Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 53, 1–22.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP318

k-cut on paths and some trees*

Xing Shi Cai† Cecilia Holmgren† Luc Devroye‡ Fiona Skerman§

Abstract

We define the (random) k-cut number of a rooted graph to model the difficulty of the destruction of a resilient network. The process is as the cut model of Meir and Moon [21] except now a node must be cut k times before it is destroyed. The first order terms of the expectation and variance of $\mathcal{X}_n$, the k-cut number of a path of length n, are proved. We also show that $\mathcal{X}_n$, after rescaling, converges in distribution to a limit $\mathcal{B}_k$, which has a complicated representation. The paper then briefly discusses the k-cut number of some trees and general graphs. We conclude by some analytic results which may be of interest.

Keywords: cutting; k-cut; random trees.
AMS MSC 2010: 60C05.
Submitted to EJP on April 9, 2018, final version accepted on May 7, 2019.
Supersedes arXiv:1804.03069v2.

1 Introduction and main results

1.1 The k-cut number of a graph

Consider $\mathbb{G}_n$, a connected graph consisting of n nodes with exactly one node labeled as the *root*, which we call a *rooted* graph. Let k be a positive integer. We remove nodes from the graph as follows:

1. Choose a node uniformly at random from the component that contains the root. Cut the selected node once.
2. If this node has been cut k times, remove the node together with edges attached to it from the graph.
3. If the root has been removed, then stop. Otherwise, go to step 1.

*This work is supported by the Knut and Alice Wallenberg Foundation, the Swedish Research Council, and the Ragnar Söderberg Foundation.

†Uppsala University, Uppsala, Sweden. E-mail: xingshi.cai@math.uu.se E-mail: cecilia.holmgren@math.uu.se

‡McGill University, Montreal, Canada. E-mail: lucdevroye@gmail.com http://http://luc.devroye.org

§Uppsala University, Uppsala, Sweden. (Current affiliation: Masaryk University, Brno, Czech.)
E-mail: skerman@fi.muni.cz

We call the (random) total number of cuts needed to end this procedure the k-cut number and denote it by $\mathcal{K}(\mathbb{G}_n)$. (Note that in traditional cutting models, nodes are removed as soon as they are cut once, i.e., $k = 1$. But in our model, a node is only removed after being cut k times.)

One can also define an edge version of this process. Instead of cutting nodes, each time we choose an edge uniformly at random from the component that contains the root and cut it once. If the edge has been cut k-times then we remove it. The process stops when the root is isolated. We let $\mathcal{K}_e(\mathbb{G}_n)$ denote the number of cuts needed for the process to end.

Our model can also be applied to botnets, i.e., malicious computer networks consisting of compromised machines which are often used in spamming or attacks. The nodes in $\mathbb{G}_n$ represent the computers in a botnet, and the root represents the bot-master. The effectiveness of a botnet can be measured using the size of the component containing the root, which indicates the resources available to the bot-master [6]. To take down a botnet means to reduce the size of this root component as much as possible. If we assume that we target infected computers uniformly at random and it takes at least k attempts to fix a computer, then the k-cut number measures how difficult it is to completely isolate the bot-master.

The case $k = 1$ and $\mathbb{G}_n$ being a rooted tree has aroused great interests among mathematicians in the past few decades. The edge version of one-cut was first introduced by Meir and Moon [21] for the uniform random Cayley tree. Janson [16, 17] noticed the equivalence between one-cuts and records in trees and studied them in binary trees and conditional Galton-Watson trees. Later Addario-Berry, Broutin and Holmgren [1] gave a simpler proof for the limit distribution of one-cuts in conditional Galton-Watson trees. For one-cuts in random recursive trees, see [22, 15, 9]. For binary search trees and split trees, see [12, 13].

1.2 The k-cut number of a tree

One of the most interesting cases is when $\mathbb{G}_n = \mathbb{T}_n$, where $\mathbb{T}_n$ is a rooted tree with n nodes.

There is an equivalent way to define $\mathcal{K}(\mathbb{T}_n)$. Imagine that each node is given an alarm clock. At time zero, the alarm clock of node v is set to ring at time $T_{1,v}$, where $(T_{i,v})_{i\geq 1, v\in\mathbb{T}_n}$ are i.i.d. (independent and identically distributed) $\mathrm{Exp}(1)$ random variables. After the alarm clock of node v rings the i-th time, we set it to ring again at time $T_{i+1,v}$. Due to the memoryless property of exponential random variables (see [10, pp. 134]), at any moment, which alarm clock rings next is always uniformly distributed. Thus, if we cut a node that is still in the tree when its alarm clock rings, and remove the node with its descendants if it has already been cut k-times, then we get exactly the k-cut model. (The random variables $(T_{i,v})_{i\geq 1}$ can be seen as the holding times in a Poisson process $N(t)_v$ of parameter 1, where $N(t)_v$ is the number of cuts in v during the time $[0, t]$ and has a Poisson distribution with parameter t.)

How can we tell if a node is still in the tree? When node v's alarm clock rings for the r-th time for some $r \leq k$, and no node above v has already rung k times, we say v has become an r-*record*. And when a node becomes an r-record, it must still be in the tree. Thus, summing the number of r-records over $r \in \{1, \ldots, k\}$, we again get the k-cut number $\mathcal{K}(\mathbb{T}_n)$. One node can be a 1-record, a 2-record, etc., at the same time, so it can be counted multiple times. Note that if a node is an r-record, then it must also be an i-record for $i \in \{1, \ldots, r-1\}$.

To be more precise, we define $\mathcal{K}(\mathbb{T}_n)$ as a function of $(T_{i,v})_{i\geq 1, v\in\mathbb{T}_n}$. Let

$$G_{r,v} \stackrel{\text{def}}{=} \sum_{i=1}^{r} T_{i,v},$$

i.e., $G_{r,v}$ is the moment when the alarm clock of node v rings for the r-th time. Then $G_{r,v}$ has a gamma distribution with parameters $(r,1)$ (see [10, Theorem 2.1.12]), which we denote by $\mathrm{Gamma}(r)$. Let

$$I_{r,v} \stackrel{\text{def}}{=} [\![G_{r,v} < \min\{G_{k,u} : u \in \mathbb{T}_n, u \text{ is an ancestor of } v\}]\!], \tag{1.1}$$

where $[\![\cdot]\!]$ denotes the Iverson bracket, i.e., $[\![S]\!] = 1$ if the statement S is true and $[\![S]\!] = 0$ otherwise. In other words, $I_{r,v}$ is the indicator random variable for node v being an r-record. Let

$$\mathcal{K}_r(\mathbb{T}_n) \stackrel{\text{def}}{=} \sum_{v\in\mathbb{T}_n} I_{r,v}, \qquad \mathcal{K}(\mathbb{T}_n) \stackrel{\text{def}}{=} \sum_{r=1}^{k} \mathcal{K}_r(\mathbb{T}_n).$$

Then $\mathcal{K}_r(\mathbb{T}_n)$ is the number of r-records and $\mathcal{K}(\mathbb{T}_n)$ is the total number of records.

