Discrete Mathematics and Theoretical Computer Science ....
.... **vol. 25:2 #17 (2023)**

On the protected nodes in exponential recursive trees

Mehri Javanian[1*] Rafik Aguech[2†]

[1] *Department of Statistics, university of Zanjan, Iran*
[2] *Department of mathematics , University of Monastir, Tunisia*

revisions 21st Dec. 2022, 12th June 2023, 18th Aug. 2023, 30th Aug. 2023; accepted 5th Sep. 2023.

The exponential recursive trees model several kinds of networks. At each step of growing of these trees, each node independently attracts a new node with probability p, or fails to do with probability $1 - p$. Here, we investigate the number of protected nodes, total path length of protected nodes, and a mean study of the protected node profile of such trees.

Keywords: exponential recursive trees, protected nodes, limiting distribution, Wasserstein metric, contraction method

---

1 Introduction

A social networking site is an Internet-based platform which people use to build social relationships with friends, family, colleagues, customers, or clients. Social networking can have a social purpose, a business purpose, or both, through sites like Facebook, Twitter, LinkedIn, Instagram, TikTok, Snapchat, Pinterest, Reddit, Tumblr, Telegram, WhatsApp and YouTube. These sites allow people and corporations to connect with one another so they can develop relationships very quickly and so they can share information, ideas, and messages. Social networking has become a significant base for marketers seeking to engage customers. In order to model certain aspects of fast-growing networks, the references Feng and Mahmoud (2018) and Mahmoud (2022) introduce exponential binary trees and exponential recursive trees, respectively.

A rooted tree grown on $n+1$ nodes labeled distinctly with the numbers $1, 2, \dots, n, n+1$; is a *recursive tree* of *age* n, T_n, that is built by attaching, at the nth step, the new node $n + 1$ to one node of a recursive tree of age $n - 1$, T_{n-1}, according to some distribution on the set $\{1, \dots, n\}$ and independently of the structure of the tree T_{n-1}. As an example, Figure 1 shows a sequence of recursive trees growing from a single node labeled with 1, T_0, into a tree of age 8, T_8, in 8 steps. See a survey of results of recursive trees in Smythe and Mahmoud (1995).

The recursive trees are slow-growing where one node is added at each step. So, the recursive tree models cannot be suitable for fast-growing phenomena (e.g., the Corona virus spreads very quickly from

---

*Corresponding author. E-mail: javanian@znu.ac.ir.
†E-mail: rafik.aguech@ipeit.rnu.tn

ISSN 1365–8050

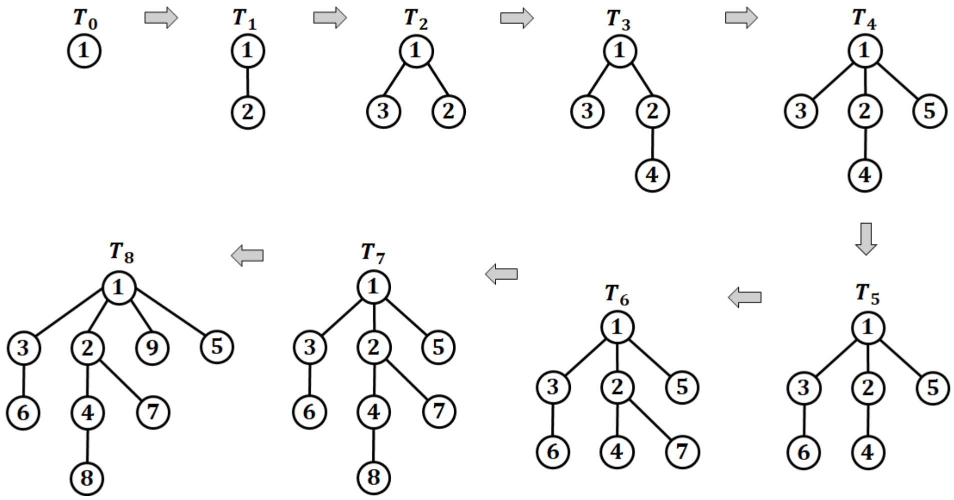


Fig. 1: The evolution of a recursive tree in 8 steps.

person to person). In Mahmoud (2022), a fast-growing analogue of recursive trees has been defined as follows: Initially, T_0 is a tree of a single root node. For $n \geq 1$, at the nth step, a tree of age n, T_n, is constructed when every node of the tree of age $n-1$, T_{n-1}, independently attracts a child with probability $p \in (0,1)$, or not to attract with probability $q := 1-p$. After the nth step, the obtained tree is called *exponential recursive tree (ERT) of index* p. The trees in Figure 2 (a) illustrate a sequence of ERT of index p growing from T_0 into T_7, a tree of probability $p^{19}q^{14}$ in the 7*th* step.

By a *protected node* in a rooted tree, we mean a node that is not a leaf and not all of its children are leaves, for instance see Figure 2 (b). For many types of random trees, protected nodes have been investigated, see e.g. Devroye and Janson (2014), Fuchs et al. (2016), Javanian et al. (2022), Mahmoud and Ward (2015). In this paper, we study the number of protected nodes in exponential recursive trees (ERT). Here, we present the asymptotic expectation, variance and characterizing of the limiting distribution of the number of protected nodes. Via contraction method, we also show the convergence in distribution for the sum of depths of all protected nodes, i.e., the total path length of all protected nodes in ERT of index p. Finally, we derive the expectation of protected node profile, i.e., the number of protected nodes at the same level.

2 Setting and Preliminary Lemmas

Let T_n be an exponential recursive tree of age n and index p. Define

- X_n := the number of protected nodes in T_n,
- $\mathcal{R}_n$:= the event that the root of T_n is protected,
- $|T_n|$:= the number of nodes in T_n (the size of T_n).

At the first step, if the root of T_0 fails to attract a child, then T_1 will be a root and after $n-1$ steps the root of T_1 will produce an exponential recursive tree T'_{n-1}. Alternatively, if the root of T_0 attracts a child,

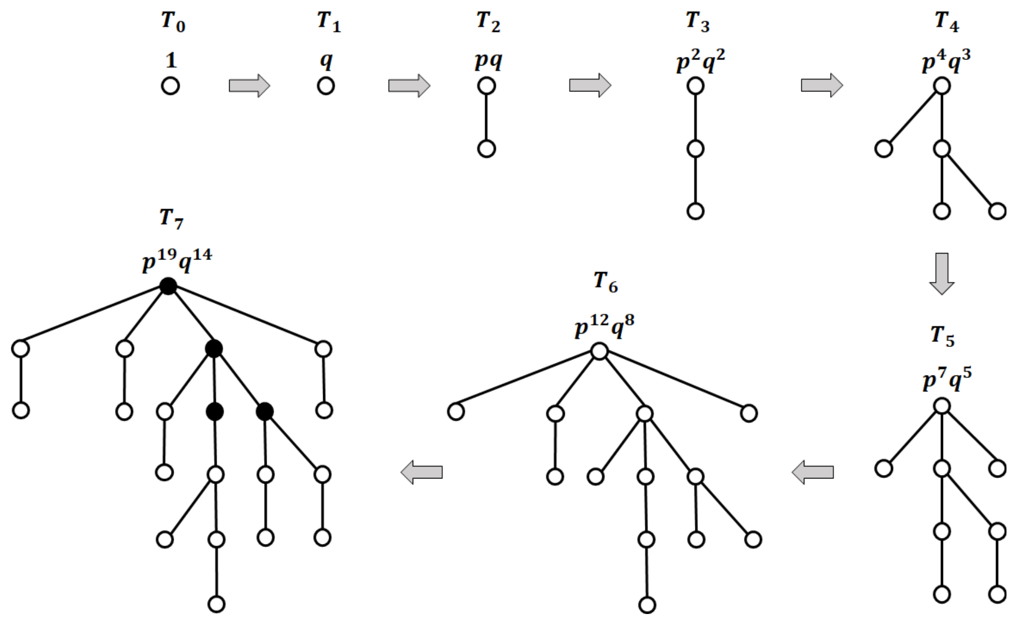


Fig. 2: **(a)** The evolution of an exponential recursive tree (ERT) of index p in 7 steps. The probability of an ERT appears above it. **(b)** In T_7, protected nodes are shown in black.

then, in $n-1$ steps, the child and the root will independently develop two exponential recursive trees T''_{n-1} and T'''_{n-1}, respectively. Let us indicate the number of protected nodes in T'_{n-1}, T''_{n-1} and T'''_{n-1} by X'_{n-1}, X''_{n-1} and X'''_{n-1}, respectively. Therefore, X'_{n-1}, X''_{n-1} and X'''_{n-1} are independent copies of X_{n-1}. We denote the events that the trees T'_{n-1}, T''_{n-1} and T'''_{n-1} have protected roots by $\mathcal{R}'_{n-1}$, $\mathcal{R}''_{n-1}$ and $\mathcal{R}'''_{n-1}$, respectively.

