



**HAL**  
open science

## On the protected nodes in exponential recursive trees

Mehri Javanian, Rafik Aguech

► **To cite this version:**

Mehri Javanian, Rafik Aguech. On the protected nodes in exponential recursive trees. 2022. hal-03908994v1

**HAL Id: hal-03908994**

**<https://hal.science/hal-03908994v1>**

Preprint submitted on 21 Dec 2022 (v1), last revised 21 Nov 2023 (v5)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the protected nodes in exponential recursive trees

Mehri Javanian\*      Rafik Aguech†

August 3, 2022

## Abstract

The exponential recursive trees model several kinds of networks. At each step of growing of these trees, each node 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.

## 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 [4] and [10] 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  $n$ th 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 [13].

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 person to person). In [10], a fast-growing analogue of recursive

---

\*Department of Statistics, Faculty of Sciences, University of Zanjan, Iran. E-mail: javanian@znu.ac.ir

†Department of mathematics , University of Monastir, Tunisia. E-mail: rafik.aguech@ipeit.rnu.tn

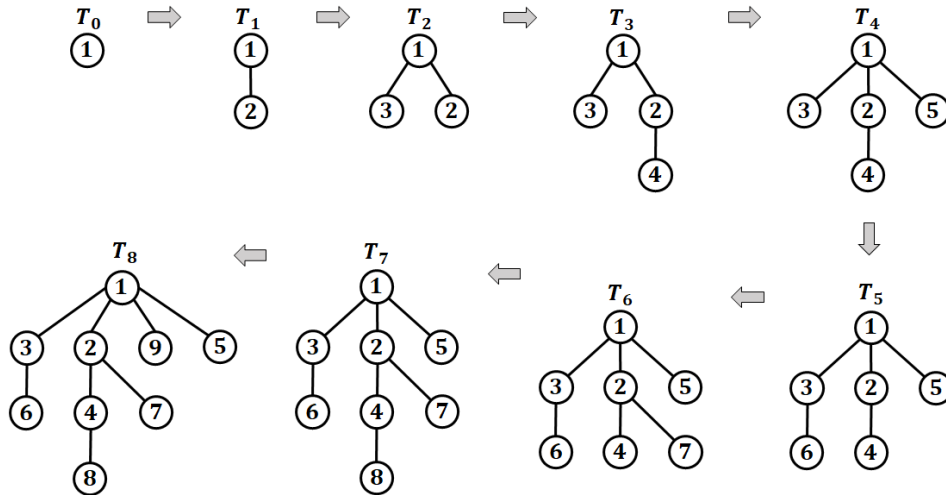


Figure 1: The evolution of a recursive tree in 8 steps.

trees has been defined as follows: Initially,  $T_0$  is a tree of a single root node. For  $n \geq 1$ , at the  $n$ th 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  $n$ th 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 7th 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. [3, 6, 8, 9]. 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  $T_n$ ).

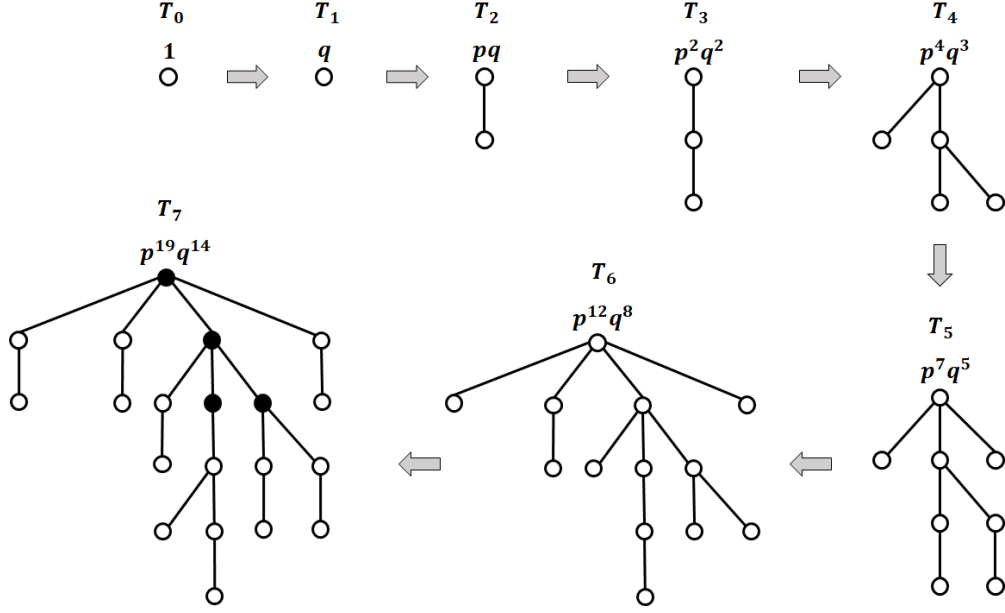


Figure 2: **(a)** The evolution of an exponential recursive tree (ERT) of index  $p$  in 7 steps. The probability of an ERT appears above it. **(b)** In  $T_7$ , protected nodes are shown in black.

