(1+\delta)\right)-o_n(1),$$

where the trees $\mathcal{T}_{n,p_i^\pm(\epsilon)},\mathcal{T}'_{n,p_i^\pm(\epsilon)}$ are uniform temporal trees that are all conditionally independent given $\left(p_1,\dots,p_{2^L}\right)$.

Proof of Lemma 4.2. Let $\mathcal{E}$ be the event that all vertices with $\imath(v)\leq L$ have children of index $(\imath(v)+1),\dots,L+1$. That is, $\mathcal{E}$ is the event that the out-degree of each vertex in $I(j)$ is at least $L-j+1$ for each $j\in\{0,\dots,L\}$. Since the degree of fixed-index vertices tends to infinity as $n\to\infty$ and L is a fixed integer, we know that $\mathbf{P}(\mathcal{E})\to 1$ as $n\to\infty$.

Let $v\in\mathcal{T}_{n,p}$ be a fixed vertex of rank k, and suppose that $p(v)$ has ℓ children listed in order of increasing rank $v_1,\dots,v_\ell$ where $v=v_k$. Recall the definition that $\mathcal{T}_{n,p}(v)$ is the subtree rooted at v in $\mathcal{T}_{n,p}$. We define $\mathcal{F}(v)$ to be the collection of all the subtrees rooted at higher rank siblings of v, that is,

$$\mathcal{F}(v)=\bigcup_{i=k+1}^{\ell}\mathcal{T}_{n,p}\left(v_i\right).$$

A simple observation is that every vertex in $\bigcup_{j\geq L+1}I(j)$ is in exactly one set $\mathcal{T}_{n,p}(v)$ or $\mathcal{F}(v)$ for some $v\in I(L+1)$, that is,

$$\bigcup_{j\geq L+1}I(j)=\bigcup_{v\in I(L+1)}\left(\mathcal{T}_{n,p}(v)\bigcup\mathcal{F}(v)\right).$$

Since all the sets on the right-hand side are disjoint and $\mid I(0)\bigcup\dots\bigcup I(L+1)\mid$ is finite, this implies that

$$\frac{1}{e^{np}}\mid\mathcal{T}_{n,p}\mid=\frac{1}{e^{np}}\sum_{v\in I(L+1)}\mid\mathcal{T}_{n,p}(v)\mid+\frac{1}{e^{np}}\sum_{v\in I(L+1)}\mid\mathcal{F}(v)\mid+o_n(1).\tag{1}$$

For any fixed $k\geq 0$, define $\mathcal{T}_{n,p}(k)$ to be the induced subtree of $\mathcal{T}_{n,p}$ on the vertex set $I(0)\bigcup\cdots\bigcup I(k)$. Let $\epsilon>0$. Since $\mathcal{T}_{n,p}(L+1)$ is a finite tree, we can apply the law of large numbers to the spacings from the uniform spacings coupling

$\left(S_{v,i} : v \in \mathcal{T}_{n,p}(L+1), i \in \{1, \dots, L+1\}\right)$ to assert that, for a sequence of independent Exponential(1) random variables $\left(E_{v,i}\right)_{v \in T_n, 1 \le i \le n+1}$,

$$\mathbf{P}\left(S_{\le}\right) := \mathbf{P}\left(\bigcap_{v \in \mathcal{T}_{n,p}(L+1)} \left\{\ell_v \le \ell_v^-(\epsilon)\right\}\right) \to 1,$$

and

$$\mathbf{P}\left(S_{\ge}\right) := \mathbf{P}\left(\bigcap_{v \in \mathcal{T}_{n,p}(L+1)} \left\{\ell_v \ge \ell_v^+(\epsilon)\right\}\right) \to 1,$$

as $n \to \infty$, where

$$\ell_v^-(\epsilon) := p - \frac{1-\epsilon}{n} \sum_{u \in P^-(v)} \sum_{i=1}^{r(u)} E_{p(u),i}, \quad \text{and} \quad \ell_v^+(\epsilon) := p - \frac{1+\epsilon}{n} \sum_{u \in P^-(v)} \sum_{i=1}^{r(u)} E_{p(u),i}.$$

Conditioning on $S_{\le}$ and $S_{\ge}$ allows us to remove many dependencies between vertex labels.

Next, on the event $\mathcal{E}$, we recursively define a one-to-one mapping ϕ of the vertices in a $\mathcal{T}_{n,p}(L+1)$ with the vertices of $\mathcal{T}^*(L)$ (the tree truncated at generation L) and assign a corresponding edge labeling $\left(X_e\right)_{e \in \mathcal{T}^*(L)}$. The following properties hold for our mapping and together provide our branching random walk description:

i. In $\mathcal{T}^*(L)$, the collection of random variables $\left(\sum_{e \in P(v)} X_e\right)_{v \in \mathcal{T}^*(L)}$ is a branching random walk with step size $\frac{1+\epsilon}{n} E$.
ii. For all $v \in \mathcal{T}_{n,p}(L+1)$, the value of $\phi(v)$ determines $\ell_v^+(\epsilon)$. That is, $\ell_v^+ = p - \sum_{e \in P(\phi(v))} X_e$.
iii. The jth generation of $\mathcal{T}^*(L)$ contains the vertices of index $j+1$ in $\mathcal{T}_{n,p}(L+1)$.

We only describe the construction for $\left(\ell_v^+(\epsilon) : v \in I(L+1)\right)$, though the same can be done for $\left(\ell_v^-(\epsilon) : v \in I(L+1)\right)$ using a similar procedure.

First, the root of $\mathcal{T}_{n,p}(L+1)$, which we denote with ρ, is mapped to the root of $\mathcal{T}^*(L)$. The unique index-1 vertex in $\mathcal{T}_{n,p}(L+1)$—call it c—, is mapped to the unique child of the root in $\mathcal{T}^*(L)$. The edge going into this vertex in $\mathcal{T}^*(L)$ is given the label $X_e = \frac{1+\epsilon}{n} E_{\rho,1}$. The three properties above hold for this base case.

Now, suppose that ϕ has been defined for $I(0), \dots, I(k)$, $k \le L$ and that the three properties hold for the partial assignment. Each vertex in $v \in I(k)$ has exactly one child c_v and one sibling s_v of index $k+1$. As $v \in I(k)$, its image $\phi(v)$ in $\mathcal{T}^*$ is already defined, and by assumption the left and right children of $\phi(v)$, $\phi(v)_\ell$ and $\phi(v)_r$, are not in the image $\phi(I(0) \bigcup \cdots \bigcup I(k))$. Define $\phi\left(c_v\right) = \phi(v)_\ell$ and $\phi\left(s_v\right) = \phi(v)_r$. We give the edges $e_1 = \left\{\phi(v), \phi(v)_\ell\right\}$ and $e_2 = \left\{\phi(v), \phi(v)_r\right\}$ the labels $\frac{1+\epsilon}{n} E_{v,1}$ and $\frac{1+\epsilon}{n} E_{p(v),r(v)+1}$ respectively. Point (iii) holds by definition for this extension of ϕ as all index-$(k+1)$ vertices are either a direct sibling or child of an index-k vertex previously placed into the tree. Moreover, by assuming that (ii) holds for the k-th step implies that it still holds for the $k+1$-th step. Since the exponential random variables $E_{v,1}$ and $E_{p(v),r(v)+1}$ have not been used in the construction previously (this is a consequence of assuming that property (ii) holds for the first k levels), property (i) holds for the extension as well (Figure 4).

This construction is almost sufficient to complete the proof, though there is still an approximation for $\mathcal{F}(v)$ with a tree that is needed. We delay proving this fact until the appendix, though we record and use the result. □

Lemma 4.3. *Let $\delta > 0$ and let $v \in \mathcal{T}_{n,p}$ be a vertex of fixed finite index. There are random trees $\mathcal{T}^+(v)$ and $\mathcal{T}^-(v)$ that are conditionally independent of the labels of v and its lower rank siblings in T_n given $\left(\ell_v^{\pm}(\epsilon) : v \in I(L+1)\right)$ (and in particular $T_n(v)$), such that $|\mathcal{T}^-(v)| \overset{\mathcal{L}}{=} |\mathcal{T}_{n,\ell_v^-(\epsilon)}|$, $|\mathcal{T}^+(v)| \overset{\mathcal{L}}{=} |\mathcal{T}_{n,\ell_v^+(\epsilon)}|$ and $\mathbf{P}(|\mathcal{F}(v)| \le (1+\delta)|\mathcal{T}^-(v)|) \to 1$ and $\mathbf{P}\left(|\mathcal{F}(v)| \ge (1-\delta)|\mathcal{T}^+(v)|\right) \to 1$ as $n \to \infty$.*

In the above lemma, $\mathcal{F}(v)$ is already independent of everything other than the label of its siblings and its parent, so the above-noted independence lines up with the desired result. The approximation for $\mathcal{T}_{n,p}(v)$ is a little bit simpler. If we remove paths that are non-decreasing and have labels all below $\ell_v^-(\epsilon)$ instead of ℓ_v from T_n, then the resulting tree $\mathcal{T}_2^-(v)$

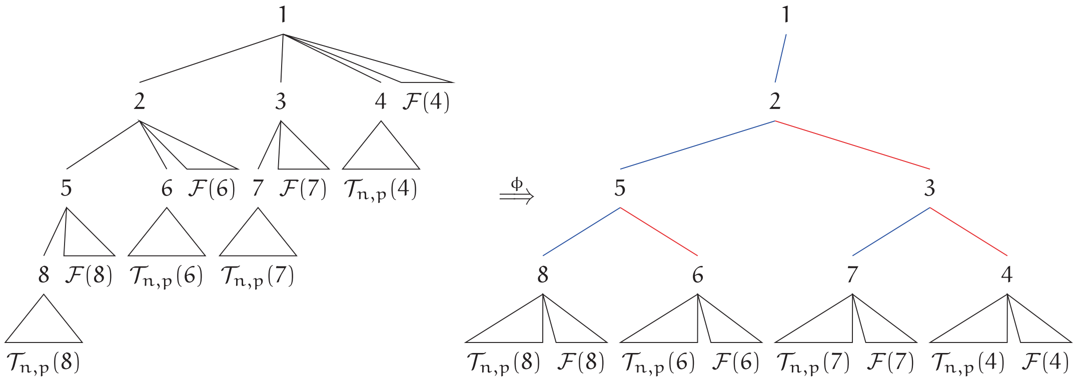

FIGURE 4 | The mapping ϕ up to index $L = 3$. The left tree is $\mathcal{T}_{n,p}$ and the right is the binary tree $\mathcal{T}^*$ with the labelling obtained from ϕ. The vertices are ordered from left to right in order of increasing index in $\mathcal{T}_{n,p}$. A left child (blue edge) in $\mathcal{T}^*$ corresponds to moving down to the vertex's child of the smallest index in $\mathcal{T}_{n,p}$, and a right child (red edge) corresponds to moving to a vertex's sibling of the smallest index in $\mathcal{T}_{n,p}$.

is at least as large as $\mathcal{T}_{n,p}(v)$ and is conditionally independent of the rest of T_n when we condition on the label $\ell_v^-(\epsilon)$ as desired.