1.3 The k-cut number of a path

Let $\mathbb{P}_n$ be a one-ary tree (a path) consisting of n nodes labeled $1,\ldots,n$ from the root to the leaf. To simplify notations, from now on we use $I_{r,i}, G_{r,i}$, and $T_{r,i}$ to represent $I_{r,v}, G_{r,v}$ and $T_{r,v}$ respectively for a node v at depth i. Then (1.1) can be written as

$$I_{r,i+1} \stackrel{\text{def}}{=} [\![G_{r,j} < \min\{G_{k,j} : 1 \leq j \leq i\}]\!]. \tag{1.2}$$

Let $\mathcal{X}_n \stackrel{\text{def}}{=} \mathcal{K}(\mathbb{P}_n)$ and $\mathcal{X}_{n,r} = \mathcal{K}_r(\mathbb{P}_n)$. In this paper, we mainly consider $\mathcal{X}_n$ and we let $k \geq 2$ be a fixed integer.

The first motivation of this choice is that, as shown in Section 4, $\mathbb{P}_n$ is the fastest to cut among all graphs. (We make this statement precise in Lemma 4.1.) Thus $\mathcal{X}_n$ provides a universal stochastic lower bound for $\mathcal{K}(\mathbb{G}_n)$. Moreover, our results on $\mathcal{X}_n$ can immediately be extended to some trees of simple structures: see Section 4. Finally, as shown below, $\mathcal{X}_n$ generalizes the well-known record number in permutations and has different limit distributions when $k=1$, the usual cut-model, and $k \geq 2$, our extended model.

The name record comes from the classic definition of *records* in random permutations. Let $\sigma_1,\ldots,\sigma_n$ be a uniform random permutation of $\{1,\ldots,n\}$. If $\sigma_i < \min_{1\leq j<i}\sigma_j$, then i is called a *(strictly lower) record*. Let $\mathcal{R}_n$ denote the number of records in $\sigma_1,\ldots,\sigma_n$. Let $W_1,\ldots,W_n$ be i.i.d. random variables with a common continuous distribution. Since the relative order of $W_1,\ldots,W_n$ also gives a uniform random permutation, we can equivalently define σ_i as the rank of W_i. As gamma distributions are continuous, we can in fact let $W_i = G_{k,i}$. Thus, being a record in a uniform permutation is equivalent to being a k-record and $\mathcal{R}_n \stackrel{\mathcal{L}}{=} \mathcal{X}_{n,k}$. Moreover, when $k=1$, $\mathcal{R}_n \stackrel{\mathcal{L}}{=} \mathcal{X}_n$.

Starting from Chandler [5]'s article [5] in 1952, the theory of records has been widely studied due to its applications in statistics, computer science, and physics. For more recent surveys on this topic, see [2].

A well-known result of $\mathcal{R}_n$ (and thus also $\mathcal{X}_{n,k}$) [25] is that $(I_{k,j})_{1\leq j\leq n}$ are independent. It follows from the Lindeberg–Lévy–Feller Theorem that

$$\frac{\mathbf{E}\left[\mathcal{R}_n\right]}{\log n} \to 1, \qquad \frac{\mathcal{R}_n}{\log n} \xrightarrow{a.s.} 1, \qquad \mathcal{L}\left(\frac{\mathcal{R}_n - \log n}{\sqrt{\log n}}\right) \xrightarrow{d} \mathcal{N}(0,1),$$

where $\mathcal{N}(0,1)$ denotes the standard normal distribution.

In the following, Theorem 1.1 gives the expectation of $\mathcal{X}_{n,r}$ which implies that the number of one-records dominates the number of other records. Subsequently Theorem 1.2 and Theorem 1.3 estimate the variance and higher moments of $\mathcal{X}_{n,1}$.

Theorem 1.1. *For all fixed $k \in \mathbb{N}$,*

$$\mathbf{E}\left[\mathcal{X}_{n,r}\right] \sim \begin{cases} \eta_{k,r} n^{1-\frac{r}{k}} & (1 \le r < k), \\ \log n & (r = k), \end{cases}$$

where the constants $\eta_{k,r}$ are defined by

$$\eta_{k,r} \stackrel{\text{def}}{=} \frac{(k!)^{\frac{r}{k}}}{k-r} \frac{\Gamma\left(\frac{r}{k}\right)}{\Gamma(r)},$$

where $\Gamma(z)$ denotes the gamma function. Therefore $\mathbf{E}\left[\mathcal{X}_n\right] \sim \mathbf{E}\left[\mathcal{X}_{n,1}\right]$. Also, for $k = 2$,

$$\mathbf{E}\left[\mathcal{X}_n\right] \sim \mathbf{E}\left[\mathcal{X}_{n,1}\right] \sim \sqrt{2\pi n}.$$

Theorem 1.2. *For all fixed $k \in \{2, 3, \dots\}$,*

$$\mathbf{E}\left[\mathcal{X}_{n,1}(\mathcal{X}_{n,1} - 1)\right] \sim \mathbf{E}\left[(\mathcal{X}_{n,1})^2\right] \sim \gamma_k n^{2-\frac{2}{k}},$$

where

$$\gamma_k = \frac{\Gamma\left(\frac{2}{k}\right)(k!)^{\frac{2}{k}}}{k-1} + 2\lambda_k,$$

and

$$\lambda_k = \begin{cases} \dfrac{\pi \cot\left(\frac{\pi}{k}\right)\Gamma\left(\frac{2}{k}\right)(k!)^{\frac{2}{k}}}{2(k-2)(k-1)} & k > 2, \\ \dfrac{\pi^2}{4} & k = 2. \end{cases}$$

Therefore

$$\operatorname{Var}\left(\mathcal{X}_{n,1}\right) \sim \left(\gamma_k - \eta_{k,1}^2\right) n^{2-\frac{2}{k}}.$$

In particular, when $k = 2$

$$\operatorname{Var}\left(\mathcal{X}_{n,1}\right) \sim \left(\frac{\pi^2}{2} + 2 - 2\pi\right) n.$$

Theorem 1.3. *For all fixed $k \in \{2, 3, \dots\}$ and $\ell \in \mathbb{N}$*

$$\limsup_{n\to\infty} \mathbf{E}\left[\left(\frac{\mathcal{X}_{n,1}}{n^{1-\frac{1}{k}}}\right)^{\ell}\right] \le \rho_{k,\ell} \stackrel{\text{def}}{=} \ell!\Gamma\left(\ell + 1 - \frac{\ell}{k}\right)^{-1}\left(\frac{\pi}{k}(k!)^{1/k}\sin\left(\frac{\pi}{k}\right)^{-1}\right)^{\ell}.$$

The upper bound is tight for $\ell = 1$ since $\rho_{k,1} = \eta_{k,1}$.