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, 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

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

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

**Lemma 1.** *By the above setting, for  $n \geq 2$ , the probability of  $\mathcal{R}_n$  is*

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

*Proof.* We can observe that

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

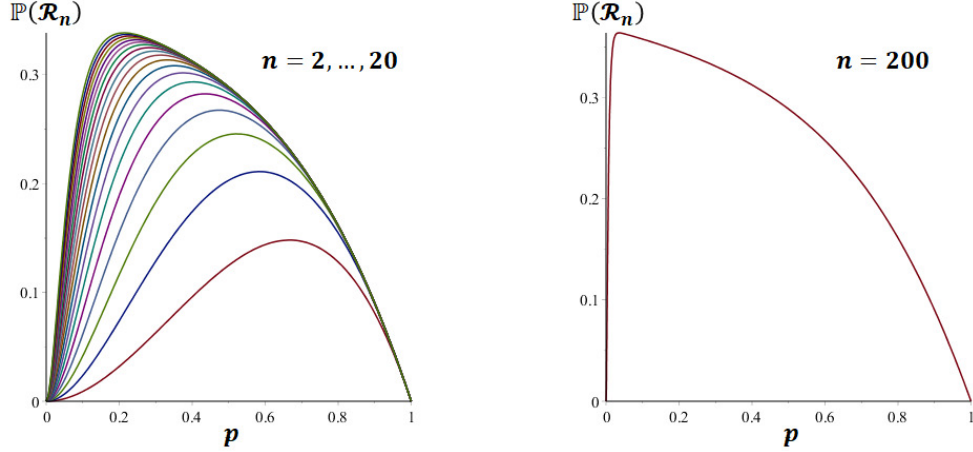


Figure 3: Graphs of  $\mathbb{P}(\mathcal{R}_n)$  versus  $p$ , for  $n = 2, \dots, 20$  and  $n = 200$ . For each  $p$ ,  $\mathbb{P}(\mathcal{R}_n)$  is increasing in  $n$ . E.g., the bottom and top curves are the graphs of  $\mathbb{P}(\mathcal{R}_2)$  and  $\mathbb{P}(\mathcal{R}_{20})$ .

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  $\mathbb{P}(\mathcal{R}_n)$ .  $\square$

**Lemma 2.** *By the above setting, we have the distributional equation*

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

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).  $\square$

**Lemma 3.** For  $m \geq 1$ , we have the following recurrences for  $\mathbb{E}[X_n^m]$  and  $\mathbb{E}[X_n^m \mathbb{1}_{\mathcal{R}_n}]$ :

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

*Proof.* Raise both sides of (3) to the  $m$ th 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

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

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

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

Therefore, the equation (4) can be simplified as

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

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 same distribution in the subtrees and

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

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

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

Taking the expectation of this equation, it implies the second recurrence of the Lemma.  $\square$

### 3 The Expectation and Variance

In this section, we obtain the expectation and variance of  $X_n$ , using Lemma 1 and Lemma 3.

**Theorem 4.** *Let  $X_n$  be the number of protected nodes in an exponential recursive tree of age  $n$  and index  $p$ . For  $q := 1 - p$ ,*

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

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

*Proof.* In Lemma 3, set  $m = 1$  to obtain the recurrence

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

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

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

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

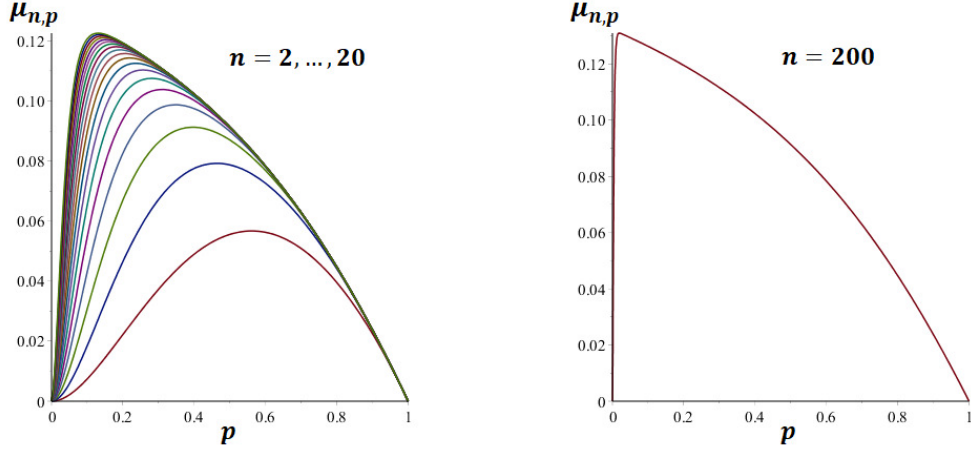


Figure 4: Graphs of  $\mu_{n,p}$  versus  $p \in (0, 1)$ , for  $n = 2, \dots, 20$  and  $n = 200$ .