Using this lemma, and the fact that the events $S_\leq$ and $\mathcal{E}$ have probability tending to 1 as $n \to \infty$ we see that, for any $x > 0$ and any $\delta > 0$, (1) gives

$$\mathbf{P}\left(\frac{1}{e^{np}}|\mathcal{T}_{n,p}| > x\right) \leq \mathbf{P}\left(\frac{1}{e^{np}}\left(\sum_{v\in I(L+1)} |\mathcal{T}_{n,p}(v)| + |\mathcal{F}(v)|\right) + o_n(1) > x\right) + o_n(1)$$
$$\leq \mathbf{P}\left(\frac{(1+\delta)}{e^{np}}\left(\sum_{v\in I(L+1)} |\mathcal{T}_2^-(v)| + |\mathcal{T}^-(v)|\right) > x\right) + o_n(1).$$

Replacing $\mathcal{T}_2^-(v)$ and $\mathcal{T}^-(v)$ in the above computation with $\mathcal{T}_2^+(v)$ and $\mathcal{T}^+(v)$ yields the corresponding lower bound, completing the proof. ◻

Applying Lemmas 4.1 and 4.2 transforms the analysis of $|\mathcal{T}_{n,p}|$ into the analysis of a branching random walk on $\mathcal{T}^*$. Next, we establish some results for the walk.

Lemma 4.4. *Let $(p_i)_{i=1}^{2^L}$, $(p_i^\pm(\epsilon))_{i=1}^{2^L}$ be as in Lemma 4.2 and set $q_i = p - p_i$, $q_i^\pm(\epsilon) = p - p_i^\pm(\epsilon)$. Introduce the notation $Q_i = nq_i$ and $Q_{i,\epsilon}^\pm = nq_i^\pm(\epsilon)$. For any $n \geq 1$,*

i. *$X_L := \sum_{i=1}^{2^L} e^{-Q_i} \overset{>\mathcal{L}}{\to} \frac{E}{2}$ as $L \to \infty$.*

ii. *Let $X_L^\pm(\epsilon) := \sum_{i=1}^{2^L} e^{-Q_{i,\epsilon}^\pm}$. There is a sequence $\epsilon(L)$ for $L \geq 0$ such that $X_L^\pm(\epsilon(L)) \overset{>\mathcal{L}}{\to} \frac{E}{2}$.*

iii. *There is a sequence $\epsilon(L)$ for $L \geq 0$ such that $\mathbf{E}\left[\frac{\max_{1\leq i\leq 2^L}\exp\left(-Q_{i,\epsilon(L)}^\pm\right)}{X_L^\pm(\epsilon(L))}\right] \to 0$ as $L \to \infty$. Moreover, $\mathbf{E}\left[\frac{\max_{1\leq i\leq 2^L}\exp(-Q_i)}{X_L}\right] \to 0$ as $L \to \infty$.*

Remark. The random variables X_L and $X_L^\pm(\epsilon)$ do not depend on the parameter n. The step sizes of the branching random walks defining each p_i and $p_i^\pm(\epsilon)$ are distributed like $\frac{1}{n}E$ and $\frac{1\pm\epsilon}{n}E$ respectively, so the random variables $Q_i = nq_i$ and $Q_{i,\epsilon}^\pm = nq_i^\pm(\epsilon)$ do not depend on n due to cancellations.

The proof of Lemma 4.4 can be found in the Appendix. All of the required lemmas have now been presented and we are prepared to prove Theorem 2.1.

Proof of Theorem 2.1. Let $x > 0$. By Lemma 4.2 we know that, for any $\epsilon > 0$, $\delta \in (0,1)$, and $L \geq 0$

$$\mathbf{P}\left(\tfrac{1}{e^{np}}|\mathcal{T}_{n,p}| > x\right) \leq \mathbf{P}\left(\tfrac{1}{e^{np}}\sum_{i=1}^{2^L}\left(|\mathcal{T}_{n,p_i^-(\epsilon)}| + |\mathcal{T}'_{n,p_i^-(\epsilon)}|\right) > (1-\delta)x\right) + o_n(1).$$

Let $\delta_2 \in (0,1)$, let $M = \sum_{i=1}^{2^L}\left(e^{n\left(p-q_i^-(\epsilon)\right)} + e^{n\left(p-q_i^-(\epsilon)\right)}\right)$, and let

$$E_{n,\delta_2} = \left\{ \left| \sum_{i=1}^{2^L} \left(|\mathcal{T}_{n,p_i^-(\epsilon)}| + |\mathcal{T}'_{n,p_i^-(\epsilon)}| \right) - M \right| \leq \delta_2 M \right\}.$$

We note that M is the mean of the sum $\sum_{i=1}^{2^L}\left(|\mathcal{T}_{n,p_i^-(\epsilon)}| + |\mathcal{T}'_{n,p_i^-(\epsilon)}|\right)$. Applying Lemma 4.1 we get,

$$\begin{aligned}
\mathbf{P}\left(\frac{1}{e^{np}}|\mathcal{T}_{n,p}| > x\right) &\leq \mathbf{P}\left(\frac{1}{e^{np}} \sum_{i=1}^{2^L} \left(|\mathcal{T}_{n,p_i^-(\epsilon)}| + |\mathcal{T}'_{n,p_i^-(\epsilon)}| \right) > (1-\delta)x \;\middle|\; E_{n,\delta_2} \right) + \mathbf{P}\left(E^c_{n,\delta_2}\right) + o_n(1) \\
&\leq \mathbf{P}\left(\frac{1}{e^{np}} \sum_{i=1}^{2^L} \left(e^{n\left(p-q_i^-(\epsilon)\right)} + e^{n\left(p-q_i^-(\epsilon)\right)} \right) > \frac{1-\delta}{1+\delta_2} x \right) + \frac{5}{\delta_2^2} \mathbf{E}\left[\frac{\max_{1\leq i\leq 2^L} e^{-Q_i}}{\sum_{i=1}^{2^L} e^{-Q_i}} \right] + o_n(1) \\
&= \mathbf{P}\left(\sum_{i=1}^{2^L} 2e^{-Q^-_{i,\epsilon}} > \frac{1-\delta}{1+\delta_2} x \right) + \frac{5}{\delta_2^2} \mathbf{E}\left[\frac{\max_{1\leq i\leq 2^L} e^{-Q_i}}{\sum_{i=1}^{2^L} e^{-Q_i}} \right] + o_n(1).
\end{aligned}$$

Note that the first two terms of the upper bound do not have any dependence on n, and therefore we have

$$\limsup_{n\to\infty} \mathbf{P}\left(\frac{1}{e^{np}}|\mathcal{T}_{n,p}| > x\right) \leq \mathbf{P}\left(\sum_{i=1}^{2^L} 2e^{-Q^-_{i,\epsilon}} > \frac{1-\delta}{1+\delta_2} x \right) + \frac{5}{\delta_2^2} \mathbf{E}\left[\frac{\max_{1\leq i\leq 2^L} e^{-Q_i}}{\sum_{i=1}^{2^L} e^{-Q_i}} \right],$$

for any $\epsilon > 0$, $\delta, \delta_2 \in (0,1)$, and any $L \geq 0$. As δ is arbitrarily small, continuity of measure implies that the inequality can be strengthened to

$$\limsup_{n\to\infty} \mathbf{P}\left(\frac{1}{e^{np}}|\mathcal{T}_{n,p}| > x\right) \leq \mathbf{P}\left(\sum_{i=1}^{2^L} 2e^{-Q^-_{i,\epsilon}} > \left(1+\delta_2\right)^{-1} x \right) + \frac{5}{\delta_2^2} \mathbf{E}\left[\frac{\max_{1\leq i\leq 2^L} e^{-Q_i}}{\sum_{i=1}^{2^L} e^{-Q_i}} \right]. \tag{2}$$

Then, Lemma 4.4 asserts that, for any $\delta_2, \delta_3 \in (0,1)$ we can choose L large and ϵ small enough such that

$$\mathbf{P}\left(\sum_{i=1}^{2^L} 2e^{-Q^-_{i,\epsilon}} > \left(1+\delta_2\right)^{-1} x \right) \leq \mathbf{P}\left(E > \left(1+\delta_2\right)^{-1} x\right) + \delta_3$$

and

$$\frac{5}{\delta_2^2} \mathbf{E}\left[\frac{\max_{1\leq i\leq 2^L} e^{-Q_i}}{\sum_{i=1}^{2^L} e^{-Q_i}} \right] \leq \frac{5}{\delta_2^2}\delta_2^3.$$

Combining these two inequalities with (2) we get.

$$\limsup_{n\to\infty} \mathbf{P}\left(\frac{1}{e^{np}}|\mathcal{T}_{n,p}| > x\right) \leq \mathbf{P}\left(E > \left(1+\delta_2\right)^{-1} x\right) + \delta_3 + 5\delta_2,$$

where δ_2, δ_3 are arbitrary. Taking $\delta_2, \delta_3 \downarrow 0$ and using the continuity of measure gives that

$$\limsup_{n\to\infty} \mathbf{P}\left(\frac{1}{e^{np}}|\mathcal{T}_{n,p}| > x\right) \leq \mathbf{P}(E > x).$$

Repeating an almost identical procedure, we obtain a matching lower bound,

$$\liminf_{n\to\infty} \mathbf{P}\left(\frac{1}{e^{np}}|\mathcal{T}_{n,p}| > x\right) \geq \mathbf{P}(E > x).$$

This completes the proof of the first statement in Theorem 2.1.