The above theorems imply that the correct rescaling parameter should be $n^{1-\frac{1}{k}}$. However, unlike the case $k = 1$, when $k \ge 2$ the limit distribution of $\mathcal{X}_n/n^{1-\frac{1}{k}}$ has a rather complicated representation $\mathcal{B}_k$ defined as follows: Let $U_1, E_1, U_2, E_2, \dots$ be mutually independent random variables with $E_j \stackrel{\mathcal{L}}{=} \operatorname{Exp}(1)$ and $U_j \stackrel{\mathcal{L}}{=} \operatorname{Unif}[0,1]$. Let

$$S_p \stackrel{\text{def}}{=} \left(k! \sum_{1 \le s \le p} \left(\prod_{s \le j < p} U_j\right) E_s\right)^{\frac{1}{k}}, \tag{1.3}$$

$$B_p \stackrel{\text{def}}{=} (1-U_p)\left(\prod_{1\le j<p} U_j\right)^{1-\frac{1}{k}} S_p, \tag{1.4}$$

$$\mathcal{B}_k \stackrel{\text{def}}{=} \sum_{1\le p} B_p, \tag{1.5}$$

where we use the convention that an empty product equals one.

Remark 1.4. An equivalent recursive definition of S_p is

$$S_p = \begin{cases} k!E_1 & (p=1), \\ \left(U_{p-1}S_{p-1}^k + k!E_p\right)^{\frac{1}{k}} & (p\ge 2). \end{cases}$$

Theorem 1.5. *Let $k \in \{2,3,\dots\}$. Let $\mathcal{L}(\mathcal{B}_k)$ denote the distribution of $\mathcal{B}_k$. Then*

$$\mathcal{L}\left(\frac{\mathcal{X}_n}{n^{1-\frac{1}{k}}}\right) \xrightarrow{d} \mathcal{L}(\mathcal{B}_k).$$

Thus, by Theorem 1.1, 1.2 and 1.3, the convergence also holds in L^p for all $p>0$ and

$$\mathbf{E}\left[\mathcal{B}_k\right] = \eta_{k,1}, \qquad \mathbf{E}\left[\mathcal{B}_k^2\right] = \gamma_k, \qquad \mathbf{E}\left[\mathcal{B}_k^p\right] \in [\eta_{k,1}^p, \rho_{k,p}] \qquad (p\in\mathbb{N}).$$

Remark 1.6. The idea behind $\mathcal{B}_k$ is that we split the path into segments according to the positions of k-records, then we count the numbers of one-records in every segment, each of which converges to a B_p in the sum (1.5). This will be made rigorous in Section 3. We will also see that $\mathcal{B}_k$ has a density close to a normal distribution in Section 3.4.

Remark 1.7. It is easy to see that $\mathcal{X}_{n+1}^{\mathrm{e}} \stackrel{\text{def}}{=} \mathcal{K}_e(P_{n+1}) \stackrel{\mathcal{L}}{=} \mathcal{X}_n$ by treating each edge on a length $n+1$ path as a node on a length n path.

The rest of the paper is organized as follows: Section 2 proves the moment results Theorem 1.1, 1.2, and 1.3. Section 3 deals with the distributional result Theorem 1.5. Section 4 discusses some easy results for general graphs and trees. Finally, Section 5 collects analytic results used in the proofs, which may themselves be of interest.

2 The moments

2.1 The expectation

In this section we prove Theorem 1.1.

Lemma 2.1. *Uniformly for all $i\ge 1$ and $r\in\{1,\dots,k\}$,*

$$\mathbf{E}\left[I_{r,i+1}\right] = \left(1+O\left(i^{-\frac{1}{2k}}\right)\right)\frac{(k!)^{\frac{r}{k}}}{k}\frac{\Gamma\left(\frac{r}{k}\right)}{\Gamma(r)} i^{-\frac{r}{k}}.$$

Proof. By (1.2), $\mathbf{E}\left[I_{r,i+1}\right] = \mathbf{P}\left\{G_{k,1} > G_{r,i+1},\dots,G_{k,i} > G_{r,i+1}\right\}$. Thus conditioning on $G_{r,i+1}=x$ yields $\mathbf{E}\left[I_{r,i+1}\right] = \int_0^\infty x^{r-1}\mathrm{e}^{-x}/\Gamma(r)\mathbf{P}\left\{G_{k,1}>x\right\}^i \,\mathrm{d}x$. Theorem 2.1 thus follows from Theorem 5.2. □

Proof of Theorem 1.1. A simple computation shows that for $a\in(0,1)$

$$\sum_{1\le i\le n}\frac{1}{i^a} = \frac{1}{1-a}n^{1-a} + O(1). \tag{2.1}$$

It then follows from Theorem 2.1 that for $r\in\{1,\dots,k-1\}$

$$\mathbf{E}\left[\mathcal{X}_{n,r}\right] = \sum_{0\le i<n}\mathbf{E}\left[I_{r,i+1}\right] = \frac{(k!)^{\frac{r}{k}}}{k}\frac{\Gamma\left(\frac{r}{k}\right)}{\Gamma(r)}\frac{1}{1-\frac{r}{k}}n^{1-\frac{r}{k}} + O\left(n^{1-\frac{r}{k}-\frac{1}{2k}}\right) + O(1).$$

When $r=k$, $\mathbf{E}\left[\mathcal{X}_{n,k}\right] = \mathbf{E}\left[\mathcal{R}_n\right] \sim \log(n)$ is already well-known. □

2.2 The variance

In this section we prove Theorem 1.2.