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

**Lemma 5.** *Let  $Y_n := X_n \mathbb{1}_{\mathcal{R}_n}$ , then*

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

where  $\alpha_p := \lim_{n \rightarrow \infty} \alpha_{n,p}$  (see Figure 5) and the functions  $\mathbb{P}(\mathcal{R}_n)$  and  $\mu_{n,p}$  are obtained in Lemma 1 and Theorem 4, respectively.

*Proof.* In Lemma 3, set  $m = 1$  to find the recurrence

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

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

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

This implies the claim. □



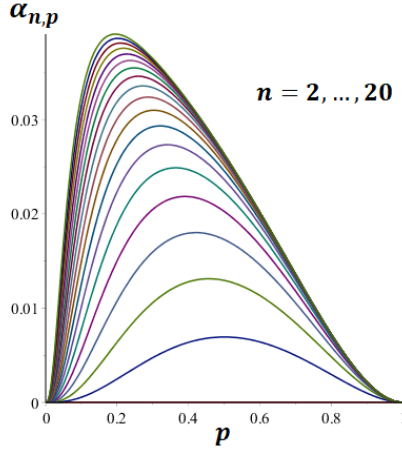


Figure 5: Graphs of  $\alpha_{n,p}$  versus  $p \in (0, 1)$ , for  $n = 2, \dots, 20$ .

**Theorem 6.** *By the above definitions for  $\mathbb{P}(\mathcal{R}_n)$ ,  $\mu_{n,p}$  and  $\alpha_{n,p}$ , we have*

$$\begin{aligned} \mathbb{E}[X_n^2] &= \sum_{j=1}^{n-1} \left( 2p\mu_{j,p}^2(p+1)^{-n+j-1} + 2pq^j(\mu_{j,p} - \alpha_{j,p})(p+1)^{-n-1} \right. \\ &\quad \left. + pq^j(1 - q^j - \mathbb{P}(\mathcal{R}_j))(p+1)^{-n-j-1} \right) (p+1)^{2n} =: \beta_{n,p}(p+1)^{2n}. \end{aligned}$$

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

*Proof.* In Lemma 3, set  $m = 2$  to find the recurrence

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

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

$$\begin{aligned} \mathbb{E}[X_n^2] &= \sum_{j=0}^{n-2} (p+1)^j \left( 2p\mathbb{E}^2[X_{n-j-1}] + 2pq^{n-j-1}\mathbb{E}[X_{n-j-1}] \right. \\ &\quad \left. - 2pq^{n-j-1}\mathbb{E}[Y_{n-j-1}] + pq^{n-j-1}(1 - q^{n-j-1} + \mathbb{P}(\mathcal{R}_{n-j-1})) \right). \end{aligned} \quad (9)$$

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

$$\begin{aligned} \mathbb{E}[X_n^2] &= \sum_{j=0}^{n-2} \left( 2p\mu_{n-j-1,p}^2(p+1)^{-j-2} + 2pq^{n-j-1}(\mu_{n-j-1,p} - \alpha_{n-j-1,p})(p+1)^{-n-1} \right. \\ &\quad \left. + pq^{n-j-1}(p+1)^{-2n+j}(1 - q^{n-j-1} - \mathbb{P}(\mathcal{R}_{n-j-1})) \right) (p+1)^{2n}. \end{aligned}$$

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

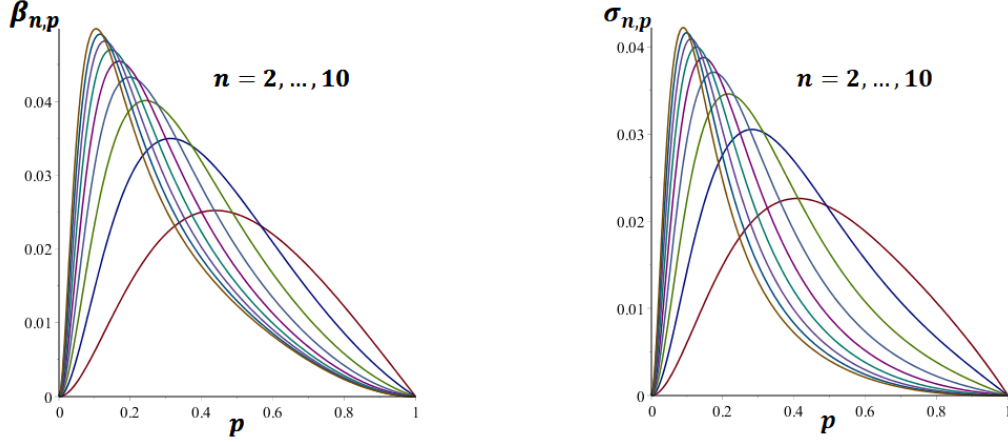


Figure 6: Graphs of  $\beta_{n,p}$  and  $\sigma_{n,p}$  versus  $p \in (0, 1)$ , for  $n = 2, \dots, 10$ .