If $\mathbb{1}_A$ denotes the indicator function of an event A, then we define

$$\mathbb{I} := \mathbb{1}_{\{|T_1|=2\}}, \quad \mathbb{J}_n := \mathbb{1}_{\{|T'''_n|=1\}\cap\{|T''_n|\geq 2\}}, \quad \mathbb{G}_n := \mathbb{1}_{\{|T''_n|=1\}\cap\mathcal{R}'''_n}. \tag{1}$$

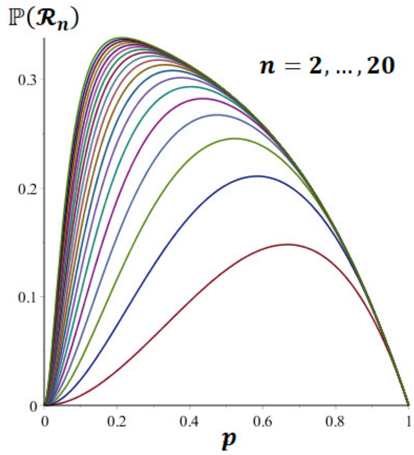
The proof of our results requires to prove the following Lemmas.

Lemma 2.1 *By the above setting, for $n \geq 2$, the probability of $\mathcal{R}_n$ is*

$$\mathbb{P}(\mathcal{R}_n) = \sum_{k=1}^{n-1} pq^k(1-q^k) \prod_{i=k+1}^{n-1} (1-pq^i). \tag{2}$$

Proof: We can observe that

$$\begin{aligned}\mathcal{R}_n = &\Big[\{|T''_{n-1}| \geq 2\} \cap \mathcal{R}'''_{n-1} \cap \{|T_1| = 2\}\Big] \cup \Big[\mathcal{R}'_{n-1} \cap \{|T_1| = 1\}\Big] \\ &\cup \Big[\{|T''_{n-1}| \geq 2\} \cap \{|T'''_{n-1}| = 1\} \cap \{|T_1| = 2\}\Big]\end{aligned}$$



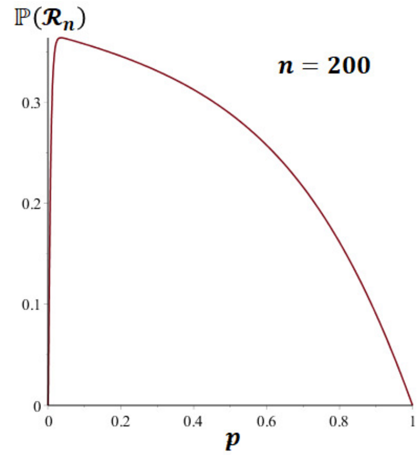


Fig. 3: Graphs of $\mathbb{P}(\mathcal{R}_n)$ versus p, for $n = 2, \ldots, 20$ and $n = 200$. For each p, $\mathbb{P}(\mathcal{R}_n)$ is increasing in n. E.g., the bottom and top curves are the graphs of $\mathbb{P}(\mathcal{R}_2)$ and $\mathbb{P}(\mathcal{R}_{20})$.

is credible for $n \geq 2$. Since $\mathbb{P}(|T''_n| \geq 2) = 1 - \mathbb{P}(|T''_n| = 1) = 1 - q^n$ and the trees T''_{n-1} and T'''_{n-1} are developed independently, then

$$\begin{aligned}\mathbb{P}(\mathcal{R}_n) &= p(1 - q^{n-1})\mathbb{P}(\mathcal{R}'''_{n-1}) + q\mathbb{P}(\mathcal{R}'_{n-1}) + pq^{n-1}(1 - q^{n-1}) \\ &= (1 - pq^{n-1})\mathbb{P}(\mathcal{R}_{n-1}) + pq^{n-1}(1 - q^{n-1}).\end{aligned}$$

Iterating the recurrence we obtain the claim. See Figure 3 as a plot for $\mathbb{P}(\mathcal{R}_n)$. □

Lemma 2.2 *By the above setting, we have the distributional equation*

$$X_n \stackrel{d}{=} X'_{n-1}(1 - \mathbb{I}) + (X''_{n-1} + X'''_{n-1} + \mathbb{J}_{n-1} - \mathbb{G}_{n-1})\mathbb{I}, \tag{3}$$

where the symbol $\stackrel{d}{=}$ *denotes the equality in distribution.*

Proof: In order to construct an exponential recursive tree $\mathcal{T}_n$, a tree distributed like T_n, we can attach the root of the tree T'''_{n-1} to the root of the tree T''_{n-1} by adding an edge. Consequently, if $|T''_{n-1}| \geq 2$ and T'''_{n-1} is a leaf, then the number of protected nodes in $\mathcal{T}_n$ is increased by 1; as the attaching causes the root of $\mathcal{T}_n$ to be protected. From the other point of view, if $|T''_{n-1}| = 1$ and the root of T'''_{n-1} is protected, then this protection is lost after the attaching. Hence we obtain the assertion by (1). □

Lemma 2.3 *For $m \geq 1$, we have the following recurrences for $\mathbb{E}[X_n^m]$ and $\mathbb{E}[X_n^m \mathbb{1}_{\mathcal{R}_n}]$:*

$$\mathbb{E}[X_n^m] = (p+1)\mathbb{E}[X_{n-1}^m] + p\sum_{k=1}^{m-1}\binom{m}{k}\mathbb{E}[X_{n-1}^k]\mathbb{E}[X_{n-1}^{m-k}] + pq^{n-1}\sum_{k=1}^{m-1}\binom{m}{k}\mathbb{E}[X_{n-1}^k]$$
$$+ pq^{n-1}\sum_{k=1}^{m-1}\binom{m}{k}(-1)^{m-k}\mathbb{E}[X_{n-1}^k\mathbb{1}_{\mathcal{R}_{n-1}}] + pq^{n-1}\big(1-q^{n-1}+(-1)^m\mathbb{P}(\mathcal{R}_{n-1})\big),$$
$$\mathbb{E}[X_n^m\mathbb{1}_{\mathcal{R}_n}] = (1-pq^{n-1})\mathbb{E}[X_{n-1}^m\mathbb{1}_{\mathcal{R}_{n-1}}] + p\sum_{k=1}^{m}\binom{m}{k}\mathbb{E}[X_{n-1}^k]\mathbb{E}[X_{n-1}^{m-k}\mathbb{1}_{\mathcal{R}_{n-1}}]$$
$$+ pq^{n-1}\sum_{k=1}^{m}\binom{m}{k}\mathbb{E}[X_{n-1}^k] + pq^{n-1}(1-q^{n-1}).$$

Proof: Raise both sides of (3) to the mth power. So we get

$$X_n^m \stackrel{d}{=} (X'_{n-1})^m(1-\mathbb{I}) + \sum_{\substack{i,j,k,l\geq 0\\ i+j+k+l=m}}\binom{m}{i,j,k,l}(X''_{n-1})^i(X'''_{n-1})^j\mathbb{J}_{n-1}^k(-\mathbb{G}_{n-1})^l\mathbb{I}.$$