Let $v_1, \ldots, v_m$ be the children of the root of rank $1, \ldots, m$ in $\mathcal{T}_{n,p}$. Using the uniform spacings coupling, we can take a vector of spacings $(S_1, \ldots, S_{n+1})$ such that

$$\left(\ell_{v_1}, \ldots, \ell_{v_m}\right) = \left(p - S_1, \ldots, p - S_m\right).$$

By the law of large numbers,

$$X_n := \left(nS_1, \ldots, nS_m\right) \xrightarrow{\mathbb{P}} \left(E_1, \ldots, E_m\right), \tag{3}$$

as $n \to \infty$, where $E_1, \ldots, E_m$ are independent Exponential(1) random variables. Using Lemma 2.1 from Devroye (1986) we can write the joint density function for the vector X_n as

$$f_n\left(x_1, \ldots, x_m\right) = \frac{1}{n} \cdot \frac{n!}{(n-m)!n^m}\left(1 - \frac{\sum_{i=1}^m x_i}{n^2}\right)^n$$

for all $\left(x_1, \ldots, x_m\right) \in \Delta_m := \left\{\left(x_1, \ldots, x_m\right) \;:\; x_i \geq 0,\; \forall 1 \leq i \leq m,\; x_1 + \ldots + x_m \leq n\right\}$. Note that $\frac{n!}{(n-m)!n^m} \leq 1$. For $x, \in \Delta_m$ we have, by using the binomial theorem, the triangle inequality, and the inequality $\binom{n}{i} \leq n^i/i!$,

$$| f_n(x) - f_n(y) | \leq \frac{1}{n}\sum_{i=0}^{n-m}\binom{n}{i}\left|\left\|\frac{x}{n^2}\right\|_1^i - \left\|\frac{y}{n^2}\right\|_1^i\right| \leq \frac{1}{n}\sum_{i=0}^{\infty}\frac{1}{i!}\left|\left\|\frac{x}{n}\right\|_1^i - \left\|\frac{y}{n}\right\|_1^i\right|.$$

Now, suppose that $\| y \|_1 \leq \| x \|_1$. Then, we have,

$$| f_n(x) - f_n(y) | \leq \frac{1}{n}\sum_{i=0}^{\infty}\frac{1}{i!}\left|\left\|\frac{x}{n}\right\|_1^i - \left\|\frac{y}{n}\right\|_1^i\right| = \frac{1}{n}\left(e^{\|x\|_1/n} - e^{\|y\|_1/n}\right).$$

By the same argument for $\| y \|_1 \geq \| x \|_1$, we deduce that

$$| f_n(x) - f_n(y) | \leq \frac{1}{n}\left|e^{\|x\|_1/n} - e^{\|y\|_1/n}\right|.$$

for all $n \geq m$. Since $\| x \|_1, \| y \|_1 \leq n$ by definition, it is straightforward to see that this inequality yields uniform equicontinuity for the sequence of functions $\left(f_n\right)_{n=m}^{\infty}$. Combining this equicontinuity with (3), we can conclude that $\| f_n - f \|_\infty \to 0$ as $n \to \infty$, where f is the density function for the vector $\left(E_1, \ldots, E_m\right)$ (see e.g., Boos (1985) Lemma 1). Now, using the conditional independence of the trees $\mathcal{T}_{n,p}\left(v_1\right), \ldots, \mathcal{T}_{n,p}\left(v_m\right)$ upon the labels of $v_1, .., v_m$ and we fact that $\ell_{v_i} = p - \left(S_1 + \ldots + S_i\right)$ we have, for $x = \left(x_1, \ldots, x_m\right) \in \Delta_m$,

$$\mathbf{P}\left(\frac{|\mathcal{T}_{n,p}\left(v_1\right)|}{e^{np}} \geq y_1, \ldots, \frac{|\mathcal{T}_{n,p}\left(v_m\right)|}{e^{np}} \geq y_m \;\middle|\; X_n = x\right) = \prod_{i=1}^m \mathbf{P}\left(\frac{|\mathcal{T}_{n,p}\left(v_i\right)|}{e^{n\ell_{v_i}}} e^{-\left(x_1 + \ldots + x_i\right)} \geq y_i\right). \tag{4}$$

Then, using the first part of this theorem, we have

$$\mathbf{P}\left(\frac{|\mathcal{T}_{n,p}\left(v_1\right)|}{e^{np}} \geq y_1, \ldots, \frac{|\mathcal{T}_{n,p}\left(v_m\right)|}{e^{np}} \geq y_m \;\middle|\; X_n = x\right) \to \prod_{i=1}^m \mathbf{P}\left(E e^{-\left(x_1 + \ldots + x_i\right)} \geq y_i\right). \tag{5}$$

as $n \to \infty$, where E is an Exponential(1) random variable. Combining the fact that $\| f_n - f \|_\infty \to 0$ as $n \to \infty$ with (5) and applying the dominated convergence theorem, we obtain, for $F_1, \ldots, F_m$ all independent Exponential(1) random variables,

$$\begin{aligned}\mathbf{P}\left(\frac{|\mathcal{T}_{n,p}\left(v_1\right)|}{e^{np}} \geq y_1, \ldots, \frac{|\mathcal{T}_{n,p}\left(v_m\right)|}{e^{np}} \geq y_m\right) &= \int_{\Delta_m} \mathbf{P}\left(\frac{|\mathcal{T}_{n,p}\left(v_1\right)|}{e^{np}} \geq y_1, \ldots, \frac{|\mathcal{T}_{n,p}\left(v_m\right)|}{e^{np}} \geq y_m \;\middle|\; X_n = x\right) f_n(x)\; dx \\ &\to \int_{\Delta_m} \prod_{i=1}^m \mathbf{P}\left(E e^{-\left(x_1 + \cdots + x_i\right)} \geq y_i\right) f(x)\; dx\end{aligned}$$

$$= \int_{\Delta_m} \mathbf{P}\Big(F_1 e^{-x_1} \geq y_1, \ldots, F_m e^{-(x_1+\ldots+x_m)} \Big) f(x)\ dx$$

$$= \mathbf{P}\Big(F_1 e^{-E_1} \geq y_1, \ldots, F_m e^{-(E_1+\cdots+E_m)} \geq y_m \Big).$$

From here, noting that e^{-F_i} is a Uniform$[0,1]$ random variable, the desired result follows. □

5 | The Height

In this section, we prove Theorem 2.2, and begin with a technical lemma that uses Cramér's large deviations theorem Cramér (1938, 1944); Cramér and Touchette (2022).

Lemma 5.1. *Let $k \geq 0$ be a fixed integer, and let K be a uniform random variable on $\{1, \ldots, k\}$. For $i \in \{1, \ldots, k\}$, let G_i be independent gamma (i) random variables and let $X_1, \ldots, X_n$ be independently distributed as G_K. Then, there is a sequence $\phi(k)$ with $\phi(k) \to 0$ as $k \to \infty$, such that for any $0 < x < \frac{k+1}{2}$,*

$$\mathbf{P}\left(\sum_{i=1}^{n} X_i \leq nx \right) \geq \exp\left(-n\left(\log\left(\frac{k}{ex} \right) + \phi(k) + o_n(1) \right) \right). \tag{6}$$

Proof. By Cramér's theorem (Klenke 2008, Theorem 23.11), we have that, for $0 < x < \frac{k+1}{2}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}\left(\sum_{i=1}^{n} X_i \leq nx \right) \geq -\inf_{0<y<x} I(y) = -I(x)$$

where

$$I(x) = \sup_{\lambda \in R}\left(\lambda x - \log \mathbf{E} e^{\lambda X_1} \right)$$

is the Legendre-Fenchel transform of the cumulant-generating function of X_1 (i.e., the logarithm of the moment generating function $\mathbf{E}e^{\lambda X_1}$). Observe that, for $\lambda \in (0,1)$,

$$\mathbf{E}\left(e^{\lambda X_1} \right) = \frac{1}{k} \sum_{i=1}^{k} \mathbf{E}\left(e^{\lambda G_i} \right)$$

$$= \frac{1}{k} \sum_{i=1}^{k} \frac{1}{(1-\lambda)^i}$$

$$= \frac{-1}{k\lambda}\left(1 - \frac{1}{(1-\lambda)^k} \right),$$

and for $\lambda \geq 1$, $\mathbf{E}e^{\lambda X_1} = \infty$. Altogether,

$$\log \mathbf{E} e^{\lambda X_1} = \begin{cases} \infty, & \lambda \geq 1 \\ 1, & \lambda = 0 \\ \frac{-1}{k\lambda}\left(1 - (1-\lambda)^{-k} \right), & \text{otherwise} \end{cases}$$

The function $J(\lambda) := \lambda x - \log \mathbf{E} e^{\lambda X_1}$ is concave. To see this, recall that the moment generating function is always log-convex and that any linear function minus a convex function is concave. It can also be seen to be continuous by similar logic. Moreover, one can compute that $J(0) = 0$ and $J'(0) = x - \frac{k+1}{2}$. Together these facts imply that $J(\lambda) \leq 0$ for all $\lambda \geq 0$ when $0 < x < \frac{k+1}{2}$. Moreover, for $x > 0$, we have that $\lim_{\lambda \to -\infty} J(\lambda) = -\infty$. Noting that $J(-1/x) > 0$, the above facts imply that $J(\lambda)$ attains a global maximum at some $\lambda_* \in (\infty, 0)$ when $0 < x < \frac{k+1}{2}$. Thus,

$$\begin{aligned} I(x) &= \sup_{\lambda<0}\left(\lambda x + \log(k) + \log(-\lambda) - \log\left(1 - \frac{1}{(1-\lambda)^k}\right)\right) \\ &\leq \log(k) + \sup_{\lambda>0}(\lambda x + \log(-\lambda)) + \underbrace{\log\left(1 - \frac{1}{\left(1-\lambda_*\right)^k}\right)}_{:=\phi(k)} \\ &= \log(k) - \log(x) - 1 + \phi(k) \end{aligned}$$

for all fixed $0 < x < \frac{k+1}{2}$. Note that $\phi(k) \to 0$ as $k \to \infty$. From here, recalling (6) completes the proof. □

Now we are ready to prove Theorem 2.2. For simplicity, we only present the proof for $p = 1$. The extension to the general case is immediate.