**Corollary 7.** *By the above definitions for  $\mu_{n,p}$  and  $\beta_{n,p}$ , we have*

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

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

*Proof.* Using  $\text{Var}[X_n] = \mathbb{E}[X_n^2] - \mathbb{E}^2[X_n]$  and Theorem 6, the proof is straightforward.  $\square$

## 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 8.** *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

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

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

*Proof.* Scale the two recurrences in Lemma 3 by  $(p+1)^{nm}$  to get

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

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

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

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

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

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

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 [4] and in page 12 of [1], 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 [2],  $\frac{X_n}{(p+1)^n}$  converges in distribution to a unique limit  $X_*$ .  $\square$

## 5 The Total path length

Let denote by  $I_n$  be the total path length of protected nodes in an exponential recursive tree after  $n$  steps: is the sum of depths of all protected nodes. At step one if the root recruits

a new child then the total path length at step  $n$  is the total path length of the root at step  $n - 1$  (for the rest of steps) plus the total path length of the sub-tree with root the first child of the root of the tree  $T_n$  and we have to add to all protected nodes of this sub-tree 1 ( the distance between the first child and the root) , but if the root does not recruits a child at step 1, then the total path length at step  $n$  will be the same as the total path length at a step  $n - 1$ . Let  $I_n$  be the total path length of protected nodes in an exponential recursive tree at time  $n$ , the following lemma holds

**Lemma 9.**

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

where  $I_k^{(1)}$ ,  $I_k^{(2)}$  and  $\tilde{I}_k^{(1)}$  are independent copies of  $I_k$  and  $\tilde{X}_k$  is the number of protected nodes of an exponential recursive tree at time  $k$ .

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

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

**Lemma 10.** *The total path length  $I_n$  satisfies, almost surely,*

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

where  $I_{n-k}$  is the total path length of the sub-tree with root the  $k^{\text{th}}$  child of the root of hole tree.

## 5.1 Mean of $I_n$

Denote, for all integer  $n$ , by and  $x_n = \mathbb{E}[X_n]$ .

**Lemma 11.** *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  $\mu_p$  are given respectively in Theorem 4

*Proof.* Denote, for all integer  $n$ , by  $i_n = \mathbb{E}[I_n]$ , from the lemma (9) we deduce that

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

By iteration we conclude that

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

Since, for all  $n$ ,  $x_n = \mu_{n,p} (1 + p)^n$  and  $\lim_n \mu_{n,p} := \mu_p$  exists, the claimed of the lemma follows.  $\square$

## 5.2 Second moment of $I_n$

To obtain the variance of  $I_n$  we need to compute  $b_n := \mathbb{E}[I_n X_n]$ . Using the notations (1), we have

**Lemma 12.**

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

*Proof.* The sequence  $X_n$  satisfies

$$X_n = \mathbb{I} \left( X_{n-1}^{(1)} + X_{n-1}^{(2)} + \mathbb{J}_{n-1} - \mathbb{G}_{n-1} \right) + (1 - \mathbb{I}) \tilde{X}_{n-1}^{(1)}, \quad (13)$$

where

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

Then multiplying (12) by (13) and recall that we deduce that  $\mathbb{I} := \mathbb{1}_{\{|T_1|=2\}}$ , we deduce that,

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

On one hand, observe that, almost surely, we have

$$\begin{aligned} \mathbb{J}_{n-1} X_{n-1}^{(2)} &= I_{n-1}^{(1)} \mathbb{J}_{n-1} = \mathbb{G}_{n-1} I_{n-1}^{(2)} = (I_{n-1}^{(2)} \mathbb{I}_{\{|T_{n-1}^{(2)}|=1\}}) \mathbb{I}_{\mathcal{R}_{n-1}^{(1)}} = 0, \\ \mathbb{G}_{n-1} X_{n-1}^{(2)} &= \mathbb{G}_{n-1} I_{n-1}^{(2)} = (I_{n-1}^{(2)} \mathbb{I}_{\{|T_{n-1}^{(2)}|=1\}}) \mathbb{I}_{\mathcal{R}_{n-1}^{(1)}} = 0 \end{aligned}$$

we deduce, almost surely, that

$$\mathbb{J}_{n-1} I_{n-1}^{(1)} = \mathbb{G}_{n-1} I_{n-1}^{(2)} = \mathbb{J}_{n-1} X_{n-1}^{(2)} = \mathbb{G}_{n-1} X_{n-1}^{(2)} = 0.$$