Now, separate the second term with $(k=0, l\geq 1)$, $(k\geq 1, l=0)$ and $(k=0, l=0)$. Then

$$X_n^m \stackrel{d}{=} (X'_{n-1})^m(1-\mathbb{I}) + \sum_{\substack{i,j\geq 0,\, l\geq 1\\ i+j+l=m}}\binom{m}{i,j,l}(X''_{n-1})^i(X'''_{n-1})^j(-\mathbb{G}_{n-1})^l\mathbb{I}$$
$$+ \sum_{\substack{i,j\geq 0,\, k\geq 1\\ i+j+k=m}}\binom{m}{i,j,k}(X''_{n-1})^i(X'''_{n-1})^j(\mathbb{J}_{n-1})^k\mathbb{I} + \sum_{\substack{i,j\geq 0\\ i+j=m}}\binom{m}{i,j}(X''_{n-1})^i(X'''_{n-1})^j\mathbb{I}. \qquad (4)$$

Note that, for $i,j,k,l \geq 1$,

$$(\mathbb{G}_{n-1})^l = \mathbb{G}_{n-1} = \mathbb{1}_{\{|T''_{n-1}|=1\}}\mathbb{1}_{\mathcal{R}'''_{n-1}}, \qquad (X''_{n-1})^i\mathbb{1}_{\{|T''_{n-1}|=1\}} = 0$$
$$(\mathbb{J}_{n-1})^k = \mathbb{J}_{n-1} = \mathbb{1}_{\{|T''_{n-1}|\geq 2\}}\mathbb{1}_{\{|T'''_{n-1}|=1\}}, \qquad (X'''_{n-1})^j\mathbb{1}_{\{|T'''_{n-1}|=1\}} = 0.$$

Therefore, the equation (4) can be simplified as

$$X_n^m \stackrel{d}{=} (X'_{n-1})^m(1-\mathbb{I}) + \sum_{j=0}^{m-1}\binom{m}{j}(X'''_{n-1})^j(-1)^{m-j}\mathbb{1}_{\{|T''_{n-1}|=1\}}\mathbb{1}_{\mathcal{R}'''_{n-1}}\mathbb{I}$$
$$+ \sum_{i=0}^{m-1}\binom{m}{i}(X''_{n-1})^i\mathbb{1}_{\{|T''_{n-1}|\geq 2\}}\mathbb{1}_{\{|T'''_{n-1}|=1\}}\mathbb{I} + \sum_{i=0}^{m}\binom{m}{i}(X''_{n-1})^i(X'''_{n-1})^{m-i}\mathbb{I}. \qquad (5)$$

Take the expectation of (5) and observe the independence in T'_{n-1}, T''_{n-1} and T'''_{n-1}, to write

$$\begin{aligned}\mathbb{E}[X_n^m] &= q\mathbb{E}[(X'_{n-1})^m] + p\mathbb{E}[(X'''_{n-1})^m] + p\mathbb{E}[(X''_{n-1})^m] + p\sum_{i=1}^{m-1}\binom{m}{i}\mathbb{E}[(X''_{n-1})^i]\mathbb{E}[(X'''_{n-1})^{m-i}]\\ &\quad + p\sum_{i=1}^{m-1}\binom{m}{i}(-1)^{m-i}\mathbb{E}[(X'''_{n-1})^i\mathbb{1}_{\mathcal{R}'''_{n-1}}]\mathbb{E}[\mathbb{1}_{\{|T''_{n-1}|=1\}}] + (-1)^m p\mathbb{E}[\mathbb{1}_{\mathcal{R}'''_{n-1}}]\mathbb{E}[\mathbb{1}_{\{|T''_{n-1}|=1\}}]\\ &\quad + p\sum_{i=1}^{m-1}\binom{m}{i}\mathbb{E}[(X''_{n-1})^i\mathbb{1}_{\{|T''_{n-1}|\geq 2\}}]\mathbb{E}[\mathbb{1}_{\{|T'''_{n-1}|=1\}}] + p\mathbb{E}[\mathbb{1}_{\{|T''_{n-1}|\geq 2\}}]\mathbb{E}[\mathbb{1}_{\{|T'''_{n-1}|=1\}}].\end{aligned}$$

This yields the first recurrence of the Lemma, by the identical distribution in the subtrees and

$$\mathbb{E}[(X''_{n-1})^i\mathbb{1}_{\{|T''_{n-1}|\geq 2\}}] = \mathbb{E}[(X''_{n-1})^i(1-\mathbb{1}_{\{|T''_{n-1}|=1\}})] = \mathbb{E}[(X''_{n-1})^i], \qquad i\geq 1.$$

By the definitions of the subtrees T'_{n-1}, T''_{n-1} and T'''_{n-1}, observe that

$$\begin{aligned}X_n^m\mathbb{1}_{\mathcal{R}_n} &\overset{d}{=} (X'_{n-1})^m\mathbb{1}_{\mathcal{R}'_{n-1}}(1-\mathbb{I}) + (X''_{n-1}+1)^m\mathbb{1}_{\{|T'''_{n-1}|=1\}}\mathbb{1}_{\{|T''_{n-1}|\geq 2\}}\mathbb{I}\\ &\quad + (X''_{n-1}+X'''_{n-1})^m\mathbb{1}_{\mathcal{R}'''_{n-1}}\mathbb{1}_{\{|T''_{n-1}|\geq 2\}}\mathbb{I}.\end{aligned}$$

Taking the expectation of this equation, it implies the second recurrence of the Lemma. □

3 The Expectation and Variance

In this section, we obtain the expectation and variance of X_n, using Lemma 2.1 and Lemma 2.3.

Theorem 3.1 *Let X_n be the number of protected nodes in an exponential recursive tree of age n and index p. For $q := 1-p$,*

$$\begin{aligned}\mathbb{E}[X_n] &= \frac{p}{q}\Bigg(\sum_{j=2}^{n}(p+1)^{-j}q^j\Big(1-q^{j-1}-\sum_{k=1}^{j-2}pq^k(1-q^k)\prod_{i=k+1}^{j-2}(1-pq^i)\Big)\Bigg)(p+1)^n\\ &=: \mu_{n,p}(p+1)^n, \qquad (\mu_{n,p}<1 \textit{ for } n\geq 1 \textit{ and } p\in(0,1)).\end{aligned} \tag{6}$$

See Figure 4 for some graphs of $\mu_{n,p}$. For $n\geq 1$, $\mu_{n,p}<\mu_{n+1,p}$. So $\mu_p := \lim_{n\to\infty}\mu_{n,p}$ exists.

Proof: In Lemma 2.3, set $m=1$ to obtain the recurrence

$$\mathbb{E}[X_n] = (p+1)\mathbb{E}[X_{n-1}] + pq^{n-1}(1-q^{n-1}) - pq^{n-1}\mathbb{P}(\mathcal{R}_{n-1}),$$

with initial condition $\mathbb{E}[X_1]=0$. Iterating this formula, we find

$$\begin{aligned}\mathbb{E}[X_n] &= \sum_{j=0}^{n-2}(p+1)^j pq^{n-j-1}\big(1-q^{n-j-1}-\mathbb{P}(\mathcal{R}_{n-j-1})\big)\\ &= \frac{p}{q}\Bigg(\sum_{j=2}^{n}(p+1)^{-j}q^j\big(1-q^{j-1}-\mathbb{P}(\mathcal{R}_{j-1})\big)\Bigg)(p+1)^n\end{aligned} \tag{7}$$

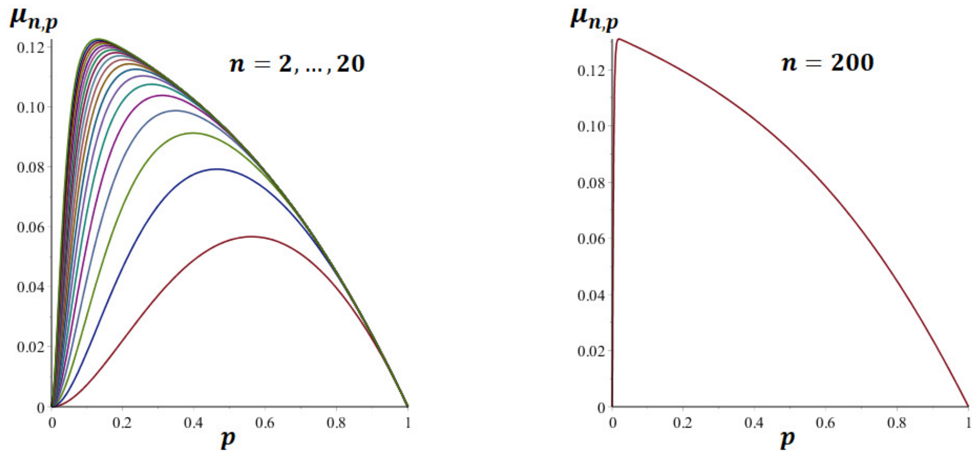


Fig. 4: Graphs of $\mu_{n,p}$ versus $p \in (0,1)$, for $n = 2, \ldots, 20$ and $n = 200$.