Proof of Theorem 2.2. Recall that a vertex $u \in T_n$ is in the temporal tree $\mathcal{T}_n$ when the sequence of labels from the root to u decreases monotonically. Thus, for a vertex in generation d of T_n,

$$\mathbf{P}\left(u \in \mathcal{T}_n\right) = \frac{1}{d!}.$$

By the union bound and Stirling's formula,

$$\mathbf{P}\left(H_n \geq en\right) \leq \frac{n^{en}}{\lceil en \rceil !} \leq \frac{n^{en}}{\Gamma(en+1)} \leq \frac{1}{\sqrt{2\pi en}} = o_n(1).$$

From here, it suffices to show that for every constant $\gamma < e$, and any integer $M > 0$,

$$\mathbf{P}\left(H_n \geq \gamma n + M\right) \to 1$$

as $n \to \infty$. We do this by exhibiting the existence of a vertex in $\mathcal{T}_n$ of depth $\geq \gamma n + M$, following a proof method that goes back to Biggins (1976, 1977). For the rest of the proof, we view our tree in function of the uniform spacings coupling of T_n. Note that this implies that the children of a vertex $v \in T_n$ all have label below ℓ_v.

Let $K > 1$ be an integer. We trim the tree T_n by, for each $v \in T_n$, keeping only the K children with the largest labels ordered from greatest to least label as $v_1, \dots, v_K$. The result of this process is an infinite K-ary tree that we denote by $\mathcal{T}_n^{(K)}$. Recall that, using the uniform spacings coupling, we may write, for $v \in \mathcal{T}_n^{(K)}$,

$$\ell_v = 1 - \sum_{u \in P^-(v)} S^*_{p(u),r(u)},$$

where $S^*_{v,i} = S_{v,1} + \cdots + S_{v,i}$ and $r(u)$ is the rank of vertex u. This means that the labels of vertices in $\mathcal{T}_n^{(K)}$ follow a generalized branching random walk in which the branching factor is K at each generation and the step sizes are distributed like $\left(S_1^*, \dots, S_K^*\right)$ where $S_i^* = S_1 + \cdots + S_i$ and $\left(S_1, \dots, S_{n+1}\right)$ is a vector of uniform spacings (for the rest of the proof we shall refer to these types of generalized step size branching random walks as just branching random walks). Altogether, this implies that

$$\mathbf{P}\left(H_n \leq \gamma n + M\right) \leq \mathbf{P}\left(\bigcap_{v \in \mathcal{T}_n^{(K)} : |v| = \lceil \gamma n + M \rceil} \left\{\sum_{u \in P^-(v)} S^*_{p(u),r(u)} \geq 1\right\}\right).$$

Thus, what matters is the largest label of any vertex at depth $\lceil \gamma n + M \rceil$ in a K-ary branching random walk in which the children have displacements distributed as $S_1^*, \dots, S_K^*$. Let D be the maximal label of any vertex at distance M from the root. The K^M vertices at distance M from the root have subtrees that behave in an i.i.d. manner, and each vertex in these subtrees has a label at most equal to D plus the total displacement within its subtree. Therefore,

$$\mathbf{P}\left(\bigcap_{v \in \mathcal{T}_n^{(K)} : |v| = \lceil \gamma n + M \rceil} \left\{ \sum_{u \in P^-(v)} S^*_{p(u),r(u)} \geq 1 \right\} \right) \leq \mathbf{P}\left(\bigcap_{v \in \mathcal{T}_n^{(K)} : |v| = \lceil \gamma n + M \rceil} \left\{ \sum_{u \in P^-(v)} S^*_{p(u),r(u)} \geq 1 - D \right\} \right)^{K^M}$$
$$\leq \mathbf{P}(D > \epsilon) + A_n(\epsilon)^{K^M}, \tag{7}$$

where

$$A_n(\epsilon) = \mathbf{P}\left(\bigcap_{v \in \mathcal{T}_n^{(K)} : |v| = \lceil \gamma n + M \rceil} \left\{ \sum_{u \in P^-(v)} S^*_{p(u),r(u)} \geq 1 - \epsilon \right\} \right).$$

Recall that, jointly over $1 \leq i \leq K$,

$$S_i^* \stackrel{\mathcal{L}}{=} \frac{E_1 + \cdots + E_i}{E_1 + \cdots + E_{n+1}},$$

where $E_1, \ldots, E_{n+1}$ are i.i.d. exponential random variables. As D is smaller than the sum of $2K^M$ random variables distributed as S_K^*, we have by Markov's inequality,

$$\mathbf{P}(D > \epsilon) \leq \frac{2K^M}{\epsilon} \frac{K}{n+1} = o_n(1).$$

We show that for special choices of $\epsilon > 0$, $K > 0$,

$$A_n(\epsilon) \leq q < 1, \tag{8}$$

for all n large enough. Then the right-hand side in (7) is upper bounded by $o_n(1) + q^{K^M}$, which can be made as small as desired by taking M large enough and letting n tend to infinity. We conclude by establishing (8).

The dependence of the distribution of S_i^* on n is a slight inconvenience, so we consider a branching random walk with larger displacements. To that end, we introduce a Bernoulli$\left(p_n\right)$ random variable B_n, where

$$p_n = \mathbf{P}\left(E_{K+1} + \cdots + E_{n+1} \leq n(1-\epsilon)\right).$$

Then, the values of the vertices in the branching random walk on $\mathcal{T}_n^{(K)}$ defined by step sizes

$$\left(1 - B_n\right) \frac{E_1 + \cdots + E_i}{n(1-\epsilon)} + B_n \quad 1 \leq i \leq K$$

dominate the values for the vertices in the original branching random walk described above. By the law of large numbers, for fixed K and ϵ, $p_n \to 0$ as $n \to \infty$. So, given $\delta > 0$, $p_n \leq \delta$ for all n large enough. Let B be a Bernoulli(δ) random variable. We introduce a final branching random walk where the family of step sizes, for $1 \leq i \leq K$, is

$$W_i = \begin{cases} E_1 + \cdots + E_i & \text{if B} = 0 \\ \infty & \text{if B} = 1 \end{cases}.$$

Let, for all $1 \leq i \leq K$ and $v \in \mathcal{T}_n^{(K)}$,

$$W_{v,i} = \begin{cases} E_{v,1} + \cdots + E_{v,i} & \text{if } \mathrm{B_v} = 0 \\ \infty & \text{if } \mathrm{B_v} = 1 \end{cases},$$

where $\left(E_{v,i} : 1 \leq i \leq K, v \in \mathcal{T}_n^{(K)}\right)$ are i.i.d. Exponential(1) random variables and B_v are independent Bernoulli(δ) random variables. For sufficiently large n we have,

$$A_n(\epsilon) \leq \mathbf{P}\left(\bigcap_{v \in \mathcal{T}_n^{(K)} : |v| = \lceil \gamma n + M \rceil} \left\{ \sum_{u \in P^-(v)} W_{p(u),r(u)} \geq n(1-\epsilon)^2 \right\} \right). \tag{9}$$

As the step sizes of this new branching random walk do not depend upon n, we can identify a supercritical Galton-Watson process with an extinction probability that upper bounds the right-hand side of (9).

Define $c = (1-2\epsilon)/\gamma$. Let W be a branching random walk with step sizes $(W_1, \ldots, W_k)$. Fix an integer L and consider all K^L vertices in generation L of W. If any of the Bernoulli random variables for the vertices in generation $\ell \le L$ is one, then the root has no children. Otherwise, we set vertices in generation L to be a child of the root (in the new Galton-Watson process) if its label is $\le c$. For each child of the root with the new Galton-Watson process, we repeat this procedure.

To be more specific, let u be a vertex with value U in W, and suppose that u is also in the Galton-Watson process. If any descendant of u up to L generations below u has its Bernoulli value set to one, then u has no children. Otherwise, we add all vertices that are L generations below u with values below $U + Lc$ as children of u. We call the resultant of this Galton-Watson process G.

If $|G| = \infty$, then the values of the vertices at level j are $\le Lj(1-2\epsilon)/\gamma$ for all $j \ge 1$. These vertices correspond to vertices in the original tree at level Lj. In particular, conditioned on survival, there are vertices at level $\lceil \gamma n \rceil$ that have value

$$\le \left\lfloor \frac{\lceil \gamma n \rceil}{L} \right\rfloor \times \frac{L(1-2\epsilon)}{\gamma} \le \frac{1}{\gamma} + (1-2\epsilon)n < (1-\epsilon)^2 n,$$

for n large enough. In particular, combined with (9), this implies that $A_n(\epsilon) \le \mathbf{P}(|G| < \infty)$.

To determine the survival probability of G, we check that the expected number of children of the root, called G_1, is larger than one. Indeed,

$$\begin{aligned}\mathbf{E}\left[G_1\right] &= (1-\delta)^{K^L} \sum_{|v|=L} \mathbf{P}(\text{the value of } v \text{ is at most Lc}) \\ &= (1-\delta)^{K^L} K^L \mathbf{P}(\text{a uniform vertex in generation L has value at most Lc}) \\ &= (1-\delta)^{K^L} K^L \mathbf{P}\left(G_{Y_1} + \cdots + G_{Y_L} \le \text{Lc}\right),\end{aligned}$$

where $Y_1, \ldots, Y_L$ are i.i.d. random integers uniformly distributed on $\{1, \ldots, K\}$, and G_m stands for a gamma random variable with parameter m. By Lemma 5.1,

$$\mathbf{P}\left(G_{Y_1} + \cdots + G_{Y_L} \le Lc\right) \ge \exp\left(-L\left(\log\left(\frac{K}{ce}\right) + \phi(K) + o_L(1)\right)\right) = \frac{(ce)^L}{K^L} e^{-L\phi(K) - o_L(L)},$$

for $c \le \frac{K+1}{2}$, and so

$$\mathbf{E}\left[G_1\right] \ge (1-\delta)^{K^L} (ce)^L e^{-L\phi(K) - o_L(L)}.$$

For fixed $\gamma < e$, we can find $\epsilon > 0$ small enough such that $ce > 1$. Then, we can choose L, K large enough and $\delta > 0$ small enough that the expected number of children is strictly larger than one. Then, with these chosen c, L, K, δ, G becomes extinct with some probability $q < 1$. As noted above, this implies that $A_n(\epsilon) \le q$ for all n large enough, finishing the proof. □

6 | Typical Depths

In this section, we prove Theorem 2.3. As in the previous section, we present the proof for the $p = 1$ case only, as the extension to the general case is straightforward.