Let $E_{i,j}$ denote the event that $[I_{1,i+1}I_{1,j+1} = 1]$. Let $A_{x,y}$ denote the event that $[G_{1,i+1} = x \cap G_{1,j+1} = y]$. Then conditioning on $A_{x,y}$

$$E_{i,j} = \left[\bigcap_{1\le s\le i} G_{k,s} > x \vee y\right] \cap [G_{k,i+1} > y] \cap \left[\bigcap_{i+2\le s\le j} G_{k,s} > y\right],$$

where $x \vee y \overset{\text{def}}{=} \max\{x, y\}$. Since conditioning on $A_{x,y}$, $G_{k,i+1} \overset{\mathcal{L}}{=} \text{Gamma}(k-1) + x$, $G_{k,s} \overset{\mathcal{L}}{=} \text{Gamma}(k)$ for $s \notin \{i+1, j+1\}$, and all these random variables are independent, we have

$$\mathbf{P}\{E_{i,j}|A_{x,y}\} = \mathbf{P}\{G_{k-1,1} + x > y\}\, \mathbf{P}\{G_{k,1} > x \vee y\}^{i}\, \mathbf{P}\{G_{k,1} > y\}^{j-i-1}. \tag{2.2}$$

It follows from $G_{1,i+1} \overset{\mathcal{L}}{=} G_{1,j+1} \overset{\mathcal{L}}{=} \text{Exp}(1)$ that

$$\begin{aligned}\mathbf{P}\{E_{i,j}\} &= \int_0^\infty \int_y^\infty \mathrm{e}^{-x-y}\mathbf{P}\{E_{i,j}|A_{x,y}\}\, \mathrm{d}x\, \mathrm{d}y + \int_0^\infty \int_0^y \mathrm{e}^{-x-y}\mathbf{P}\{E_{i,j}|A_{x,y}\}\, \mathrm{d}x\, \mathrm{d}y \\ &\overset{\text{def}}{=} A_{1,i,j} + A_{2,i,j}.\end{aligned}$$

We next estimate these two terms.

Lemma 2.2. *Let $k \in \{2, 3, \dots\}$. We have*

$$A_{2,i,j} = \left(1 + O\left(j^{-\frac{1}{2k}}\right)\right) \frac{(k!)^{\frac{2}{k}}}{k} \Gamma\left(\frac{2}{k}\right) j^{-\frac{2}{k}}.$$

Proof. In this case, $x \vee y = y$. Thus, by (2.2)

$$A_{2,i,j} = \int_0^\infty \mathrm{e}^{-y}\mathbf{P}\{G_{k,1} > y\}^{j-1} \int_0^y \mathrm{e}^{-x}\mathbf{P}\{G_{k-1,1} > y - x\}\, \mathrm{d}x\, \mathrm{d}y.$$

Note that the dependence on i disappears. Let Z denote a Poisson random variable with mean $y - x$. By the well-known connection between Poisson and gamma distributions, the inner integral in the above equals

$$\int_0^y \mathrm{e}^{-x}\mathbf{P}\{Z < k-1\}\, \mathrm{d}x = \int_0^y \mathrm{e}^{-x} \sum_{\ell=0}^{k-2} \mathrm{e}^{-(y-x)} \frac{(y-x)^\ell}{\ell!}\, \mathrm{d}x = \mathrm{e}^{-y} \sum_{\ell=0}^{k-2} \frac{y^{\ell+1}}{(\ell+1)!}.$$

It then follows from Theorem 5.2 that

$$\begin{aligned}A_{2,i,j} &= \sum_{\ell=0}^{k-2} \int_0^\infty \mathrm{e}^{-2y} \frac{y^{\ell+1}}{(\ell+1)!}\mathbf{P}\{G_{k,1} > y\}^{j-1}\, \mathrm{d}y \\ &= \sum_{\ell=0}^{k-2} \left(1 + O\left(j^{-\frac{1}{2k}}\right)\right) \frac{(k!)^{\frac{\ell+2}{k}}}{k(\ell+1)!} \Gamma\left(\frac{\ell+2}{k}\right) j^{-\frac{\ell+2}{k}} \\ &= \left(1 + O\left(j^{-\frac{1}{2k}}\right)\right) \frac{(k!)^{\frac{2}{k}}}{k} \Gamma\left(\frac{2}{k}\right) j^{-\frac{2}{k}}.\end{aligned}$$

□

Lemma 2.3. *Let $k \in \{2, 3, \dots\}$. Let $a = i$ and $b = j - i - 1$. Then for all $a \ge 1$ and $b \ge 1$,*

$$A_{1,i,j} = \xi_k(a, b) + O\left(\left(a^{-\frac{1}{2k}} + b^{-\frac{1}{2k}}\right)\left(a^{-\frac{2}{k}} + b^{-\frac{2}{k}}\right)\right),$$

where

$$\xi_k(a, b) \overset{\text{def}}{=} \int_0^\infty \int_y^\infty \exp\left(-a\frac{x^k}{k!} - b\frac{y^k}{k!}\right) \mathrm{d}x\, \mathrm{d}y.$$

Proof. In this case, $x \vee y = x$ and $y - x < 0$. Thus, by (2.2) and Theorem 5.2

$$
\begin{aligned}
A_{1,i,j} &= \int_0^\infty \int_y^\infty \mathrm{e}^{-x}\mathrm{e}^{-y}\mathbf{P}\left\{G_{k,1} > x\right\}^i \mathbf{P}\left\{G_{k,1} > y\right\}^{j-i-1} \,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^\infty \int_y^\infty \mathrm{e}^{-x-y} \left(\frac{\Gamma(k,x)}{\Gamma(k)}\right)^a \left(\frac{\Gamma(k,y)}{\Gamma(k)}\right)^b \mathrm{d}x\,\mathrm{d}y, \qquad (2.3)
\end{aligned}
$$

where $\Gamma(\ell, z)$ denotes the upper incomplete gamma function.

Let $\mathcal{S}$ be the integration area of (2.3). Let $x_0 = a^{-\alpha}$ and $y_0 = b^{-\alpha}$ where $\alpha = \frac{1}{2}\left(\frac{1}{k} + \frac{1}{k+1}\right)$. Let

$$
\mathcal{S}_0 = \mathcal{S} \cap \left\{(x,y) \in \mathbb{R}^2 : x < x_0, y < y_0\right\}.
$$

We split (2.3) into two parts $A_{1,1}$ and $A_{1,2}$ with integration area $\mathcal{S}_0$ and $\mathcal{S} \setminus \mathcal{S}_0$ respectively.