Substituting $\mathbb{P}(\mathcal{R}_{j-1})$ from (2) into (7), we get the exact solution in the Theorem. $\square$

In the second moment recurrence, we need $\mathbb{E}[X_n \mathbb{1}_{\mathcal{R}_n}]$ that is given in the following lemma.

Lemma 3.1 *Let $Y_n := X_n \mathbb{1}_{\mathcal{R}_n}$, then*

$$\begin{aligned}\mathbb{E}[Y_n] &= \left(p \sum_{k=1}^{n-1} \mu_{n-k,p}(p+1)^{-k}\big(q^{n-k} + \mathbb{P}(\mathcal{R}_{n-k})\big) \prod_{i=n-k+1}^{n-1} (1 - pq^i)\right)(p+1)^n + \mathbb{P}(\mathcal{R}_n) \\ &=: \alpha_{n,p}(p+1)^n + \mathbb{P}(\mathcal{R}_n) = \alpha_p (p+1)^n + \mathcal{O}(1), \end{aligned} \tag{8}$$

where $\alpha_p := \lim_{n\to\infty} \alpha_{n,p}$ (see Figure 5) and the functions $\mathbb{P}(\mathcal{R}_n)$ and $\mu_{n,p}$ are obtained in Lemma 2.1 and Theorem 3.1, respectively.

Proof: In Lemma 2.3, set $m = 1$ to find the recurrence

$$\mathbb{E}[Y_n] = (1 - pq^{n-1})\mathbb{E}[Y_{n-1}] + p\big(q^{n-1} + \mathbb{P}(\mathcal{R}_{n-1})\big)\mathbb{E}[X_{n-1}] + pq^{n-1}(1 - q^{n-1})$$

with initial condition $\mathbb{E}[Y_1] = 0$. Iterating this formula, we obtain

$$\begin{aligned}\mathbb{E}[Y_n] &= \sum_{k=1}^{n-1} \Big(p\big(q^k + \mathbb{P}(\mathcal{R}_k)\big)\mathbb{E}[X_k] + pq^k(1 - q^k)\Big) \prod_{i=k+1}^{n-1} (1 - pq^i) \\ &= p \sum_{k=1}^{n-1} \mu_{k,p}(p+1)^k \big(q^k + \mathbb{P}(\mathcal{R}_k)\big) \prod_{i=k+1}^{n-1} (1 - pq^i) + \mathbb{P}(\mathcal{R}_n).\end{aligned}$$

This implies the claim. $\square$

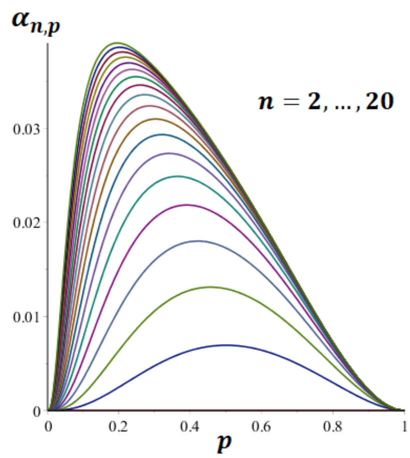


Fig. 5: Graphs of $\alpha_{n,p}$ versus $p \in (0,1)$, for $n = 2, \ldots, 20$.

Theorem 3.2 *By the above definitions for* $\mathbb{P}(\mathcal{R}_n)$*,* $\mu_{n,p}$ *and* $\alpha_{n,p}$*, we have*

$$\mathbb{E}[X_n^2] = \sum_{j=1}^{n-1} \Big(2p\mu_{j,p}^2 (p+1)^{-n+j-1} + 2pq^j(\mu_{j,p} - \alpha_{j,p})(p+1)^{-n-1} + pq^j\big(1 - q^j - \mathbb{P}(\mathcal{R}_j)\big)(p+1)^{-n-j-1}\Big)(p+1)^{2n} =: \beta_{n,p}(p+1)^{2n}.$$

See Figure 6 for some graphs of $\beta_{n,p}$ *with* $\beta_p := \lim_{n\to\infty} \beta_{n,p}$*.*

Proof: In Lemma 2.3, set $m = 2$ to find the recurrence

$$\mathbb{E}[X_n^2] = (p+1)\mathbb{E}[X_{n-1}^2] + 2p\mathbb{E}^2[X_{n-1}] + 2pq^{n-1}\mathbb{E}[X_{n-1}] - 2pq^{n-1}\mathbb{E}[X_{n-1}\mathbb{1}_{\mathcal{R}_{n-1}}] + pq^{n-1}\big(1 - q^{n-1} + \mathbb{P}(\mathcal{R}_{n-1})\big),$$

with boundary condition $\mathbb{E}[X_1^2] = 0$. This standard linear recurrence has the solution

$$\mathbb{E}[X_n^2] = \sum_{j=0}^{n-2} (p+1)^j \Big(2p\mathbb{E}^2[X_{n-j-1}] + 2pq^{n-j-1}\mathbb{E}[X_{n-j-1}] - 2pq^{n-j-1}\mathbb{E}[Y_{n-j-1}] + pq^{n-j-1}\big(1 - q^{n-j-1} + \mathbb{P}(\mathcal{R}_{n-j-1})\big)\Big). \tag{9}$$

After substituting (2), (6), (8) in (9), we have

$$\mathbb{E}[X_n^2] = \sum_{j=0}^{n-2} \Big(2p\mu_{n-j-1,p}^2 (p+1)^{-j-2} + 2pq^{n-j-1}(\mu_{n-j-1,p} - \alpha_{n-j-1,p})(p+1)^{-n-1} + pq^{n-j-1}(p+1)^{-2n+j}\big(1 - q^{n-j-1} - \mathbb{P}(\mathcal{R}_{n-j-1})\big)\Big)(p+1)^{2n}.$$

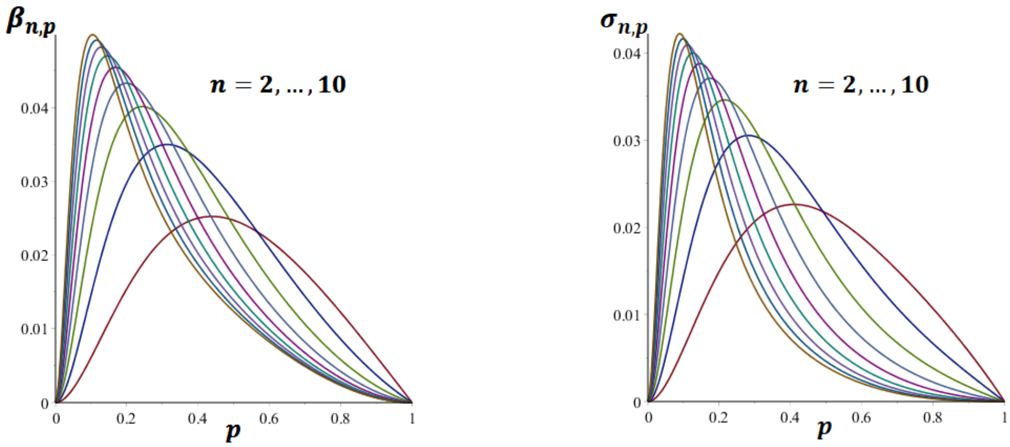


Fig. 6: Graphs of $\beta_{n,p}$ and $\sigma_{n,p}$ versus $p \in (0,1)$, for $n = 2, \ldots, 10$.