Let $(\mathcal{Z}_k)_{k \ge 0}$ be the number of vertices in generation k of the tree $\mathcal{T}_n$. Let H_n be the height of $\mathcal{T}_n$. Since we argued previously that $H_n/n \overset{\mathbb{P}}{\to} e$, it is sufficient to argue that $\mathcal{Z}_1 + \ldots + \mathcal{Z}_{\lfloor (1-\epsilon)n \rfloor}$ and $\mathcal{Z}_{\lfloor (1+\epsilon)n \rfloor} + \ldots + \mathcal{Z}_{\lfloor 2en \rfloor}$ are both negligible compared to $|\mathcal{T}_n|$ for any $\epsilon \in (0, 1)$. One may use Markov's inequality, the union bound, and Stirling's formula. Indeed, for $\delta < \epsilon^2$,

$$\begin{aligned}
\mathbf{P}\left(\mathcal{Z}_1+\cdots+\mathcal{Z}_{\lfloor(1-\epsilon)n\rfloor}>e^{(1-\delta)n}\right) &\leq \sum_{i=1}^{\lfloor(1-\epsilon)n\rfloor}\mathbf{P}\left(\mathcal{Z}_i>\frac{1}{(1-\epsilon)n}e^{(1-\delta)n}\right)\\
&\leq (1-\epsilon)^2n^2\frac{\mathbf{E}\left[\mathcal{Z}_{\lfloor(1-\epsilon)n\rfloor}\right]}{e^{(1-\delta)n}}\\
&\leq \frac{(1-\epsilon)^2n^2n^{(1-\epsilon)n}}{(\lfloor(1-\epsilon)n\rfloor)!e^{(1-\delta)n}}\\
&\leq C\frac{n^3e^{(1-\epsilon)n-(1-\delta)n}}{(1-\epsilon)^{(1-\epsilon)n}}\\
&\leq Cn^3\exp(\delta n-\epsilon n+(1-\epsilon)\log(1-\epsilon)n)\\
&\leq Cn^3\exp(\delta n-\epsilon n+(1-\epsilon)\epsilon n)\to 0,
\end{aligned}$$

where $C>0$ is a constant. For the other side, we repeat the same computation, with the only changes being that $(1-\epsilon)$ is replaced with $(1+\epsilon)$ and that the expected values are slightly different for the $\mathcal{Z}_i$s. Indeed, for $\delta<\epsilon^3$,

$$\begin{aligned}
\mathbf{P}\left(\mathcal{Z}_{\lfloor(1+\epsilon)n\rfloor}+\ldots+\mathcal{Z}_{\lfloor 2en\rfloor}>e^{(1-\delta)n}\right) &\leq \frac{(2e)^2n^2n^{(1+\epsilon)n}}{(\lfloor(1+\epsilon)n\rfloor)!e^{(1-\delta)n}}\\
&\leq C'\frac{n^3e^{(1+\epsilon)n}}{(1+\epsilon)^{(1+\epsilon)n}e^{(1-\delta)n}}\\
&\leq C'n^3\exp(\epsilon n-\delta n-(1+\epsilon)\log(1+\epsilon)n)\\
&\leq C'n^3\exp\left(\epsilon n-\delta n-(1+\epsilon)\left(\epsilon-\epsilon^2\right)n\right)\to 0,
\end{aligned}$$

for some constant $C'>0$. Finally, applying Theorem 2.1 allows us to assert that $\mathbf{P}\left(|\mathcal{T}_n|\leq e^{(1-\delta/2)n}\right)\to 0$ as $n\to\infty$. Choosing $\delta=\epsilon^4$ (recall that we assume $\epsilon<1$) we are able to conclude that

$$\frac{\mathcal{Z}_1+\cdots+\mathcal{Z}_{\lfloor(1-\epsilon)n\rfloor}+\mathcal{Z}_{\lfloor(1+\epsilon)n\rfloor}+\cdots+\mathcal{Z}_{\lfloor 2en\rfloor}}{|\mathcal{T}_n|}\xrightarrow{\mathbb{P}}0,$$

as $n\to\infty$ for any $\epsilon>0$. From the remarks at the beginning of the proof, this means that

$$\frac{\mathcal{Z}_{\lfloor(1-\epsilon)n\rfloor}+\cdots+\mathcal{Z}_{\lfloor(1+\epsilon)n\rfloor}}{|\mathcal{T}_n|}\xrightarrow{\mathbb{P}}1,$$

as $n\to\infty$ for any $\epsilon>0$, which means that uniformly chosen vertices will be between depth $(1-\epsilon)n$ and $(1+\epsilon)n$ for any $\epsilon>0$ with probability tending to 1 as $n\to\infty$.

7 | The Expected Degree Distribution

In this section, we prove Theorem 2.4. Once again, for the sake of clarity of the presentation, we only consider the case $p=1$. The extension to the general case is immediate.

Consider a vertex u of depth ℓ in T_n. The vertex has degree at least k in $\mathcal{T}_n$ if and only if the labels of the ℓ edges on the path from the root to u are decreasing, moreover, at least k of the n labels of the out-edges of u have labels less than the minimum edge label on the path from the root to u. The probability of this event equals

$$\frac{1}{\ell!}\cdot\frac{n}{n+\ell}\cdot\frac{n-1}{n-1+\ell}\cdots\frac{n-k+1}{n-k+1+\ell}.$$

Thus, the expected number of vertices of outdegree at least k is

$$\mathbf{E}L_{n,\geq k}=\sum_{\ell=1}^{\infty}\frac{n^\ell}{\ell!}\prod_{i=0}^{k-1}\frac{1}{1+\frac{\ell}{n-i}}=e^n\mathbf{E}\prod_{i=0}^{k-1}\frac{1}{1+\frac{X_n}{n-i}},$$

where X_n is a Poisson(n) random variable. Since $X_n/(n-i) \to 1$ in probability for each fixed $i \in \{0, 1, \ldots, k-1\}$, and for every $k \geq 1$,

$$\lim_{n\to\infty} \frac{\mathbf{E}L_{n,\geq k}}{e^n} = 2^{-k},$$

which implies Theorem 2.4 for $p = 1$.

Acknowledgments

Caelan Atamanchuk acknowledges the support of the NSERC Postgraduate Scholarship-Doctoral. Luc Devroye acknowledges the support of NSERC grant RGPIN-2024-04164. Gábor Lugosi acknowledges the support of Ayudas Fundación BBVA a Proyectos de Investigación Cientfica 2021 and the Spanish Ministry of Economy and Competitiveness grant PID2022-138268NB-I00, financed by MCIN/AEI/10.13039/501100011033, FSE+MTM2015-67304-P, and FEDER, EU. The authors would like to thank the two referees. Their careful reading and comments greatly improved the quality and clarity of the paper.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2024-04164); Spanish Ministry of Economy and Competitiveness (PID2022-138268NB-I00); NSERC Postgraduate Scholarship; Fundación BBVA a Proyectos de Investigación Científica 2021.

Data Availability Statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

References

Addario-Berry, L., and L. Eslava. 2018. "High Degrees in Random Recursive Trees." *Random Structures and Algorithms* 52, no. 4: 560–575.

Addario-Berry, L., and B. Reed. 2009. "Minima in Branching Random Walks." *Annals of Probability* 37: 1044–1079.

Angel, O., A. Ferber, B. Sudakov, and V. Tassion. 2020. "Long Monotone Trails in Random Edge-Labellings of Random Graphs." *Combinatorics, Probability and Computing* 29, no. 1: 22–30.

Atamanchuk, C., L. Devroye, and G. Lugosi. 2024. "On the Size of Temporal Cliques in Subcritical Random Temporal Graphs." arXiv, 2404.04462.

Becker, R., A. Casteigts, P. Crescenzi, et al. 2022. "Giant Components in Random Temporal Graphs." arXiv, 2205.14888.

Berestycki, J., É. Brunet, and Z. Shi. 2013. "The Number of Accessible Paths in the Hypercube." *Bernoulli* 22, no. 2: 653–680.

Berestycki, J., É. Brunet, and Z. Shi. 2014. "Accessibility Percolation With Backsteps." arXiv, 1401.6894.

Biggins, J. 1976. "The First and Last-Birth Problems for a Multitype Age-Dependent Branching Process." *Advances in Applied Probability* 8: 446–459.

Biggins, J. 1977. "Chernoff's Theorem in the Branching Random Walk." *Journal of Applied Probability* 14: 630–636.

Boos, D. D. 1985. "A Converse to Scheffe's Theorem." *Annals of Statistics* 13: 423–427.

Broutin, N., N. Kamčev, and G. Lugosi. 2023. "Increasing Paths in Random Temporal Graphs." arXiv, 2306.11401.

Casteigts, A., M. Raskin, M. Renken, and V. Zamaraev. 2024. "Sharp Thresholds in Random Simple Temporal Graphs." *SIAM Journal on Computing* 53, no. 2: 346–388.

Coath, C. D., R. C. Steele, and W. F. Lunnon. 2013. "Statistical Bias in Isotope Ratios." *Journal of Analytical Atomic Spectrometry* 28, no. 1: 52–58.

Cramér, H. 1938. *Sur un nouveau théorème-limite de la théorie des probabilités. In Colloque consacré à la Théorie des Probabilités*. Vol. 736, 2–23. Hermann.

Cramér, H. 1944. "On a New Limit Theorem of the Theory of Probability." *Uspekhi Matematicheskikh Nauk* 10: 166–178.

Cramér, H., and H. Touchette. 2022. "On a New Limit Theorem in Probability Theory (Translation of 'sur un nouveau théorème-limite de la théorie des probabilités')." arXiv 1802.05988.

Devroye, L. 1986. *Non-Uniform Random Variate Generation*. Springer-Verlag.

Devroye, L. 1987. "Branching Processes in the Analysis of the Heights of Trees." *Acta Informatica* 24, no. 3: 277–298.

Devroye, L. 1988. "Applications of the Theory of Records in the Study of Random Trees." *Acta Informatica* 26, no. 1–2: 123–130.