On the other hand, from

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

we obtain

$$\begin{aligned} \mathbb{E}[I_{n-1}^{(2)} \mathbb{J}_{n-1}] &= \mathbb{E}[I_{n-1}^{(2)}] \mathbb{E}[\mathbb{I}_{\{|T_{n-1}^{(1)}|=1\}}] = i_{n-1} q^{n-1} \\ \mathbb{E}[I_{n-1}^{(1)} \mathbb{G}_{n-1}] &= \mathbb{E}[I_{n-1}^{(1)} \mathbb{I}_{\mathcal{R}_{n-1}^{(1)}}] \mathbb{E}[\mathbb{I}_{\{|T_{n-1}^{(2)}|=1\}}] = q^{n-1} \mathbb{E}[I_{n-1}^{(1)} \mathbb{I}_{\mathcal{R}_{n-1}^{(1)}}]. \end{aligned}$$

$$\begin{aligned} b_n &= p \left( b_{n-1} + x_{n-1} i_{n-1} + \mathbb{E}[I_{n-1}^{(1)} \mathbb{J}_{n-1}] - \mathbb{E}[I_{n-1}^{(1)} \mathbb{G}_{n-1}] \right) \\ &\quad + p \left( x_{n-1} i_{n-1} + b_{n-1} + \mathbb{E}[I_{n-1}^{(2)} \mathbb{J}_{n-1}] - \mathbb{E}[I_{n-1}^{(2)} \mathbb{G}_{n-1}] \right) \\ &\quad + x_{n-1}^2 + \mathbb{E}[(X_{n-1}^{(2)})^2] - \mathbb{E}[\mathbb{G}_{n-1} X_{n-1}^{(2)}] + q b_{n-1} \\ &= (1+p) b_{n-1} + 2p x_{n-1} i_{n-1} + p \mathbb{E}[I_{n-1}^{(1)} \mathbb{J}_{n-1}] \\ &\quad + p \left( \mathbb{E}[I_{n-1}^{(2)} \mathbb{J}_{n-1}] - \mathbb{E}[\mathbb{G}_{n-1} X_{n-1}^{(2)}] - \mathbb{E}[I_{n-1}^{(2)} \mathbb{G}_{n-1}] - \mathbb{E}[I_{n-1}^{(1)} \mathbb{G}_{n-1}] \right). \end{aligned}$$

□

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

**Lemma 13.**

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

*Proof.* Since

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

we have

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

Multiplying by  $I_n$ , we obtain

$$\begin{aligned} I_n \mathbb{I}_{\mathcal{R}_n} &= (1 - \mathbb{I}) \tilde{I}_{n-1}^{(1)} \mathbb{I}_{\tilde{\mathcal{R}}_{n-1}^{(1)}} + \mathbb{I}_{\{|T_{n-1}^{(1)}|=1\}} \mathbb{I}_{\{|T_{n-1}^{(2)}|\geq 2\}} (I_{n-1}^{(2)} + X_{n-1}^{(2)}) \mathbb{I} \\ &\quad + (I_{n-1}^{(2)} + I_{n-1}^{(1)} + X_{n-1}^{(2)}) \mathbb{I}_{\mathcal{R}_{n-1}^{(1)}} \mathbb{I}_{\{|T_{n-1}^{(2)}|\geq 2\}} \mathbb{I}. \end{aligned}$$

The result of the lemma is obtained by taking the expectation of this equation.  $\square$

Using lemmas (12) and (13), we deduce

**Proposition 14.** *The second moment of  $I_n$  satisfies the following recursion*

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

where the general solution is

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

### 5.3 Convergence in Distribution

The aim of this section is to use the contraction method by Rosler and Neininger to state the limiting distribution of  $\hat{I}_n := I_n/n(1+p)^n$ . By equation (12) we have

$$\frac{I_n}{n(1+p)^n} \stackrel{d}{=} \mathbb{I} \left( \frac{I_{n-1}^{(1)}}{n(1+p)^n} + \frac{I_{n-1}^{(2)}}{(n)(1+p)^n} + \frac{X_{n-1}^{(2)}}{n(1+p)^n} \right) + (1 - \mathbb{I}) \frac{\tilde{I}_{n-1}^{(1)}}{n(1+p)^n} \quad (14)$$

All the conditions of the contraction method are satisfied (Theorem 3 of [12], page 8), then we deduce that

**Theorem 15.** *The normalized total path length:  $\hat{I}_n = I_n/n(1+p)^n$  converges in distribution to some random variable  $\hat{I}$  with distribution the unique solution the following equation in distribution*

$$\hat{I} \stackrel{d}{=} \frac{\text{Ber}(p)}{1+p} (\hat{I}^{(1)} + \hat{I}^{(2)}) + \frac{1 - \text{Ber}(p)}{1+p} \hat{I}^{(3)}, \quad (15)$$

where  $\hat{I}^{(i)}$ ,  $i = 1, 2, 3$ , are independent copies of  $\hat{I}$  and independent of  $\text{Ber}(p)$ .