Let $\beta = \frac{1}{2(k+1)}$. Let $x_1 = a^\beta/k!$ and $y_1 = b^\beta/k!$. It follows from Theorem 5.1 and Theorem 5.4 that

$$
\begin{aligned}
A_{1,1} &= \left(1 + O\left(a^{-\frac{1}{2k}} + b^{-\frac{1}{2k}}\right)\right) \iint_{\mathcal{S}_0} \exp\left(-a\frac{x^k}{k!} - b\frac{y^k}{k!}\right) \mathrm{d}x\,\mathrm{d}y \\
&= \left(1 + O\left(a^{-\frac{1}{2k}} + b^{-\frac{1}{2k}}\right)\right) \xi_k(a,b) + O\left(\mathrm{e}^{-x_1} + \mathrm{e}^{-y_1}\right) \\
&= \xi_k(a,b) + O\left(\left(a^{-\frac{1}{2k}} + b^{-\frac{1}{2k}}\right)\left(a^{-\frac{2}{k}} + b^{-\frac{2}{k}}\right)\right).
\end{aligned}
$$

It is not difficult to verify that

$$
A_{1,2} = O\left(\left(\frac{\Gamma(k,x_0)}{\Gamma(k)}\right)^{-a} + \left(\frac{\Gamma(k,y_0)}{\Gamma(k)}\right)^{-b}\right) = O\left(\mathrm{e}^{-x_1} + \mathrm{e}^{-y_1}\right). \qquad \square
$$

Proof of Theorem 1.2. We have

$$
\mathbf{E}\left[\mathcal{X}_{n,1}\left(\mathcal{X}_{n,1} - 1\right)\right] = 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbf{P}\left\{E_{i,j}\right\} = 2\sum_{i=0}^{n-2}\sum_{j=i+1}^{n-1}\left(A_{1,i,j} + A_{2,i,j}\right). \qquad (2.4)
$$

Thus, by Theorem 2.2 and (2.1),

$$
\begin{aligned}
\sum_{i=0}^{n-2}\sum_{j=i+1}^{n-1} A_{2,i,j} &= \sum_{i=0}^{n-2}\sum_{j=i+1}^{n-1}\left[\frac{(k!)^{\frac{2}{k}}\Gamma\left(\frac{2}{k}\right)}{k} j^{-\frac{2}{k}} + O\left(j^{-\frac{5}{2k}}\right)\right] \\
&= \frac{(k!)^{\frac{2}{k}}\Gamma\left(\frac{2}{k}\right)}{2(k-1)} n^{2-\frac{2}{k}} + O\left(n^{2-\frac{5}{2k}}\right). \qquad (2.5)
\end{aligned}
$$

For $A_{1,i,j}$, it follows from Theorem 2.3 that

$$
\begin{aligned}
\sum_{i=0}^{n-2}\sum_{j=i+1}^{n-1} A_{1,i,j} &= \sum_{a=1}^{n-1}\sum_{b=1}^{n-a}\xi_k(a,b) + O\left(n^{2-\frac{5}{2k}}\right) \\
&= \int_0^n \int_0^{n-a} \xi_k(a,b)\,\mathrm{d}b\,\mathrm{d}a + O\left(n^{2-\frac{5}{2k}}\right) \\
&= n^{2-\frac{2}{k}} \int_0^1 \int_0^{1-s} \xi_k(s,t)\,\mathrm{d}t\,\mathrm{d}s + O\left(n^{2-\frac{5}{2k}}\right) \\
&= \lambda_k n^{2-\frac{2}{k}} + O\left(n^{2-\frac{5}{2k}}\right), \qquad (2.6)
\end{aligned}
$$

where the last step follows from Theorem 5.5. Theorem 1.2 follows by putting (2.5), (2.6) into (2.4). $\square$

2.3 Higher moments

In this section we prove Theorem 1.3.

The computations of higher moments of $\mathcal{X}_{n,1}$ are rather complicated. However, an upper bound is readily available. Let $(x)_\ell \stackrel{\text{def}}{=} x(x-1)\dots(x-\ell+1)$. For $\ell \ge 1$,

$$\begin{aligned}
\mathbf{E}\left[(\mathcal{X}_{n,1})_\ell\right] &= \ell! \sum_{1\le i_1<i_2\cdots<i_\ell\le n} \mathbf{E}\left[I_{1,i_1} I_{1,i_2}\cdots I_{1,i_\ell}\right] \\
&\le \ell! \sum_{1\le i_1<i_2\cdots<i_\ell\le n} \mathbf{E}\left[I_{1,i_1}\right]\mathbf{E}\left[I_{1,i_2-i_1}\right]\cdots \mathbf{E}\left[I_{1,i_\ell-i_{\ell-1}}\right] \\
&= \ell! \sum_{(a_1,\dots,a_\ell)\in\mathcal{S}_{n,\ell}} \prod_{j=1}^{\ell} \mathbf{E}\left[I_{1,a_j}\right], \qquad (2.7)
\end{aligned}$$

where

$$\mathcal{S}_{n,\ell} \stackrel{\text{def}}{=} \left\{(a_1,a_2,\dots,a_\ell)\in\mathbb{N}^\ell : a_1\ge 0,\dots,a_\ell\ge 0, \sum_{j=1}^{\ell} a_j \le n-\ell\right\}.$$

The above inequality holds since if i_j is a one-record in the whole path, then it must also be a one-record in the segment $(i_{j-1}+1,\dots,i_j)$ ignoring everything else, and what happens in each of such segments are independent. It follows from Theorem 2.1 that (2.7) equals

$$\begin{aligned}
&\ell! \sum_{(a_1,\dots,a_\ell)\in\mathcal{S}_{n,\ell}} \prod_{j=1}^{\ell} \left(1+O\left(a_j^{-\frac{1}{2k}}\right)\right) \frac{(k!)^{\frac{1}{k}}}{k}\Gamma\left(\frac{1}{k}\right) a_j^{-\frac{1}{k}} \\
&= \ell! n^{\ell\left(1-\frac{1}{k}\right)} \left(\frac{(k!)^{\frac{1}{k}}}{k}\Gamma\left(\frac{1}{k}\right)\right)^{\ell} \sum_{(a_1,\dots,a_\ell)\in\mathcal{S}_{n,\ell}} \prod_{j=1}^{\ell} \left(1+O\left(a_j^{-\frac{1}{2k}}\right)\right)\left(\frac{a_j}{n}\right)^{-\frac{1}{k}} \frac{1}{n} \\
&\sim n^{\ell\left(1-\frac{1}{k}\right)} \ell! \left(\frac{(k!)^{\frac{1}{k}}}{k}\Gamma\left(\frac{1}{k}\right)\right)^{\ell} \int_{A_\ell} \prod_{j=1}^{\ell} x_j^{-\frac{1}{k}} \,\mathrm{d}(x_1,\dots,x_\ell) \\
&= n^{\ell\left(1-\frac{1}{k}\right)} \ell! \left(\frac{(k!)^{\frac{1}{k}}}{k}\Gamma\left(\frac{1}{k}\right)\right)^{\ell} \Gamma\left(\frac{k-1}{k}\right)^{\ell} \Gamma\left(1+\ell-\frac{\ell}{k}\right)^{-1} \stackrel{\text{def}}{=} n^{\ell\left(1-\frac{1}{k}\right)} \rho_{k,\ell},
\end{aligned}$$

where A_ℓ is the simplex $\{(x_1,\dots,x_\ell) : x_1>0,\dots,x_\ell>0, x_1+\dots+x_\ell<1\}$. The above integral is known as the *beta* integral [24, 5.14.1].