This yields the claimed result by changing $n - j - 1 \to j$ in the range of the above sum. □

Corollary 3.1 *By the above definitions for $\mu_{n,p}$ and $\beta_{n,p}$, we have*

$$\mathrm{Var}[X_n] = (\beta_{n,p} - \mu_{n,p}^2)(p+1)^{2n} =: \sigma_{n,p}(p+1)^{2n}.$$

See Figure 6 for some graphs of $\sigma_{n,p}$ with $\lim_{n\to\infty} \sigma_{n,p} =: \sigma_p^2 = \beta_p - \mu_p^2$.

Proof: Using $\mathrm{Var}[X_n] = \mathbb{E}[X_n^2] - \mathbb{E}^2[X_n]$ and Theorem 3.2, the proof is straightforward. □

4 Convergence in Distribution

Here, we characterize the limiting distribution of $\frac{X_n}{(p+1)^n}$, i.e., a scaled version of X_n, the number of protected nodes in an ERT of age n and index p, by its moments.

Theorem 4.1 *Let X_n be the number of protected nodes in an exponential recursive tree of age n and index p. We have the convergence in distribution*

$$\frac{X_n}{(p+1)^n} \xrightarrow{\mathcal{D}} X_*,$$

where the limiting random variable X_* has moments $b_m := \mathbb{E}[X_*^m]$ defined inductively by

$$b_m = \frac{p}{(p+1)^m - (p+1)} \sum_{i=1}^{m-1} \binom{m}{i} b_i b_{m-i}, \qquad m \geq 2,$$

with $b_1 = \mathbb{E}[X_*] = \mu_p = \lim_{n\to\infty} \mu_{n,p}$ at the basis of the induction.

Proof: Scale the two recurrences in Lemma 2.3 by $(p+1)^{nm}$ to get

$$\begin{aligned}\mathbb{E}\Big[\Big(\frac{X_n}{(p+1)^n}\Big)^m\Big] &= \frac{1}{(p+1)^{m-1}}\mathbb{E}\Big[\Big(\frac{X_{n-1}}{(p+1)^{n-1}}\Big)^m\Big]\\ &\quad+\frac{p}{(p+1)^m}\sum_{k=1}^{m-1}\binom{m}{k}\mathbb{E}\Big[\Big(\frac{X_{n-1}}{(p+1)^{n-1}}\Big)^k\Big]\mathbb{E}\Big[\Big(\frac{X_{n-1}}{(p+1)^{n-1}}\Big)^{m-k}\Big]\\ &\quad+pq^{n-1}\sum_{k=1}^{m-1}\binom{m}{k}\frac{1}{(p+1)^{(m-k)n+k}}\mathbb{E}\Big[\Big(\frac{X_{n-1}}{(p+1)^{n-1}}\Big)^k\Big]\\ &\quad+pq^{n-1}\sum_{k=1}^{m-1}\binom{m}{k}\frac{(-1)^{m-k}}{(p+1)^{(m-k)n+k}}\mathbb{E}\Big[\Big(\frac{X_{n-1}}{(p+1)^{n-1}}\Big)^k\mathbb{1}_{\mathcal{R}_{n-1}}\Big]\\ &\quad+\frac{1}{(p+1)^{nm}}\cdot pq^{n-1}\big(1-q^{n-1}+(-1)^m\mathbb{P}(\mathcal{R}_{n-1})\big), \end{aligned} \tag{10}$$

$$\begin{aligned}\mathbb{E}\Big[\Big(\frac{X_n}{(p+1)^n}\Big)^m\mathbb{1}_{\mathcal{R}_n}\Big] &= \frac{1-pq^{n-1}}{(p+1)^m}\mathbb{E}\Big[\Big(\frac{X_{n-1}}{(p+1)^{n-1}}\Big)^m\mathbb{1}_{\mathcal{R}_{n-1}}\Big]+\frac{1}{(p+1)^{nm}}\cdot pq^{n-1}(1-q^{n-1})\\ &\quad+\frac{p}{(p+1)^m}\sum_{k=1}^{m}\binom{m}{k}\mathbb{E}\Big[\Big(\frac{X_{n-1}}{(p+1)^{n-1}}\Big)^k\Big]\mathbb{E}\Big[\Big(\frac{X_{n-1}}{(p+1)^{n-1}}\Big)^{m-k}\mathbb{1}_{\mathcal{R}_{n-1}}\Big]\\ &\quad+pq^{n-1}\sum_{k=1}^{m}\binom{m}{k}\frac{1}{(p+1)^{(m-k)n+k}}\mathbb{E}\Big[\Big(\frac{X_{n-1}}{(p+1)^{n-1}}\Big)^k\Big]. \end{aligned} \tag{11}$$

According to the two recurrences (10) and (11) with finite limits

$$\lim_{n\to\infty}\mathbb{E}\Big[\frac{X_n}{(p+1)^n}\Big]=\mu_p,\quad \text{and} \quad \lim_{n\to\infty}\mathbb{E}\Big[\frac{X_n}{(p+1)^n}\mathbb{1}_{\mathcal{R}_n}\Big]=\alpha_p,$$

we see that $b_m := \lim_{n\to\infty}\mathbb{E}\big[\big(\frac{X_n}{(p+1)^n}\big)^m\big]$ exists, by induction on m, and we get

$$(p+1)^m b_m = (p+1)b_m + p\sum_{i=1}^{m-1}\binom{m}{i}b_i b_{m-i}, \qquad m\ge 2.$$

Using this recurrence and $b_1=\mu_p<1$, we can conclude that $\frac{b_m}{m!}<1$ by induction on m (this induction is shown in page 7 of Mahmoud (2022) and in page 12 of Aguech et al. (2022), as well). Subsequently, for $|z|<1$, the series $\sum_{m=0}^{\infty}\frac{b_m}{m!}z^m$ converges. Therefore, by Theorem 30.1 in Billingsley (2012), $\frac{X_n}{(p+1)^n}$ converges in distribution to a unique limit X_*. □

5 The Total path length

Let I_n be the total path length of protected nodes in an exponential recursive tree T_n (the tree of size n, at time $n-1$), which is the sum of the depths of all protected nodes in T_n. There are two scenarios at the first step: Either the root recruits a node (let us call it v), or it does not. In the first scenario, by time

$n-1$, node v has acquired a subtree (call it $T_{n-1}^{(1)}$) of total path length $I_{n-1}^{(1)}$ (measured from v), with $I_{n-1}^{(1)}$ distributed like I_{n-1}. Measured from the root of T_n, each protected node in $T_{n-1}^{(1)}$ is at depth 1 plus its depth in $T_{n-1}^{(1)}$; there is a total number of protected nodes in $T_{n-1}^{(1)}$ distributed like X_{n-1}. Therefore, the contribution of $T_{n-1}^{(1)}$ to the total path length of protected nodes in T_n is distributed like $I_{n-1}^{(1)} + X_{n-1}$.

In the meantime, the root is actively recruiting. Nodes not in $T_{n-1}^{(1)}$ (including) the root of T_n form a tree of total path length $I_{n-1}^{(2)}$, with $I_{n-1}^{(2)}$ distributed like I_{n-1}. In this scenario, T_n has a total path length of protected nodes distributed like $I_{n-1}^{(1)} + I_{n-1}^{(2)} + X_{n-1}$.

In the second scenario (failure to recruit in the first step), by time $n-1$, the tree T_n has a total path length of protected nodes $\tilde{I}_{n-1}^{(1)}$ distributed like I_{n-1}.

Lemma 5.1

$$I_n \overset{d}{=} \mathbb{I}\left(I_{n-1}^{(1)} + I_{n-1}^{(2)} + X_{n-1}\right) + (1-\mathbb{I})\ \tilde{I}_{n-1}^{(1)} \tag{12}$$

where $I_n^{(1)}$, $(I_n^{(2)}, X_n)$, $\tilde{I}_n^{(1)}$ and $\mathbb{I}$ are independent, $I_n^{(i)} \overset{d}{=} I_n$ for $i = 1, 2$, $\mathbb{I}$ is a Bernoulli random variable with success probability p, and X_n is the number of protected nodes of an exponential recursive tree at age n.