Devroye, L. 1998. "Universal Limit Laws for Depths in Random Trees." *SIAM Journal on Computing* 28, no. 2: 409–432.

Devroye, L., and J. Lu. 1995. "The Strong Convergence of Maximal Degrees in Uniform Random Recursive Trees and Dags." *Random Structures and Algorithms* 7, no. 1: 1–14.

Drmota, M. 2009. *Random Trees: An Interplay Between Combinatorics and Probability*. Springer Science & Business Media.

Durrett, R. 2019. *Probability: Theory and Examples*. Vol. 49. Cambridge University Press.

Eslava, L. 2022. "Depth of Vertices With High Degree in Random Recursive Trees." *Latin American Journal of Probability and Mathematical Statistics* 19, no. 1: 839–857.

Holme, P., and J. Saramäki. 2012. "Temporal Networks." *Physics Reports* 519, no. 3: 97–125.

Holme, P., and J. Saramäki. 2013. "Temporal Networks as a Modeling Framework." In *Temporal Networks*, 1–14. Springer.

Holme, P., and J. Saramäki. 2019. *Temporal Network Theory*. Vol. 2. Springer.

Hosseinzadeh, M. M., M. Cannataro, P. H. Guzzi, and R. Dondi. 2022. "Temporal Networks in Biology and Medicine: A Survey on Models, Algorithms, and Tools." *Network Modeling Analysis in Health Informatics and Bioinformatics* 12, no. 1: 10.

Janson, S. 2005. "Asymptotic Degree Distribution in Random Recursive Trees." *Random Structures and Algorithms* 26, no. 1–2: 69–83.

Klenke, A. 2008. *Probability Theory*. Springer.

Martinsson, A. 2015. "Accessibility Percolation and First-Passage Percolation on the Hypercube," arXiv, 1501.02206.

Meir, A., and J. W. Moon. 1978. "On the Altitude of Nodes in Random Trees." *Canadian Journal of Mathematics* 30, no. 5: 997–1015.

Mertzios, G. B., S. Nikoletseas, C. Raptopoulos, and P. G. Spirakis. 2024. "On the Existence of δ-Temporal Cliques in Random Simple Temporal Graphs." arXiv, 2404.07147.

Nowak, S., and J. Krug. 2013. "Accessibility Percolation on n-Trees." *Europhysics Letters* 101, no. 6: 66004.

Pittel, B. 1994. "Note on the Heights of Random Recursive Trees and Random m-Ary Search Trees." *Random Structures & Algorithms* 5, no. 2: 337–347.

Roberts, M., and L. Zhao. 2013. "Increasing Paths in Regular Trees." *Electronic Communications in Probability* 18, no. 87: 1–10.

Sanjay Kumar, N. B., and B. S. Panda. 2020. "Analysis of Social Network Metrics Based on the Model of Random Recursive Tree." *Journal of Interdisciplinary Mathematics* 23, no. 1: 237–246.

Shi, Z. 2015. *Branching Random Walks*. Vol. 2151. Springer.

Appendix A

Proof of Lemma 4.1. Recall that at the beginning of Section 4 we showed that $\mathbf{E}\mid\mathcal{T}_{n,p}\mid = e^{np}$. The second moment of $\mid\mathcal{T}_{n,p}\mid = \sum_{v\in T_n}\mathbf{1}_{\{v\in\mathcal{T}_{n,p}\}}$ is a sum, over pairs of vertices of T_n, the probability that both vertices exist in $\mathcal{T}_{n,p}$. We may split the sum based on where the pairs of paths stop overlapping,

$$\mathbf{E}\left|\mathcal{T}_{n,p}\right|^2 \leq \mathbf{E}\mid\mathcal{T}_{n,p}\mid + 2\sum_{k=0}^{\infty}\sum_{\ell=k}^{\infty}\sum_{m=0}^{k}\underbrace{n^{k-m}\binom{n}{2}n^{m-1}n^{\ell-k+m-1}p^{\ell+m}}_{\text{I}}\underbrace{\frac{1}{m!}\frac{1}{(\ell-k+m)!}}_{\text{II}}\underbrace{\frac{(\ell-k+2m)!}{(\ell+m)!}}_{\text{III}}.$$

Term I comes from choosing the pairs of paths and ensuring the paths have edge labels below p, term II is the probability that the portion of both paths after the overlap is decreasing (m and $\ell-k+m$ edges), and term III is the probability that the $(k-m)$ edges in the overlap are such that both paths as a whole are decreasing. Rearranging this expression and swapping the order of summation gives

$$\begin{aligned}\mathbf{E}\left|\mathcal{T}_{n,p}\right|^2 &\leq \mathbf{E}\mid\mathcal{T}_{n,p}\mid + \sum_{k=0}^{\infty}\sum_{\ell=k}^{\infty}\sum_{m=0}^{k}\frac{(np)^{\ell+m}}{(\ell+m)!}\binom{\ell-k+2m}{m}\\ &\leq \mathbf{E}\mid\mathcal{T}_{n,p}\mid + \sum_{\ell=0}^{\infty}\sum_{m=0}^{\ell}\frac{(np)^{\ell+m}}{(\ell+m)!}\left(\sum_{k=m}^{\ell}\binom{\ell-k+2m}{m}\right).\end{aligned}$$

Writing $\binom{\ell-k+2m}{m} := f(k)$, we can bound the innermost sum with a geometric series. Since,

$$\frac{f(k+1)}{f(k)} = \frac{(\ell-k+m)!}{(\ell-k+m)!}\frac{(\ell-k+2m-1)!}{(\ell-k+m-1)!} = \frac{\ell-k+m}{\ell-k+2m} \leq \frac{\ell}{\ell+m},$$

for all $k \geq m$, we have that

$$\sum_{k=m}^{\ell}\binom{\ell-k+2m}{m} \leq f(m)\sum_{j=0}^{\infty}\left(\frac{\ell}{\ell+m}\right)^{j} = f(m)\frac{1}{1-\frac{\ell}{\ell+m}} = f(m)\left(1+\frac{\ell}{m}\right).$$

Splitting off the $m = 0$ term from the original expression and applying this bound gives

$$\mathbf{E}\left|\mathcal{T}_{n,p}\right|^{2} \leq \underbrace{\mathbf{E}\mid\mathcal{T}_{n,p}\mid}_{\mathrm{I}} + \underbrace{\sum_{\ell=0}^{\infty}\frac{(np)^{\ell}}{\ell!}2^{\ell}}_{\mathrm{II}} + \underbrace{\sum_{\ell=0}^{\infty}\sum_{m=1}^{\ell}\frac{(np)^{\ell+m}}{(\ell+m)!}\binom{\ell+m}{m}\left(1+\frac{\ell}{m}\right)}_{\mathrm{III}}.$$

Clearly $\mathrm{I} = e^{np}$, $\mathrm{II} = e^{2np}$, and III can be bounded as,

$$\begin{aligned}
\mathrm{III} &\leq 2\sum_{\ell=0}^{\infty}\sum_{m=1}^{\ell}\frac{(np)^{\ell}}{\ell!}\frac{(np)^{m}}{m!}\frac{\ell}{m} \\
&\leq 2e^{2np}\sum_{\ell=0}^{\infty}\sum_{m=1}^{\ell}\mathbf{P}(X_n=\ell)\mathbf{P}(Y_n=m)\frac{\ell}{m} \\
&\leq 2e^{n}\mathbf{E}\left[\frac{X_n}{Y_n}\mathbf{1}_{\{1\leq Y_n\leq X_n\}}\right] \\
&\leq 4e^{2n}\mathbf{E}\left[\frac{X_n}{Y_n+1}\right],
\end{aligned}$$

where X_n and Y_n are independent Poisson(np) random variables. Finally, the value of the expected ratio given above is known to be $(1-e^{-np})$ (see, e.g., Coath et al. (2013)), so we can put the expressions for I, II, and III together to get that

$$\mathbf{E}\left|\mathcal{T}_{n,p}\right|^{2} \leq e^{np} + e^{2np} + 4e^{2np}\left(1-e^{-np}\right) \leq 5e^{2np} = 5\left(\mathbf{E}|\mathcal{T}_{n,p}|\right)^{2}.$$

The second statement of the lemma is a quick corollary of the first. Since the trees are independent of one another, an application of (i) yields

$$\mathrm{Var}\left(\sum_{i=1}^{m}|\mathcal{T}_{n,p_i}|\right) \leq \sum_{i=1}^{m}\mathbf{E}\left|\mathcal{T}_{n,p_i}\right|^{2} \leq 5\sum_{i=1}^{m}e^{2np_i}.$$

Applying Chebyshev's inequality and upper bounding $e^{np_i} \leq \max_{1\leq I\leq m} e^{np_i}$ gives

$$\mathbf{P}\left(\left|\sum_{i=1}^{m}|\mathcal{T}_{n,p_i}| - \mu\right| > \epsilon\mu\right) \leq \frac{5}{\epsilon^2}\frac{\left(\max_{1\leq i\leq m} e^{np_i}\right)\sum_{i=1}^{m}e^{np_i}}{\left(\sum_{i=1}^{m}e^{np_i}\right)^{2}}.$$

Factoring an e^{np} (recall that $q_i = p - p_i$) from both the numerator and denominator completes the proof:

$$\mathbf{P}\left(\left|\sum_{i=1}^{m}|\mathcal{T}_{n,p_i}| - \mu\right| > \epsilon\mu\right) \leq \frac{5}{\epsilon^2}\frac{\max_{1\leq i\leq m} e^{np}e^{-nq_i}}{e^{np}\sum_{i=1}^{m}e^{-nq_i}} = \frac{5}{\epsilon^2}\frac{\max_{1\leq i\leq m} e^{-nq_i}}{\sum_{i=1}^{m}e^{-nq_i}}.$$