*Proof.* The proof is based on the Lemma 3.1 page 502 of the paper [11].

Consider the transformation  $T : \mathcal{M}_2 \rightarrow \mathcal{M}_2$  defined by

$$T(\mu) \stackrel{d}{=} \frac{\text{Ber}(p)}{1+p} (\hat{I}^{(1)} + \hat{I}^{(2)}) + \frac{(1 - \text{Ber}(p))}{1+p} \hat{I}^{(3)}, \quad (16)$$

where  $\hat{I}^{(i)}$ ,  $i = 1, 2, 3$  and  $\text{Ber}(p)$  are independent and  $\hat{I}^{(i)}$  have  $\mu$  as distribution.

1. At a first step we have to prove that with respect to the metric  $\mathcal{L}_2^2$  defined on  $\mathcal{M}_2$  by

$$\mathcal{L}_2^2(\mu, \nu) := \min \{ \mathbb{E}[|X - Y|^2], \text{ where } X, Y \text{ have } \mu \text{ and } \nu \text{ as distribution, respectively} \},$$

the transformation have a unique fixed point. In fact let  $\mu$  and  $\nu$  be two measures of  $\mathcal{M}_2$

$$\begin{aligned} T(\mu) &\stackrel{d}{=} \frac{\text{Ber}(p)}{1+p} (\hat{I}_\mu^{(1)} + \hat{I}_\mu^{(2)}) + \frac{(1 - \text{Ber}(p))}{1+p} \hat{I}_\mu^{(3)}, \\ T(\nu) &\stackrel{d}{=} \frac{\text{Ber}(p)}{1+p} (\hat{I}_\nu^{(1)} + \hat{I}_\nu^{(2)}) + \frac{(1 - \text{Ber}(p))}{1+p} \hat{I}_\nu^{(3)}, \end{aligned}$$

$$\begin{aligned} \mathcal{L}_2^2(T(\mu), T(\nu)) &= \mathbb{E}[|T(\mu) - T(\nu)|^2] \\ &= 2 \frac{\mathbb{E}[(\text{Ber}(p))^2]}{(1+p)^2} \mathbb{E}[\hat{I}_\mu - \hat{I}_\nu]^2 + \frac{\mathbb{E}[(1 - \text{Ber}(p))^2]}{(1+p)^2} \mathbb{E}[\hat{I}_\mu - \hat{I}_\nu]^2 \\ &= \frac{1}{1+p} \mathbb{E}[\hat{I}_\mu - \hat{I}_\nu]^2 \\ &\leq \frac{1}{1+p} \mathcal{L}_2^2(\mu, \nu). \end{aligned}$$

Since  $1/(1+p) < 1$  we deduce that  $T$  is a contracted transform and then it has only one fixed point.

2. Heuristically speaking, it is clear that if  $\hat{I}_n$  converges in distribution, the limit will be the unique fixed point of the transformation  $T$ . We have to prove that

$$\lim_n \mathcal{L}_2^2(\hat{I}, \hat{I}_n) = 0,$$



to conclude that  $\hat{I}_n$  converges in distribution and in  $L_2$  to  $\hat{I}$ . We have