3 Convergence to the k-cut distribution

By Theorem 1.1 and Markov's inequality, $\mathcal{X}_{n,r}/n^{1-\frac{1}{k}} \xrightarrow{p} 0$ for $r\in\{2,\dots,k\}$. So it suffices to prove Theorem 1.5 for $\mathcal{X}_{n,1}$ instead of $\mathcal{X}_n$. Throughout Section 3, unless otherwise emphasized, we assume that $k\ge 2$.

The idea of the proof is to condition on the positions and values of the k-records, and study the distribution of the number of one-records between two consecutive k-records.

We use $(R_{n,q})_{q\ge 1}$ to denote the k-record values and $(P_{n,q})_{q\ge1}$ the positions of these k-records. To be precise, let $R_{n,0}\stackrel{\text{def}}{=} 0$, and $P_{n,0}\stackrel{\text{def}}{=} n+1$; for $q\ge1$, if $P_{n,q-1}>1$, then let

$$\begin{aligned}
R_{n,q} &\stackrel{\text{def}}{=} \min\{G_{k,j} : 1\le j< P_{n,q-1}\}, \\
P_{n,q} &\stackrel{\text{def}}{=} \operatorname{argmin}\{G_{k,j} : 1\le j< P_{n,q-1}\},
\end{aligned}$$

i.e., $P_{n,q}$ is the unique positive integer which satisfies that $G_{k,P_{n,q}} < G_{k,i}$ for all $1 \le i < P_{n,q-1}$; otherwise let $P_{n,q} = 1$ and $R_{n,q} = \infty$. Note that $R_{n,1}$ is simply the minimum of n i.i.d. $\mathrm{Gamma}(k)$ random variables.

According to $(P_{n,q})_{q\ge1}$, we split $\mathcal{X}_{n,1}$ into the following sum

$$\mathcal{X}_{n,1} = \sum_{1\le j\le n} I_{1,j} = \mathcal{X}_{n,k} + \sum_{1\le q}\sum_{1\le j} [\![P_{n,q-1} > j > P_{n,q}]\!]\, I_{1,j} \stackrel{\text{def}}{=} \mathcal{X}_{n,k} + \sum_{1\le q} B_{n,q}. \tag{3.1}$$

Figure 1 gives an example of $(B_{n,q})_{q\ge1}$ for $n = 12$. It depicts the positions of the k-records and the one-records. It also shows the values and the summation ranges for $(B_{n,q})_{q\ge1}$.

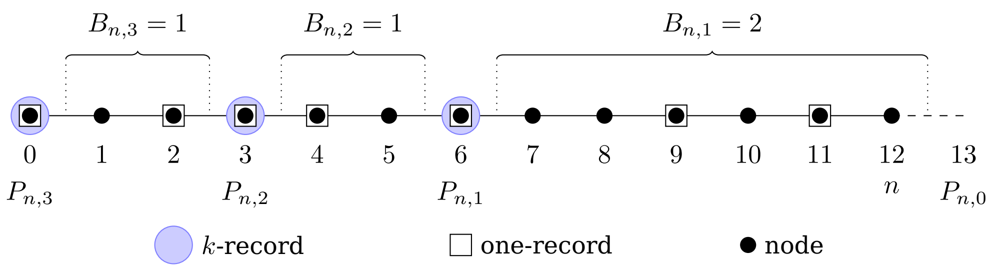


Figure 1: An example of $(B_{n,q})_{q\ge1}$ for $n = 12$.

Recall that $T_{r,j} \stackrel{\mathcal{L}}{=} \mathrm{Exp}(1)$, is the lapse of time between the alarm clock of j rings for the $(r-1)$-st time and the r-th time. Conditioning on $(R_{n,q}, P_{n,q})_{n\ge1,q\ge1}$, for $j \in (P_{n,q}, P_{n,q-1})$, we have

$$\mathbf{E}\left[I_{1,j}\right] = \mathbf{P}\left\{T_{1,j} < R_{n,q} \,|G_{k,j} > R_{n,q}\right\}.$$

Then the distribution of $B_{n,q}$ conditioning on $(R_{n,q}, P_{n,q})_{n\ge1,q\ge1}$ is simply that of

$$\mathrm{Bin}\left(P_{n,q-1} - P_{n,q} - 1, \mathbf{P}\left\{T_{1,j} < R_{n,q} \,|G_{k,j} > R_{n,q}\right\}\right),$$

where $\mathrm{Bin}(m, q)$ denotes a binomial (m, q) random variable. When $R_{n,q}$ is small and $P_{n,q-1} - P_{n,q}$ is large, this is roughly

$$\mathrm{Bin}\left(P_{n,q-1} - P_{n,q}, \mathbf{P}\left\{T_{1,j} < R_{n,q}\right\}\right) \stackrel{\mathcal{L}}{=} \mathrm{Bin}\left(P_{n,q-1} - P_{n,q}, 1 - \mathrm{e}^{-R_{n,q}}\right). \tag{3.2}$$

Therefore, we first study a slightly simplified model. Let $(T^*_{r,j})_{r\ge1,j\ge1}$ be i.i.d. $\mathrm{Exp}(1)$ which are also independent from $(T_{r,j})_{r\ge1,j\ge1}$. Let

$$I^*_j \stackrel{\text{def}}{=} [\![T^*_{1,j} < \min\{G_{k,i} : 1 \le i \le j\}]\!], \qquad \mathcal{X}^*_n \stackrel{\text{def}}{=} \sum_{1\le j\le n} I^*_j.$$

We say a node j is an *alt-one-record* if $I^*_j = 1$. As in (3.1), we can write

$$\mathcal{X}^*_n = \sum_{1\le j\le n} I^*_j = \sum_{1\le q}\sum_{1\le j} [\![P_{n,q-1} > j \ge P_{n,q}]\!]\, I^*_j \stackrel{\text{def}}{=} \sum_{1\le q} B^*_{n,q}. \tag{3.3}$$

Then conditioning on $(R_{n,q}, P_{n,q})_{n\ge1,q\ge1}$, $B^*_{n,q}$ has exactly the distribution as (3.2). Figure 2 gives an example of $(B^*_{n,q})_{q\ge1}$ for $n = 12$. It shows the positions of alt-one-records, as well as the values and the summation ranges of$(B^*_{n,q})_{q\ge1}$.