Remark 1 *The internal path length can be obtained by another method, if we denote for all k by*

$$\mathcal{N}_k = \{\text{at step } k \text{ the root recruits a new child}\},$$

then, I_n satisfies, almost surely,

$$I_n = \sum_{k=1}^{n} \mathbb{I}_{\mathcal{N}_k}\, I_{n-k} + X_n - \mathbb{I}_{\mathcal{R}_n}, \tag{13}$$

where I_{n-k} and $\mathcal{N}_k$ are independent.

5.1 Mean of I_n

Define, for all integer n, $x_n := \mathbb{E}[X_n]$ and $i_n := \mathbb{E}[I_n]$. From the Lemma 5.1, we deduce that

$$i_n = p\,(2\,i_{n-1} + x_{n-1}) + q\,i_{n-1} = (1+p)\,i_{n-1} + p\,x_{n-1}. \tag{14}$$

Proposition 5.1 *The mean i_n of I_n is given by*

$$\mathbb{E}[I_n] = p\ (1+p)^n \sum_{k=1}^{n} \mu_{k,\,p},$$

and

$$\lim_n \mathbb{E}\left[\frac{I_n}{n\,(1+p)^n}\right] = p\,\mu_p,$$

where $\mu_{k,\,p}$ and μ_p are given respectively in Theorem 3.1.

Proof: By iteration, we conclude that the solution of the recurrence (14) is

$$i_n = p\sum_{k=1}^{n-1}(1+p)^{n-k}x_k.$$

Since, for all k, $x_k = \mu_{k,\,p}(1+p)^k$ and $\lim_n\ \mu_{n,\,p} := \mu_p$, the claim of the lemma follows. □

5.2 Second moment of I_n

To obtain the variance of I_n we need to compute, at first, $b_n := \mathbb{E}[I_n X_n]$.

Lemma 5.2 *Using the notations* (1)*, we have*

$$b_n = (1+p)\,b_{n-1} + 2\,p\,x_{n-1}\,i_{n-1} + p\,q^{n-1}\,\left(i_{n-1} - \mathbb{E}\left[I_{n-1}\,\mathbb{I}_{\mathcal{R}_{n-1}}\right]\right).$$

Proof: The sequence X_n satisfies

$$X_n = \mathbb{I}\left(X^{(1)}_{n-1} + X^{(2)}_{n-1} + \mathbb{J}_{n-1} - \mathbb{G}_{n-1}\right) + (1-\mathbb{I})\,\tilde{X}^{(1)}_{n-1}, \tag{15}$$

where

$$\mathbb{J}_n = \mathbb{I}_{\{|T^{(1)}_n|=1\}}\,\mathbb{I}_{\{|T^{(2)}_n|\geq 2\}}, \qquad \mathbb{G}_n = \mathbb{I}_{\{|T^{(2)}_n|=1\}}\,\mathbb{I}_{\mathcal{R}^{(1)}_n}.$$

Multiplying (12) by (15) and recall that $\mathbb{I} := \mathbb{1}_{\{|T_1|=2\}}$, we deduce,

$$\begin{aligned}
X_n\,I_n = &\ \mathbb{I}\left(X^{(1)}_{n-1}\,I^{(1)}_{n-1} + X^{(2)}_{n-1}\,I^{(1)}_{n-1} + \mathbb{J}_{n-1}\,I^{(1)}_{n-1} - \mathbb{G}_{n-1}\,I^{(1)}_{n-1}\right)\\
&+ \mathbb{I}\left(X^{(1)}_{n-1}\,I^{(2)}_{n-1} + X^{(2)}_{n-1}\,I^{(2)}_{n-1} + \mathbb{J}_{n-1}\,I^{(2)}_{n-1} - \mathbb{G}_{n-1}\,I^{(2)}_{n-1}\right)\\
&+ \mathbb{I}\left(X^{(1)}_{n-1}\,X^{(2)}_{n-1} + X^{(2)}_{n-1}\,X^{(2)}_{n-1} + \mathbb{J}_{n-1}\,X^{(2)}_{n-1} - \mathbb{G}_{n-1}\,X^{(2)}_{n-1}\right)\\
&+ (1-\mathbb{I})\,\tilde{X}^{(1)}_{n-1}\,\tilde{I}^{(1)}_{n-1}.
\end{aligned}$$

On one hand, observe that, almost surely,

$$\begin{aligned}
\mathbb{J}_{n-1}\,X^{(2)}_{n-1} = I^{(1)}_{n-1}\,\mathbb{J}_{n-1} = \mathbb{G}_{n-1}\,I^{(2)}_{n-1} &= \left(I^{(2)}_{n-1}\,\mathbb{I}_{\{|T^{(2)}_{n-1}|=1\}}\right)\mathbb{I}_{\mathcal{R}^{(1)}_{n-1}} = 0,\\
\mathbb{G}_{n-1}\,X^{(2)}_{n-1} = \mathbb{G}_{n-1}\,I^{(2)}_{n-1} &= \left(I^{(2)}_{n-1}\,\mathbb{I}_{\{|T^{(2)}_{n-1}|=1\}}\right)\mathbb{I}_{\mathcal{R}^{(1)}_{n-1}} = 0.
\end{aligned}$$

We deduce, almost surely, that

$$\mathbb{J}_{n-1}\,I^{(1)}_{n-1} = \mathbb{G}_{n-1}\,I^{(2)}_{n-1} = \mathbb{J}_{n-1}\,X^{(2)}_{n-1} = \mathbb{G}_{n-1}\,X^{(2)}_{n-1} = 0.$$

On the other hand, from

$$\begin{aligned}
I^{(2)}_{n-1}\,\mathbb{J}_{n-1} &= \left(I^{(2)}_{n-1}\,\mathbb{I}_{\{|T^{(2)}_{n-1}|\geq 2\}}\right)\mathbb{I}_{\{|T^{(1)}_{n-1}|=1\}} = I^{(2)}_{n-1}\,\mathbb{I}_{\{|T^{(1)}_{n-1}|=1\}},\\
I^{(1)}_{n-1}\mathbb{G}_{n-1} &= \left(I^{(1)}_{n-1}\,\mathbb{I}_{\mathcal{R}^{(1)}_{n-1}}\right)\mathbb{I}_{\{|T^{(2)}_{n-1}|=1\}},
\end{aligned}$$

we obtain

$$\mathbb{E}\big[I^{(2)}_{n-1}\,\mathbb{J}_{n-1}\big] = \mathbb{E}\big[I^{(2)}_{n-1}\big]\,\mathbb{E}\big[\mathbb{I}_{\{|T^{(1)}_{n-1}|=1\}}\big] = i_{n-1}\,q^{n-1},$$
$$\mathbb{E}\big[I^{(1)}_{n-1}\mathbb{G}_{n-1}\big] = \mathbb{E}\big[I^{(1)}_{n-1}\,\mathbb{I}_{\mathcal{R}^{(1)}_{n-1}}\big]\,\mathbb{E}\big[\mathbb{I}_{\{|T^{(2)}_{n-1}|=1\}}\big] = q^{n-1}\,\mathbb{E}\big[I^{(1)}_{n-1}\,\mathbb{I}_{\mathcal{R}^{(1)}_{n-1}}\big].$$

Then

$$\begin{aligned} b_n &= p\Big(b_{n-1} + x_{n-1}i_{n-1} + \mathbb{E}\big[I^{(1)}_{n-1}\,\mathbb{J}_{n-1}\big] - \mathbb{E}\big[I^{(1)}_{n-1}\mathbb{G}_{n-1}\big]\Big)\\ &\quad + p\Big(x_{n-1}\,i_{n-1} + b_{n-1} + \mathbb{E}\big[I^{(2)}_{n-1}\,\mathbb{J}_{n-1}\big] - \mathbb{E}\big[I^{(2)}_{n-1}\mathbb{G}_{n-1}\big]\\ &\quad + x^2_{n-1} + \mathbb{E}\big[(X^{(2)}_{n-1})^2\big] - \mathbb{E}\big[\mathbb{G}_{n-1}\,X^{(2)}_{n-1}\big]\Big) + q\,b_{n-1}\\ &\quad = (1+p)\,b_{n-1} + 2\,p\,x_{n-1}\,i_{n-1} + p\mathbb{E}\big[I^{(1)}_{n-1}\,\mathbb{J}_{n-1}\big]\\ &\quad + p\Big(\mathbb{E}\big[I^{(2)}_{n-1}\,\mathbb{J}_{n-1}\big] - \mathbb{E}\big[\mathbb{G}_{n-1}\,X^{(2)}_{n-1}\big] - \mathbb{E}\big[I^{(2)}_{n-1}\mathbb{G}_{n-1}\big] - \mathbb{E}\big[I^{(1)}_{n-1}\mathbb{G}_{n-1}\big]\Big). \end{aligned}$$

So the assertion of the lemma follows. □

To obtain a closed form of b_n, we need to compute $\mathbb{E}\big[I_n\,\mathbb{I}_{\mathcal{R}_n}\big]$.