□

Proof of Lemma 4.3. Let $v \in T_n$ have fixed finite index, let $v_1, \ldots, v_n$ be the children of v in T_n in order of decreasing label, and let $q \geq 1$ be some fixed integer not depending on n. Notice that v_q is also a finite-index vertex. Take a sequence of independent Exponential(1) random variables $\left(E_i\right)_{i\geq 1}$ such that

$$S_i := S_{v,i} = \frac{E_i}{E_1 + \cdots + E_{n+1}},$$

for all $1 \leq i \leq n+1$, where the collection $\left(S_{v,i} : v \in T_n, 1 \leq i \leq n+1\right)$ contains the spacings from the uniform spacings coupling of T_n that define the labels of vertices in T_n. The only dependence of $\left(S_i : q+1 \leq i \leq n+1\right)$ upon $\left(\ell^{-}_{v_1}(\epsilon), \ldots, \ell^{-}_{v_q}(\epsilon)\right)$ comes from the existence of the random variables $E_1, \ldots, E_q$ in the denominator. We define new spacings for $q+1 \leq i \leq n+q+1$ that are independent of $E_1, \ldots, E_q$,

$$S_i^{*} = \frac{E_{q+i}}{E_{q+1} + \cdots + E_{n+q+1}}.$$

Recall that the forest $\mathcal{F}(v_q)$ may be constructed by, for each $q+1 \le i \le n$, deleting the vertices from $T_n(v_i)$ whose unique path connecting them to v_i are not decreasing if $\ell_{v_i} > 0$, and deleting the whole tree if $\ell_{v_i} < 0$. Starting from the forest containing $T_n(v_i)$ for each $q+1 \le i \le n$, we construct a new tree T_n^-. First, add q new vertices $v_{n+1}, \dots, v_{n+q}$. Make each of these vertices the root of an independent tree distributed like T_n. We define labels for the vertices $v_{q+1}, \dots, v_{n+q}$ (replacing the old labels for $v_{q+1}, \dots, v_n$) by, for each $1 \le i \le n$

$$\ell^*_{v_{q+i}} := \ell^-_{v_q}(\epsilon) - S^*_{q+1} - \cdots - S^*_{q+i}.$$

Finally, we attach $v_{q+1}, \dots, v_{n+q}$ to some mutual shared root. If we delete vertices from T_n^- that have either negative label or are such that their unique path to the root is not decreasing, then, by construction, the resultant tree, which we denote by $\mathcal{T}_n^-$, is distributed exactly as $\mathcal{T}_{n,\ell^-_{v_q}(\epsilon)}$. Moreover, by construction, the tree T_n^- is independent of $T_n(v_1), \dots, T_n(v_q)$ when one conditions on $\left(\ell^-_{v_1}(\epsilon), \dots, \ell^-_{v_q}(\epsilon)\right)$.

Let $k_n = \lfloor n^{1/4} \rfloor$. To prove that $\mathbf{P}\left(|\mathcal{F}(v)| \le (1+\delta)|\mathcal{T}_n^-|\right) \to 1$ as $n \to \infty$ for any $\delta > 0$, it suffices the following three claims:

i. $\mathbf{P}\left(\bigcap_{i=1}^{k_n} \left\{ |\mathcal{T}_{n,p}(v_{q+i})| \le |\mathcal{T}_n^-(v_{q+i})| \right\}\right) \to 1$ as $n \to \infty$.

ii. $\frac{|\mathcal{F}(v)|}{\sum_{i=1}^{k_n} |\mathcal{T}_{n,p}(v_i)|} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$.

iii. $\frac{|\mathcal{T}_n^-|}{\sum_{i=1}^{k_n} |\mathcal{T}_n^-(v_i)|} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$.

We begin with (i). Since $\ell^-_{v_q}(\epsilon) \ge \ell_{v_q}$, it holds that, for $1 \le i \le k_n$,

$$\ell_{v_{q+i}} - \ell^*_{v_{q+i}} \le \left| \sum_{j=1}^{i} \left(S_{q+j} - S^*_{q+j} \right) \right|,$$

with probability tending to 1 as $n \to \infty$. Writing $R = \sum_{j=1}^{q} E_{n+j+1} - \sum_{j=1}^{q} E_j$ and $T = \sum_{i=1}^{n+1} E_i$ we can simplify the above to get

$$\ell_{v_{q+i}} - \ell^*_{v_{q+i}} \le \frac{\left(1 + \frac{R}{T}\right) \sum_{j=1}^{i} E_{q+j}}{\left(1 + \frac{R}{T}\right) T},$$

and so

$$\sup_{1 \le i \le k_n} \left(\ell_{v_{q+i}} - \ell^*_{v_{q+i}} \right) \le \frac{\frac{R}{T} \sum_{j=1}^{k_n} E_{q+j}}{\left(1 + \frac{R}{T}\right) T} \le \frac{\left(1 + \frac{R}{T}\right) \sum_{j=1}^{k_n} E_{q+j}}{T} \le \frac{\left(1 + \frac{R}{T^*}\right) \sum_{j=1}^{k} E_{q+j}}{T^*},$$

with probability tending to 1 as $n \to \infty$, where $T^* := \sum_{i=q+k+1}^{n+1} E_i \le T$. Note that all the terms in the final upper bound are independent. Applying the law of large numbers, we get that, for any $\delta > 0$,

$$\sup_{1 \le i \le k_n} \left(\ell_{v_{q+i}} - \ell^*_{v_{q+i}} \right) \le \frac{(1+\delta) x_n k_n}{n^2}, \tag{A1}$$

with probability tending to 1 as $n \to \infty$ for any sequence $(x_n)_{n=1}^{\infty}$ such that $x_n = \omega_n(1)$ (since R is a finite sum, it does not tend to infinity). By construction, for each $1 \le i \le k_n$, the children of v_{q+i} in both T_n and T_n^- have the same labels. Moreover, the child labels form a vector of independent Uniform[0, 1] random variables. For the vertex v_i, let these uniforms be given by $U_1^{(i)}, \dots, U_n^{(i)}$. If $\left\{U_1^{(i)}, \dots, U_n^{(i)}\right\} \bigcap \left[\ell^*_{v_{q+i}}, \ell_{v_{q+i}}\right] = \emptyset$, then every child of v_{q+i} that is in $\mathcal{T}_{n,p}$ is also in $\mathcal{T}_n^-$. Then, since the labels of these children are the same in both trees, it must hold that $|\mathcal{T}_{n,p}(v_{q+i})| \le |\mathcal{T}_n^-(v_{q+i})|$. Hence,

$$\mathbf{P}\left(\left\{ |\mathcal{T}_{n,p}(v_{q+i})| \ge |\mathcal{T}_n^-(v_{q+i})| \right\}\right) \le \mathbf{P}\left(\bigcup_{j=1}^{n} \left\{ U_j^{(i)} \in \left[\ell^*_{v_{q+i}}, \ell_{v_{q+i}}\right] \right\} \right).$$

By conditioning on (A1) and applying the union bound, we get,

$$\begin{aligned}
\mathbf{P}\left(\bigcup_{i=1}^{k_n} \{ |\mathcal{T}_{n,p}(v_{q+i})| \ge |\mathcal{T}_n^-(v_{q+i})| \} \right) &\le k_n \mathbf{P}\left(\text{Binomial}\left(n, \frac{(1+\delta)x_n k_n}{n^2}\right) > 0 \right) + o_n(1), \\
&\le \frac{(1+\delta) k_n^2 x_n}{n} + o_n(1) = o_n(1)
\end{aligned}$$

when one chooses $x_n = n^{1/4}$.

Now we focus on (ii) and (iii). Both convergences may be proved by an almost identical method, so we only present the proof for the case of $|\mathcal{T}_n^-|$. Clearly, to prove that $\frac{|\mathcal{T}_n^-|}{\sum_{i=1}^{k_n}|\mathcal{T}_n^-(v_i)|} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$, it is enough to show that

$$\frac{\sum_{i=k_n+1}^{n} |\mathcal{T}_n(v_i)|}{|\mathcal{T}_n^-|} \xrightarrow{\mathbb{P}} 0, \tag{A2}$$

as $n \to \infty$. First, observe that $\ell^*_{v_{q+i}} \le \ell^*_{v_{q+k_n}} := p_{k_n}$ for all $k_n \le i \le n$. By replacing the label of each vertex v_{q+i} in T_n^- for $k_n \le i \le n$ with $\ell^*_{v_{q+k_n}}$, we can see that it suffices to show the following two points to deduce (A2):

1. If $\mathcal{T}^{(1)}_{n,p_{k_n}}, \dots, \mathcal{T}^{(n)}_{n,p_{k_n}}$ are all distributed like $\mathcal{T}_{n,p_{k_n}}$, then, as $n \to \infty$

$$\mathbf{P}\left(\sum_{i=1}^{n}|\mathcal{T}^{(i)}_{n,p_{k_n}}| > \exp\left(n\ell^-_{v_q}(\epsilon) - n^{3/16}\right)\right) \xrightarrow{\mathbb{P}} 0.$$

2. $\mathbf{P}\left(|\mathcal{T}_n^-| \le \exp\left(n\ell^-_{v_q}(\epsilon) - n^{1/16}\right)\right) \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$

Let $0 < \delta < 1$. Recall that $p_{k_n} = \ell^-_{v_q}(\epsilon) - S^*_{q+1} - \cdots - S^*_{q+k_n}$. Applying the law of large numbers, we obtain that

$$\mathbf{P}\left(p_k \le \ell^-_{v_q}(\epsilon) - \frac{1-\delta}{n^{3/4}}\right) \to 1, \tag{A3}$$

as $n \to \infty$. Applying the union bound, Markov's inequality, and Lemma 4.1, we have

$$\begin{aligned}
\mathbf{P}\left(\sum_{i=1}^{n}|\mathcal{T}^{(i)}_{n,p_{k_n}}| > \exp\left(n\ell^-_{v_q}(\epsilon) - n^{3/16}\right)\right) &\le n\mathbf{E}\left[\mathbf{P}\left(|\mathcal{T}_{n,\ell^-_{v_q}(\epsilon)-(1-\delta)/n^{3/4}}| > \frac{1}{n}\exp\left(n\ell^-_{v_q}(\epsilon) - n^{3/16}\right) \;\middle|\; \ell^-_{v_q}(\epsilon)\right)\right] + o_n(1)\\
&\le n\mathbf{E}\left[\frac{n\exp\left(n\ell^-_{v_q}(\epsilon) - (1-\delta)n^{1/4}\right)}{\exp\left(n\ell^-_{v_q}(\epsilon) - n^{3/16}\right)}\right] + o_n(1)\\
&\le n^2\exp\left(n^{3/16} - (1-\delta)n^{1/4}\right),
\end{aligned}$$