In the rest of this section, we will first prove the following proposition:

Proposition 3.1. *For all fixed $q \in \mathbb{N}$ and $k \ge 2$,*

$$\mathcal{L}\left(\left(\frac{B^*_{n,1}}{n^{1-\frac{1}{k}}}, \ldots, \frac{B^*_{n,q}}{n^{1-\frac{1}{k}}}\right)\right) \xrightarrow{d} \mathcal{L}\left((B_1, \ldots B_q)\right),$$

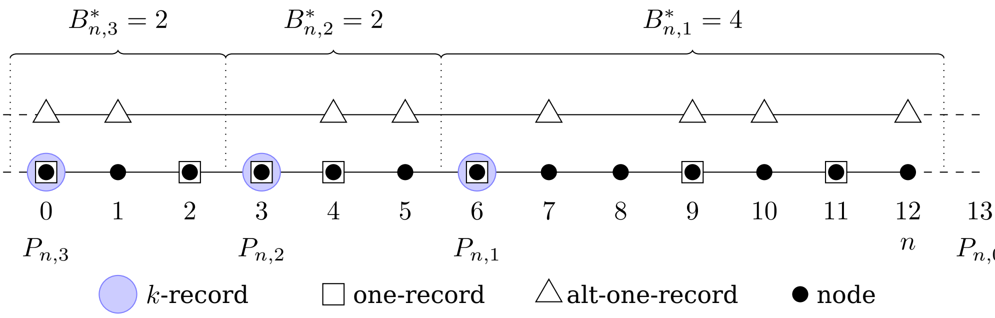


Figure 2: An example of $(B^*_{n,q})_{q\geq 1}$ for $n = 12$.

which implies by the Cramér–Wold device that

$$\mathcal{L}\left(\sum_{1\leq j\leq q} \frac{B^*_{n,j}}{n^{1-\frac{1}{k}}}\right) \xrightarrow{d} \mathcal{L}\left(\sum_{1\leq j\leq q} B_j\right), \tag{3.4}$$

Then we can prove that q can be chosen large enough so that $\sum_{q<j} B^*_{n,j}/n^{1-\frac{1}{k}}$ is negligible. Thus,

$$\mathcal{L}\left(\frac{\mathcal{X}^*_n}{n^{1-\frac{1}{k}}}\right) \stackrel{\text{def}}{=} \mathcal{L}\left(\frac{\sum_{1\leq j} B^*_{n,j}}{n^{1-\frac{1}{k}}}\right) \xrightarrow{d} \mathcal{L}\left(\sum_{1\leq j} B_j\right) \stackrel{\text{def}}{=} \mathcal{L}\left(\mathcal{B}_k\right).$$

Following this, we can use a coupling argument to show that $\mathcal{X}_{n,1}/n^{1-\frac{1}{k}}$ and $\mathcal{X}^*_n/n^{1-\frac{1}{k}}$ converge to the same limit, which finishes the proof of Theorem 1.5. The section ends with some discussions on $\mathcal{B}_k$.

3.1 Proof of Theorem 3.1

To prove (3.4), we construct a coupling by defining all the random variables being studied in one probability space. Let

$$P_{n,q} = \max\left\{\lceil U_q\left(P_{n,q-1} - 1\right)\rceil, 1\right\},$$

for $q \geq 1$, where $(U_q)_{q\geq 1}$ are i.i.d. $\mathrm{Unif}[0,1]$ random variables, independent of everything else. This is a valid coupling, since conditioning on $P_{n,q-1}$, $P_{n,q}$ is uniformly distributed on $\{1,\ldots,P_{n,q-1}-1\}$. Note that by induction on q this implies that for all $q \in \mathbb{N}$

$$\frac{P_{n,q}}{n} \xrightarrow{a.s.} \prod_{1\leq s\leq q} U_s. \tag{3.5}$$

Then conditioning on $(P_{n,q})_{q\geq 1}$, we generate the random variables $(T_{r,j})_{r\geq 1, j\geq 1}$ according to their proper conditional distribution, which determine $(G_{r,j})_{r\geq 1,j\geq 1}$ and $(R_{n,q})_{q\geq 1}$. Let $(T^*_{r,j})_{r\geq 1,j\geq 1}$ be as before.

Recall that $R_{m,1}$ is the minimum of m independent $\mathrm{Gamma}(k)$ random variables. Let $M(m,t) \stackrel{\text{def}}{=} (R_{m,1} | R_{m,1} > t)$ for $t \geq 0$. Then conditioning on $P_{n,q-1}$ and $R_{n,q-1}$, $R_{n,q} \stackrel{\mathcal{L}}{=} M(P_{n,q-1}-1, R_{n,q-1})$. The following lemma allows us to describe the limit distribution of $R_{n,q}$ conditioning on $P_{n,q-1}$ and $R_{n,q-1}$.

Lemma 3.2. *Let $k \in \mathbb{N}$. Assume that $\frac{r_m}{m} \to 1$ and $t \geq 0$. Let $H_m \stackrel{\text{def}}{=} r_m^{\frac{1}{k}} \cdot M\left(m, t r_m^{-\frac{1}{k}}\right)$. Then as $m \to \infty$,*

$$\mathcal{L}\left(H_m\right) \xrightarrow{d} \mathcal{L}\left(\left(t^k + k!E\right)^{\frac{1}{k}}\right),$$

where $E \stackrel{\mathcal{L}}{=} \operatorname{Exp}(1)$*. In particular,* $\mathcal{L}\left(m^{\frac{1}{k}} M(m,0)\right) \xrightarrow{d} \mathcal{L}\left((k!E)^{\frac{1}{k}}\right)$*. The convergence is also point-wise for the density functions. The lemma also holds if* H_m *by is replaced by*

$$H'_m \stackrel{\text{def}}{=} r_m^{\frac{1}{k}} \cdot \left(1 - \exp\left(-M\left(m, t r_m^{-\frac{1}{k}}\right)\right)\right).$$