Lemma 5.3

$$\begin{aligned} \mathbb{E}\big[I_n\,\mathbb{I}_{\mathcal{R}_n}\big] &= q\,\mathbb{E}\big[I_{n-1}\,\mathbb{I}_{\mathcal{R}_{n-1}}\big] + p\,q^{n-1}\,\left(i_{n-1}+x_{n-1}\right) + p\,\mathbb{E}[I_{n-1}]\mathbb{P}(\mathcal{R}_{n-1})\\ &\quad + p\,\left(1-q^{n-1}\right)\,\mathbb{E}\big[I_{n-1}\,\mathbb{I}_{\mathcal{R}_{n-1}}\big] + p\,x_{n-1}\,\mathbb{P}(\mathcal{R}_{n-1}).\\ &= \left(1-p\,q^{n-1}\right)\,\mathbb{E}\big[I_{n-1}\,\mathbb{I}_{\mathcal{R}_{n-1}}\big] + \left(p\,\mathbb{P}(\mathcal{R}_{n-1}) + p\,q^{n-1}\right)\left(i_{n-1}+x_{n-1}\right)\\ &= \sum_{k=0}^{n-1}\prod_{j=k+1}^{n}\left(1-p\,q^{j-1}\right)\left(p\,\mathbb{P}(\mathcal{R}_k) + p\,q^k\right)\left(i_k+x_k\right). \end{aligned}$$

Proof: Since

$$\begin{aligned} \mathcal{R}_n &= \Big[\{|T_1|=1\}\cap\tilde{\mathcal{R}}^{(1)}_{n-1}\Big] \cup \Big[\{|T_1|=2\}\cap\mathcal{R}^{(1)}_{n-1}\cap\Big\{|T^{(2)}_{n-1}|\geq 2\Big\}\Big]\\ &\quad \cup\Big[\{|T_1|=2\}\cap\Big\{|T^{(1)}_{n-1}|=1\Big\}\cap\Big\{|T^{(2)}_{n-1}|\geq 2\Big\}\Big]. \end{aligned}$$

we have

$$\mathbb{I}_{\mathcal{R}_n} = (1-\mathbb{I})\,\mathbb{I}_{\tilde{\mathcal{R}}^{(1)}_{n-1}} + \mathbb{I}_{\{|T^{(1)}_{n-1}|=1\}}\,\mathbb{I}_{\{|T^{(2)}_{n-1}|\geq 2\}}\,\mathbb{I} + \mathbb{I}_{\mathcal{R}^{(1)}_{n-1}}\,\mathbb{I}_{\{|T^{(2)}_{n-1}|\geq 2\}}\,\mathbb{I},$$

Multiplying by I_n, we obtain

$$\begin{aligned} I_n\,\mathbb{I}_{\mathcal{R}_n} &= (1-\mathbb{I})\,\tilde{I}^{(1)}_{n-1}\,\mathbb{I}_{\tilde{\mathcal{R}}^{(1)}_{n-1}} + \mathbb{I}_{\{|T^{(1)}_{n-1}|=1\}}\,\mathbb{I}_{\{|T^{(2)}_{n-1}|\geq 2\}}\left(I^{(2)}_{n-1}+X^{(2)}_{n-1}\right)\mathbb{I}\\ &\quad + \left(I^{(2)}_{n-1}+I^{(1)}_{n-1}+X^{(2)}_{n-1}\right)\mathbb{I}_{\mathcal{R}^{(1)}_{n-1}}\,\mathbb{I}_{\{|T^{(2)}_{n-1}|\geq 2\}}\,\mathbb{I}. \end{aligned}$$

The result of the lemma is obtained by taking the expectation of this equation. □

Using Lemmas 5.2 and 5.3, we deduce

Proposition 5.2 *The second moment of I_n satisfies the following recursion*

$$\mathbb{E}[I_n^2] = (p+1)\ \mathbb{E}[I_{n-1}^2] + 2p\left(i_{n-1}^2 + i_{n-1}\,x_{n-1} + b_{n-1}\right) + p\mathbb{E}[X_{n-1}^2],$$

with as general solution

$$\mathbb{E}[I_n^2] = p\sum_{k=1}^{n-1}(p+1)^{n-k}\ \left(2(i_k^2 + i_k\,x_k + b_k) + \mathbb{E}[X_k^2]\right).$$

5.3 Convergence in Distribution of I_n

The aim of this section is to use the contraction method in Roesler and Rueschendorf (2001) or the multivariate contraction method in Neininger (2001) to state the limiting distribution of $\hat{I}_n := I_n/n\,(1+p)^n$.

To apply the contraction method, we set up some notation as follows: For a random variable X, we write $X \sim \lambda$ if the distribution of X is λ, i.e. the law $\mathcal{L}(X)$ of X is λ. The symbol $\|X\|_2 := (\mathbb{E}[|X|^2])^{1/2}$ denotes the usual L_2-norm of X. The Wasserstein-metric ℓ_2 is defined on the space of probability distributions with existing second moments by

$$\ell_2(X,\,Y) := \ell_2(\mathcal{L}(X),\,\mathcal{L}(Y)) := \ell_2(\lambda,\,\nu) := \inf\{\|X-Y\|_2\,:\,X\sim\lambda,\,Y\sim\nu\}.$$

By $\mathcal{M}_2$ the space of all probability distributions λ with mean $p\,\mu_p$ (as in Proposition 5.1) and finite second moment is denoted. The metric space $(\mathcal{M}_2, \ell_2)$ is complete and convergence in ℓ_2 is equivalent to convergence in distribution plus convergence of the second moments.

By equation (12) we have

$$\begin{aligned}\frac{I_n}{n(1+p)^n} \overset{d}{=} \frac{\mathbb{I}}{1+p}\cdot\frac{n-1}{n}\cdot\Big(\frac{I_{n-1}^{(1)}}{(n-1)(1+p)^{n-1}} + \frac{I_{n-1}^{(2)}}{(n-1)(1+p)^{n-1}}\Big)\\ + \frac{1-\mathbb{I}}{1+p}\cdot\frac{n-1}{n}\cdot\frac{\tilde{I}_{n-1}^{(1)}}{(n-1)(1+p)^{n-1}} + \mathbb{I}\cdot\frac{X_{n-1}}{n(1+p)^n}.\end{aligned} \tag{16}$$

Theorem 5.1 *Let I_n and X_n be the total path length and the number of protected nodes of an exponential recursive tree at age n, respectively. The normalized total path length $\hat{I}_n := I_n/n\,(1+p)^n$ satisfies the distributional recursion*

$$\hat{I}_n \overset{d}{=} \frac{\mathbb{I}}{1+p}\cdot\frac{n-1}{n}\cdot\left(\hat{I}_{n-1}^{(1)} + \hat{I}_{n-1}^{(2)}\right) + \frac{1-\mathbb{I}}{1+p}\cdot\frac{n-1}{n}\cdot\hat{I}_{n-1}^{(1)} + \mathbb{I}\cdot\frac{X_{n-1}}{n(1+p)^n}, \tag{17}$$

where $\hat{I}_n^{(1)}$, $(\hat{I}_n^{(2)}, X_n)$, $\hat{I}_n^{(1)}$ and $\mathbb{I}$ are independent, $\hat{I}_n^{(i)} \overset{d}{=} \hat{I}_n$ for $i = 1,\,2$, and $\mathbb{I}$ is a Bernoulli random variable with success probability p, then $\hat{I}_n \xrightarrow{\mathcal{D}} \hat{I}$, as $n\to\infty$, where the random variable $\hat{I}$ is the unique distributional fixed-point of

$$\hat{I} \overset{d}{=} \frac{\mathbb{I}}{1+p}\left(\hat{I}^{(1)} + \hat{I}^{(2)}\right) + \frac{1-\mathbb{I}}{1+p}\,\hat{I}^{(1)}, \tag{18}$$

with $\hat{I}^{(i)}$, $i = 1,\,2$, are independent copies of $\hat{I}$ and independent of $\mathbb{I}$.