$$\begin{aligned}
\mathcal{L}_2^2(\hat{I}, \hat{I}_n) &\leq \frac{\mathbb{E}[\text{Ber}(p)]}{(1+p)^2} \left| \left( \hat{I}^{(1)} - \hat{I}_{n-1}^{(1)} \frac{n-1}{n} \right) + \left( \hat{I}^{(2)} - \hat{I}_{n-1}^{(2)} \frac{n-1}{n} \right) \right|_2^2 \\
&\quad + \frac{\mathbb{E}[1 - \text{Ber}(p)]}{(1+p)^2} \left| \hat{I}^{(3)} - \hat{I}_{n-1}^{(3)} \frac{n-1}{n} \right|_2^2, \\
&\quad + 2 \frac{\mathbb{E}[\text{Ber}(p)]}{(1+p)^2} \mathbb{E} \left[ \frac{\tilde{X}_{n-1}}{n(1+p)^{n-1}} \left( \hat{I}^{(2)} - \hat{I}_{n-1}^{(2)} \frac{n-1}{n} \right) \right] \\
&\quad + 2 \frac{\mathbb{E}[\text{Ber}(p)]}{(1+p)^2} \mathbb{E} \left[ \frac{\tilde{X}_{n-1}}{n(1+p)^{n-1}} \left( \hat{I}^{(1)} - \hat{I}_{n-1}^{(1)} \frac{n-1}{n} \right) \right] \\
&\quad + \frac{\mathbb{E}[\text{Ber}(p)]}{(1+p)^2} \left| \frac{\tilde{X}_{n-1}^2}{n(1+p)^{n-1}} \right|_2^2 \\
&= \frac{p}{(1+p)^2} \left( \left| \hat{I}^{(1)} - \hat{I}_{n-1}^{(1)} \frac{n-1}{n} \right|_2^2 + \left| \hat{I}^{(2)} - \hat{I}_{n-1}^{(2)} \frac{n-1}{n} \right|_2^2 \right) \\
&\quad + \frac{q}{(1+p)^2} \left| \hat{I}^{(3)} - \hat{I}_{n-1}^{(3)} \frac{n-1}{n} \right|_2^2, \\
&\quad + 2 \frac{p}{(1+p)^2} \mathbb{E} \left[ \frac{\tilde{X}_{n-1}}{n(1+p)^{n-1}} \left( \hat{I}^{(2)} - \hat{I}_{n-1}^{(2)} \frac{n-1}{n} \right) \right] \\
&\quad + 2 \frac{p}{(1+p)^2} \mathbb{E} \left[ \frac{\tilde{X}_{n-1}}{n(1+p)^{n-1}} \left( \hat{I}^{(1)} - \hat{I}_{n-1}^{(1)} \frac{n-1}{n} \right) \right] \\
&\quad + \frac{\mathbb{E}[\text{Ber}(p)]}{(1+p)^2} \left| \frac{\tilde{X}_{n-1}^2}{n(1+p)^{n-1}} \right|_2^2.
\end{aligned}$$

Then we deduce

$$\lim_n \mathcal{L}_2^2(\hat{I}, \hat{I}_n) \leq \frac{1}{1+p} \lim_n \lim_n \mathcal{L}_2^2(\hat{I}, \hat{I}_{n-1}),$$

this is true only if  $\lim_n \mathcal{L}_2^2(\hat{I}, \hat{I}_n) = 0$ .

□

## 6 Protected Node Profile

For all  $0 \leq k \leq n$  let  $P_{n,k}$  (respectively  $X_{n,k}$ ) be the number of protected nodes (be the number of nodes at step  $n$  at level  $k$ ) It is obvious that

$$P_{n+1,k} = P_{n,k} - \sum_{l=1}^{P_{n,k}} \text{Ber}^{(l)}(p) + \sum_{i=1}^{X_{n,k}} \mathbb{I}_{\tilde{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1}}$$

where  $\bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1}$  is the event that the node number  $i$  (from left to right) at level  $k$  is non protected at step  $n$  but becomes protected at step  $(n+1)$ . The the sum  $\sum_{i=1}^{X_{n,k}} \mathbb{I}_{\bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1}}$  represents the number of nodes that become protected only at step  $n$ ,

**Theorem 16.** *The expected value  $p_{n,k} := \mathbb{E}[P_{n,k}]$  satisfies, for all  $k < n$ ,*

$$p_{n,k} = (1-p)^n + p \sum_{i=2}^k q^{i-2} \sum_{l=k}^{n-i} \binom{l-1}{k-1} \binom{n-i+1}{k} p^{2k} q^{l-k} (1 - q^{n-i-l-1}).$$

*Proof.*

$$\begin{aligned} p_{n+1,k} &= p_{n,k} - p p_{n,k} + \mathbb{E} \left[ X_{n,k} \mathbb{E} \left[ \mathbb{I}_{\bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1}} \middle| X_{n,k} \right] \right] \\ &= (1-p) p_{n,k} + \mathbb{E} \left[ X_{n,k} \mathbb{P} \left( \bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1} \middle| X_{n,k} \right) \right]. \end{aligned}$$

For all  $l \geq k$ , let  $B_i^{(l)}$  be the event that the birth date of node  $i$  is at step  $l$ , then we have

$$\bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1} = \cup_{l=k}^{n-1} B_i^{(l)} \cap \bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1}.$$

We obtain

$$\mathbb{P} \left( \bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1} \middle| X_{n,k} \right) = \sum_{l=k}^{n-1} \mathbb{P} \left( \bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1} \middle| B_i^{(l)}, X_{n,k} \right) \mathbb{P} \left( B_i^{(l)} \middle| X_{n,k} \right).$$

Since the node  $i$  to be protected, all his first children must recruits other children at times between  $l+1$  and  $n$ , then denote, for all  $l+1 \leq s \leq n$ , by  $A_{n,k,l,s}$  the event defined by: the node  $i$  which have  $l$  as birth date recruits at a time  $s$  and the recruited node recruits only at time  $(n+1)$ .