which converges to 0 as $n \to \infty$. This proves point 1. Now, let $m_n = \lfloor n^{1/8} \rfloor$, and let $p_{m_n} = \ell^*_{v_m}$. Applying the law of large numbers again, we have, for any $\delta > 0$,

$$\mathbf{P}\left(p_m \ge \ell^-_{v_q}(\epsilon) - \frac{1+\delta}{n^{7/8}}\right) \to 1,$$

as $n \to \infty$. Recall that the labels of the vertices $v_{q+1}, \dots, v_{q+m_n}$ are all larger than that of v_{q+m}. Then, applying Lemma 4.1 and the Chebyshev-Cantelli inequality, we obtain

$$\begin{aligned}
\mathbf{P}\left(|\mathcal{T}_n^-| \le \exp\left(n\ell^-_{v_q}(\epsilon) - n^{1/16}\right)\right) &\le \mathbf{P}\left(|\mathcal{T}_{n,\ell^-_{v_q}(\epsilon)-(1+\delta)/n^{7/8}}| \le \exp\left(n\ell^-_{v_q}(\epsilon) - n^{1/16}\right)\right)^m + o_n(1)\\
&\le \mathbf{E}\left[\left(\frac{\mathrm{Var}|\mathcal{T}_{n,\ell^-_{v_q}(\epsilon)-(1+\delta)/n^{7/8}}|}{\mathrm{Var}|\mathcal{T}_{n,\ell^-_{v_q}(\epsilon)-(1+\delta)/n^{7/8}}| + \frac{1}{5}\mathrm{Var}|\mathcal{T}_{n,\ell^-_{v_q}(\epsilon)-(1+\delta)/n^{7/8}}|\left(1 - e^{(1+\delta)n^{1/8} - n^{1/16}}\right)^2}\right)^m\right] + o_n(1),
\end{aligned}$$

which converges to 0 as $n \to \infty$. This proves point 2, and concludes the proof of (ii) and (iii). One can construct a tree $\mathcal{T}_n^+$ in a completely analogous fashion to how $\mathcal{T}_n^-$ was constructed. Following the same proof approach that was used for $\mathcal{T}_n^-$, one can show the following three points:

i. $\mathbf{P}\left(\bigcap_{i=1}^{k_n}\left\{|\mathcal{T}_{n,p}(v_{q+i})| \ge |\mathcal{T}_n^+(v_{q+i})|\right\}\right) \to 1$ as $n \to \infty$.

ii. $\frac{|\mathcal{F}(v)|}{\sum_{i=1}^{k_n}|\mathcal{T}_{n,p}(v_i)|} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$.

iii. $\frac{|\mathcal{T}_n^+|}{\sum_{i=1}^{k_n}|\mathcal{T}_n^+(v_i)|} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$.

As with $\mathcal{T}_n^-$, we obtain the desired result. □

Proof of Lemma 4.4. Let $\left(Y_v : v \in \mathcal{T}^*\right)$ be a branching random walk with step size $\frac{E}{n}$. Note that, by definition, for all $L \geq -1$,

$$\sum_{v:|v|=L} e^{-Y_v} \stackrel{\mathcal{L}}{=} \sum_{i=1}^{2^L} e^{-Q_i}.$$

Define $X_L = \sum_{v:|v|=L} e^{-nY_v}$. Each vertex v has two children, say v_1 and v_2. These children have values $Y_{v_1} = Y_v + \frac{1}{n}E(1,v)$ and $Y_{v_2} \stackrel{\mathcal{L}}{=} Y_v + \frac{1}{n}E(2,v)$, where $E(i,v) \stackrel{\mathcal{L}}{=} \frac{E}{n}$ for any pair (i,v) and are independent of all other edge labels in the graph. Since $nE(i,v) \stackrel{\mathcal{L}}{=} E$ for all (i,v), it holds that $U(i,v) := e^{-nE(i,v)} \stackrel{\mathcal{L}}{=} \text{Uniform}[0,1]$. Thus,

$$\begin{aligned}\mathbf{E}\left[X_{L+1}|X_L\right] &= \sum_{|v|=L}\sum_{i=1}^{2}\mathbf{E}\left[U(i,v)e^{-nY_v}|X_L\right] \\ &= \sum_{|v|=L}\sum_{i=1}^{2}\frac{1}{2}\mathbf{E}\left[e^{nY_v}|X_L\right] \\ &= \mathbf{E}\left[\sum_{|v|=L} e^{-nY_v}|X_L\right] = X_L.\end{aligned}$$

Hence, X_L is a martingale in L with $\sup_L \mathbf{E}\left[X_L\right] < \infty$ for any n, and thus has an almost sure limit. Call this limit X.

Due to the structure of $\mathcal{T}^*$, going down one step in the tree reveals two copies of $\mathcal{T}^*$, both of which have an extra exponential from the first edge in all the vertex values. This structural recursion for the tree implies a distributional equality for the branching random walk:

$$X_L \stackrel{\mathcal{L}}{=} U\left(X_{L-1} + X'_{L-1}\right),$$

where $U \stackrel{\mathcal{L}}{=} \text{Uniform}[0,1]$ and X_{L-1} and X'_{L-1} are two independent copies of X_{L-1}. From this, we obtain the distributional identity, $X \stackrel{\mathcal{L}}{=} U\left(X' + X''\right)$ for X' and X'' independent of each other.

Set $a_k := \mathbf{E}\left[X^k\right]$. Using the distributional identity for X, we obtain the recursion $a_0 = 1$, $a_1 = \frac{1}{2}$,

$$a_k = \frac{1}{k+1}\sum_{i=0}^{k}\binom{k}{i}a_i a_{k-i}.$$

It is easily verified that $a_k = \frac{k!}{2^k}$ solves this recursion. Thus, $\mathbf{E}\left[X^k\right] = \frac{k!}{2^k}$ for all $k \geq 1$, which implies that $X \stackrel{\mathcal{L}}{=} \frac{E}{2}$. The uniqueness of this distributional identity follows from noticing that the exponential distribution satisfies the Stieltjes moment problem conditions (Durrett (2019)). This covers the first claim.

The second claim is a consequence of the Biggins-Hammersley-Kingman theorem (see, e.g., Addario-Berry and Reed (2009)). For our purposes, the theorem implies that both the minimum and maximum value of all vertices in the L-th generation of a branching random walk with step size $\frac{E}{n}$ is $\Theta\left(\frac{L}{n}\right)$ as $L \to \infty$. More precisely, we have constants $C_1, C_2 > 0$ such that

$$\mathbf{P}\left(C_1 L \leq \min_{1\leq i\leq 2^L} nq \leq \max_{1\leq i\leq 2^L} nq_i \leq C_2 L\right) \to 1 \quad \text{as} \quad \text{L} \to \infty.$$

If we take $\epsilon(L) = L^{-2}$, then $\min_{1\leq i\leq 2^L} n\epsilon(L)q_i$ and $\max_{1\leq i\leq 2^L} n\epsilon(L)q_i$ both converge to 0 in probability as $L \to \infty$, and so $\min_{1\leq i\leq 2^L}\exp\left(\pm n\epsilon(L)q_i\right)$ and $\max_{1\leq i\leq 2^L}\exp\left(\pm n\epsilon(L)q_i\right)$ both converge to 1 in probability as $L \to \infty$. Now, from the definition of the values $\left(q_i^\pm(\epsilon)\right)_{i=1}^{2^L}$, we obtain the bounds

$$\min_{1\leq i\leq 2^L}\exp\left(\mp n\epsilon q_i\right)X_L \leq X_L^\pm \leq \max_{1\leq i\leq 2^L}\exp\left(\mp n\epsilon q_i\right)X_L.$$

From Slutsky's theorem and the first claim (i) it holds that both the upper and lower bounds above converge to $\frac{E}{2}$ as $L \to \infty$, and so the same holds for $X_L^\pm$.

For the final claim, note that the aforementioned Biggins-Hammersley-Kingman theorem states that $\max_{1\leq i\leq 2^L}\exp\left(-nq_i^\pm(\epsilon)\right) \to 0$ almost surely as $L \to \infty$. Since

$$\max_{1\leq i\leq 2^L} e^{-n\epsilon q_i^\pm(\epsilon)} \leq 1,$$

the convergence also holds in L_1. Now, let $\eta > 0$ and let $\epsilon(L)$ be as in the second claim. Splitting up the expectation in (iii) gives the upper bound

$$\mathbf{E}\left[\frac{\max_{1\le i\le 2^L}\exp\left(-nq_i^{\pm}(\epsilon)\right)}{\sum_{i=1}^{2^L}\exp\left(-nq_i^{\pm}(\epsilon)\right)}\right] = \mathbf{E}\left[\frac{\max_{1\le i\le 2^L}\exp\left(-nq_i^{\pm}(\epsilon)\right)}{\sum_{i=1}^{2^L}\exp\left(-nq_i^{\pm}(\epsilon)\right)}\mathbf{1}_{\left\{\sum_{i=1}^{2^L}\exp\left(-nq_i^{\pm}(\epsilon)\right)<\eta\right\}}\right]$$
$$+\,\mathbf{E}\left[\frac{\max_{1\le i\le 2^L}\exp\left(-nq_i\right)}{\sum_{i=1}^{2^L}\exp\left(-nq_i^{\pm}(\epsilon)\right)}\mathbf{1}_{\left\{\sum_{i=1}^{2^L}\exp\left(-nq_i^{\pm}(\epsilon)\right)>\eta\right\}}\right]$$
$$\le \mathbf{P}\left(\sum_{i=1}^{2^L}\exp\left(-nq_i^{\pm}(\epsilon)\right) < \eta\right) + \frac{1}{\eta}\mathbf{E}\left[\max_{1\le i\le 2^L}\exp\left(-nq_i^{\pm}(\epsilon)\right)\right].$$

As $L \to \infty$ the final upper bound converges to $\mathbf{P}(E < 2\eta)$. From here, letting $\eta \downarrow 0$ completes the proof. To prove that

$$\mathbf{E}\left[\frac{\max_{1\le i\le 2^L}\exp\left(-nq_i\right)}{\sum_{i=1}^{2^L}\exp\left(-nq_i\right)}\right] \to 0$$

as $L \to \infty$, it suffices to follow the same procedure as the $\pm$ case just covered above. □