Proof: The equation (17) is an immediate consequence of the equation (16), where we define $\hat{I}_n^{(3)} := \tilde{I}_n^{(1)}/n\,(1+p)^n$. Using Theorem 4.1 and Slutsky's theorem, we have $\frac{X_{n-1}}{n(1+p)^{n-1}} \xrightarrow{\mathcal{D}} 0$, and then convergence in probability $\frac{X_{n-1}}{n(1+p)^{n-1}} \xrightarrow{\mathcal{P}} 0$. That is,

$$\hat{I}_n \stackrel{d}{=} \frac{\mathbb{I}}{1+p} \cdot \frac{n-1}{n} \cdot \big(\hat{I}_{n-1}^{(1)} + \hat{I}_{n-1}^{(2)}\big) + \frac{1-\mathbb{I}}{1+p} \cdot \frac{n-1}{n} \cdot \hat{I}_{n-1}^{(1)} + o_p(1), \tag{19}$$

where $o_p(1)$ denotes a quantity tending to zero in probability. For purposes of convergence we can, and hence will, ignore the $o_p(1)$ term.

Consider the well-defined transformation $T : \mathcal{M}_2 \to \mathcal{M}_2$,

$$T(\lambda) := \mathcal{L}\Big(\frac{\mathbb{I}}{1+p}\big(\hat{I}_\lambda^{(1)} + \hat{I}_\lambda^{(2)}\big) + \frac{1-\mathbb{I}}{1+p}\hat{I}_\lambda^{(1)}\Big), \tag{20}$$

where $\hat{I}_\lambda^{(i)} := \hat{I}^{(i)}$, $i = 1,\, 2$, and $\mathbb{I}$ are independent and $\hat{I}^{(i)}$ have λ as distribution.

At a first step, we have to prove that the transformation T has a unique fixed point with respect to the ℓ_2-metric. Let $\lambda, \nu \in \mathcal{M}_2$ be given. By (20), we have

$$\begin{aligned}
T(\lambda) &= \mathcal{L}\Big(\frac{\mathbb{I}}{1+p}\big(\hat{I}_\lambda^{(1)} + \hat{I}_\lambda^{(2)}\big) + \frac{1-\mathbb{I}}{1+p}\hat{I}_\lambda^{(1)}\Big),\\
T(\nu) &= \mathcal{L}\Big(\frac{\mathbb{I}}{1+p}\big(\hat{I}_\nu^{(1)} + \hat{I}_\nu^{(2)}\big) + \frac{1-\mathbb{I}}{1+p}\hat{I}_\nu^{(1)}\Big),\\
\ell_2^2(T(\lambda),\, T(\nu)) &\le 2\,\frac{\mathbb{E}[\mathbb{I}^2]}{(1+p)^2}\,\mathbb{E}\big[\big|\hat{I}_\lambda - \hat{I}_\nu\big|\big]^2 + \frac{\mathbb{E}\left[(1-\mathbb{I})^2\right]}{(1+p)^2}\,\mathbb{E}\big[\big|\hat{I}_\lambda - \hat{I}_\nu\big|\big]^2\\
&= \frac{1}{1+p}\,\mathbb{E}\big[\big|\hat{I}_\lambda - \hat{I}_\nu\big|\big]^2.
\end{aligned}$$

Therefore, we have $\ell_2^2(T(\lambda),\, T(\nu)) \le \frac{1}{1+p}\,\ell_2^2(\lambda,\, \nu)$. Since $\frac{1}{1+p} < 1$, we deduce that T is a contraction with respect to the ℓ_2-metric. Thus, Banach's fixed point theorem provides existence and uniqueness of solutions of the fixed point equation $T(\lambda) = \lambda$. By (19), if $\hat{I}_n$ converges in distribution to some random variable $\hat{I}$, then $\hat{I}$ satisfies (18). Consequently, the distribution of $\hat{I}$ will be λ_0, the unique fixed point of T, i.e. $\mathcal{L}(\hat{I}) = T(\lambda_0) = \lambda_0$. Therefore, we have to prove $\lim_{n\to\infty} \ell_2(\hat{I}_n,\, \hat{I}) = 0$ to conclude $\hat{I}_n \xrightarrow{\mathcal{D}} \hat{I}$.

Since $\hat{I}_n^{(i)} \stackrel{d}{=} \hat{I}_n$, $\hat{I}^{(i)} \stackrel{d}{=} \hat{I}$, for $i = 1,\, 2,\, 3$, we deduce

$$\begin{aligned}
\lim_{n\to\infty} \ell_2^2(\hat{I}_n,\, \hat{I}) &\le \frac{p}{(1+p)^2}\lim_{n\to\infty}\Big(\Big\|\frac{n-1}{n}\,\hat{I}_{n-1}^{(1)} - \hat{I}^{(1)}\Big\|_2^2 + \Big\|\frac{n-1}{n}\,\hat{I}_{n-1}^{(2)} - \hat{I}^{(2)}\Big\|_2^2\Big)\\
&\quad + \frac{1-p}{(1+p)^2}\lim_{n\to\infty}\Big\|\frac{n-1}{n}\,\hat{I}_{n-1}^{(3)} - \hat{I}^{(3)}\Big\|_2^2\\
&\quad \le \frac{1}{1+p}\lim_{n\to\infty}\big\|\hat{I}_n - \hat{I}\big\|_2.
\end{aligned}$$

Therefore, we have

$$\lim_{n\to\infty} \ell_2^2(\hat{I}_n,\, \hat{I}) \le \frac{1}{1+p}\lim_{n\to\infty} \ell_2^2(\hat{I}_n,\, \hat{I}).$$

This is true only if $\lim_{n\to\infty} \ell_2(\hat{I}_n,\, \hat{I}) = 0$. □

References

R. Aguech, S. Bose, H. M. Mahmoud, and Y. Zhang. Some properties of exponential trees. *International Journal of Computer Mathematics: Computer Systems Theory*, 7:16–32, 2022.

P. Billingsley. *Probability and Measure*. Anniversary Ed., Wiley, Hoboken, New Jersey, 2012.

L. Devroye and S. Janson. Protected nodes and fringe subtrees in some random trees. *Electronic Communications in Probability*, 19:1–10, 2014.

Y. Feng and H. M. Mahmoud. Profile of random exponential binary trees. *Methodology and Computing in Applied Probability*, 20:575–587, 2018.

M. Fuchs, C. K. Lee, and G. R. Yu. On 2-protected nodes in random digital trees. *Theoretical Computer Science*, 622:111–122, 2016.

M. Javanian, R. Imany Nabiyyi, and J. Toofanpour. Normal limit law for protected node profile of random recursive trees. *Theory of Probability and Its Applications*, 67:563–578, 2022.

H. M. Mahmoud. Profile of random exponential recursive trees. *Methodology and Computing in Applied Probability*, 24:259–275, 2022.

H. M. Mahmoud and M. D. Ward. Asymptotic properties of protected nodes in random recursive trees. *Journal of Applied Probability*, 52:290–297, 2015.

R. Neininger. On a multivariate contraction method for random recursive structures with applications to quicksort. *Random Structures and Algorithms*, 19:498–524, 2001.

U. Roesler and L. Rueschendorf. The contraction method for recursive algorithms. *Algorithmica*, 29: 3–33, 2001.

R. T. Smythe and H. M. Mahmoud. A survey of recursive trees. *Theory of Probability and Mathematical Statistics*, 51:1–27, 1995.