We conclude

$$\begin{aligned}
\mathbb{P} \left[ \bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1} \mid X_{n,k} \right] &= \sum_{l=k}^{n-1} \mathbb{P} \left[ \bar{\mathcal{R}}_{i,k}^n \cap \mathcal{R}_{i,k}^{n+1} \mid B_i^{(l)}, X_{n,k} \right] \mathbb{P} \left[ B_i^{(l)} \mid X_{n,k} \right] \\
&= \sum_{l=k}^{n-1} \mathbb{P} \left[ B_i^{(l)} \mid X_{n,k} \right] \sum_{s=l+1}^n \mathbb{P} \left[ A_{n,k,l,s} \mid B_i^{(l)}, X_{n,k} \right] \\
&= \sum_{l=k}^{n-1} \mathbb{P} \left[ B_i^{(l)} \mid X_{n,k} \right] \sum_{s=l+1}^n p q^{n-s} p \\
&= \sum_{l=k}^{n-1} p \binom{l-1}{k-1} p^{k-1} q^{l-1-(k-1)} \sum_{s=l+1}^n p q^{n-s} p \\
&= \sum_{l=k}^{n-1} p^3 \binom{l-1}{k-1} p^{k-1} q^{l-k} \frac{1 - q^{n-l-1}}{p} \\
&= \sum_{l=k}^{n-1} p^2 \binom{l-1}{k-1} p^{k-1} q^{l-k} (1 - q^{n-l-1}).
\end{aligned}$$

Finally we conclude that

$$p_{n+1,k} = (1-p)p_{n,k} + \sum_{l=k}^{n-1} p^2 \binom{l-1}{k-1} p^{k-1} q^{l-k} (1 - q^{n-l-1}) \mathbb{E}[X_{n,k}].$$

Since, it is well known that  $\mathbb{E}[X_{n,k}] = p^k, \binom{n}{k}$  we deduce

$$p_{n+1,k} = (1-p)p_{n,k} + \sum_{l=k}^{n-1} p^2 \binom{l-1}{k-1} \binom{n}{k} p^{2k-1} q^{l-k} (1 - q^{n-l-1}),$$

with starting condition  $p_{k+1,k} = p^{k+1}$ . □

## References

- [1] R. Aguech, S. Bose, H. M. Mahmoud and Y. Zhang, *Some properties of exponential trees*, Methodol Comput Appl Probab (2021).
- [2] P. Billingsley, *Probability and Measure*, Anniversary Ed., Wiley, Hoboken, New Jersey, 2012.
- [3] L. Devroye and S. Janson, *Protected nodes and fringe subtrees in some random trees*, Electron. Commun. Probab. **19** (2014), 1–10.

- [4] Y. Feng and H. M. Mahmoud, *Profile of random exponential binary trees*, Methodol Comput. Appl. Probab. **20** (2018), 575–587. <https://doi.org/10.1007/s11009-020-09831-9>
- [5] A. Frieze and M. Karoński, *Introduction to Random Graphs*, 2nd Ed. Cambridge University Press, Cambridge, UK, 2016.
- [6] M. Fuchs, C.K. Lee, and G.R. Yu, *On 2-protected nodes in random digital trees*, Theor. Comput. Sci. **622** (2016), 111–122.
- [7] M. Hofri and H. M. Mahmoud, *Algorithmics of Nonuniformity: Tools and Paradigms*, CRC Press, Boca Raton, Florida, 2018.
- [8] M. Javarian, R. Imany Nabiyi and J. Toofanpour, *Protected node profile in recursive trees*, Theory of Probability and Its Applications, accepted, 2021.
- [9] H. M. Mahmoud and M. D. Ward, *Asymptotic properties of protected nodes in random recursive trees*, J. Appl. Probab. **52** (2015), 290–297.
- [10] H. M. Mahmoud, *Profile of random exponential recursive trees*, Methodol Comput. Appl. Probab. (2021), 10.1007/s11009-020-09831-9.
- [11] R. Neininger, *On a multivariate contraction method for random recursive structures with applications to quicksort*, Random Structures and Algorithms. **19** (2001), 498–524.
- [12] U. Roesler and L. Rueschendorf, *The contraction method for recursive algorithms*, Algorithmica **29** (2001), 3–33.
- [13] R. T. Smythe and H. M. Mahmoud, *A survey of recursive trees*, Theory Probab. Math. Stat. **51** (1995), 1–27.