Proof. We only prove the lemma for H_m. Similar argument works for H'_m. Let $y_m = x/r_m^{\frac{1}{k}}$ and let $s_m = t/r_m^{\frac{1}{k}}$. By Theorem 5.1, for all fixed $x \geq t$,

$$\begin{aligned} \mathbf{P}\left\{H_m > x\right\} &= \frac{\mathbf{P}\left\{R_{m,1} \geq y_m\right\}}{\mathbf{P}\left\{R_{m,1} \geq s_m\right\}} = \left(\frac{\Gamma\left(k, y_m\right)}{\Gamma\left(k, s_m\right)}\right)^m \sim \exp\left(m\left(-\frac{y_m^k - s_m^k}{k!}\right)\right) \\ &\to \exp\left(-\frac{x^k - t^k}{k!}\right) = \mathbf{P}\left\{\left(t^k + k!E\right)^{\frac{1}{k}} > x\right\}. \end{aligned} \tag{3.6}$$

Using (3.6) and the derivative formula for the incomplete gamma functions [24, 8.8.13], it is straightforward to verify the point-wise convergence of the density functions. □

The next step is to recursively apply Theorem 3.2 to get a joint convergence in distribution for $(S_{n,1}, \dots, S_{n,q})$ as well as $(S^*_{n,1}, \dots, S^*_{n,q})$, which are basically rescaled versions of $(R_{n,1}, \dots, R_{n,q})$ defined by

$$L^*_{n,q} \stackrel{\text{def}}{=} \left(n \prod_{1 \leq j < q} U_j\right)^{\frac{1}{k}}, \quad S_{n,q} \stackrel{\text{def}}{=} L^*_{n,q} R_{n,q}, \quad S^*_{n,q} \stackrel{\text{def}}{=} L^*_{n,q}(1 - \mathrm{e}^{-R_{n,q}}).$$

Lemma 3.3. *For all fixed* $q \in \mathbb{N}$ *and* $k \in \{2, 3, \dots\}$*,*

$$\mathcal{L}\left((S_{n,1}, S_{n,2}, \dots, S_{n,q})\right) \xrightarrow{d} \mathcal{L}\left((S_1, S_2, \dots, S_q)\right).$$

The convergence is also point-wise for the joint distribution functions. The lemma holds if $S_{n,j}$ *is replaced by* $S^*_{n,j}$*.*

Proof. We only prove the lemma for $S_{n,j}$. The same argument works for $S^*_{n,j}$.

Let $\mathcal{F} = \sigma((U_j)_{j \geq 1})$ denote the sigma algebra generated by $(U_j)_{j \geq 1}$. Throughout the proof of this lemma, we will condition on $\mathcal{F}$ and treat $(U_q, P_{n,q}, L^*_{n,q})_{q \geq 0, n \geq 1}$ as if they are deterministic numbers.

Let $f_{n,1}(\cdot)$ and $f_1(\cdot)$ denote the density functions of $S_{n,1}$ and S_1 respectively. For $q > 1$, let $f_{n,q}(\cdot | y_{q-1})$ and $f_q(\cdot | y_{q-1})$ denote the density function of $S_{n,q} | S_{n,q-1} = y_{q-1}$, and $S_q | S_{q-1} = y_{q-1}$ respectively. It follows from Theorem 3.2 that for all $y_1 \geq 0$, $f_{n,1}(y_1) \to f_q(y_1)$, and for all $y_q \geq 0$, $f_{n,q}(y_q | y_{q-1}) \to f_q\left(y_q | y_{q-1}\right)$. Therefore, for all $y_1, \dots, y_q \in [0, \infty)^q$, as $n \to \infty$,

$$\begin{aligned} g_{n,q}(y_1, \dots, y_q) &\stackrel{\text{def}}{=} f_{n,q}(y_q | y_{q-1}) f_{n,q-1}(y_{q-1} | y_{q-2}) \dots f_{n,1}(y_1) \\ &\to f_q(y_q | y_{q-1}) f_{q-1}(y_{q-1} | y_{q-2}) \dots f_1(y_1) \stackrel{\text{def}}{=} g_q(y_1, \dots, y_q). \end{aligned}$$

In other words, the joint density function of $(S_{n,1}, \dots, S_{n,q})$ converges point-wise to the joint density function of $(S_1, \dots, S_q)$ conditioning on $\mathcal{F}$. Thus, the lemma still holds without conditioning on $\mathcal{F}$. □

One last ingredient needed is the next lemma which follows easily from Chernoff's bound, see, e.g., [23, pp. 43].

Lemma 3.4. *Let* $W_m \stackrel{\mathcal{L}}{=} \operatorname{Bin}(m, p_m)$*. If* $\ell_m p_m \to c \in (0, \infty)$ *and* $m/\ell_m \to \infty$*, then* $\ell_m W_m / m \xrightarrow{p} c$*.*

Proof of Theorem 3.1. As in the proof of Theorem 3.3, we condition on $\mathcal{F} = \sigma((U_j)_{j\geq 1})$ and treat $(U_j, P_{n,j}, L^*_{n,j})_{j\geq 0, n\geq 1}$ as deterministic numbers. By (3.2), conditioning on $(S^*_{n,1}, \dots, S^*_{n,q})$, $B^*_{n,1}, \dots, B^*_{n,q}$ are independent and for $j \in \{1, \dots, q\}$,

$$B^*_{n,j} | (S^*_{n,1}, \dots, S^*_{n,q}) \stackrel{\mathcal{L}}{=} \mathrm{Bin}\left(P_{n,j-1} - P_{n,j}, \frac{S^*_{n,j}}{L^*_{n,j}}\right).$$

It follows from (3.5) and Theorem 3.4 that

$$\left.\frac{B^*_{n,j}}{n^{1-\frac{1}{k}}}\right| (S^*_{n,1}, \dots, S^*_{n,q}) \xrightarrow{p} (1 - U_j)\left(\prod_{1\leq s<j} U_s\right)^{1-\frac{1}{k}} S^*_{n,j}.$$

Now by Theorem 3.3, the joint density function of $(S^*_{n,1}, \dots, S^*_{n,q})$ converges point-wise to that of $(S_1, \dots, S_q)$. Therefore, jointly, conditioning on $\mathcal{F} = \sigma((U_j)_{j\geq 1})$,

$$\mathcal{L}\left(\left(\frac{B^*_{n,1}}{n^{1-\frac{1}{k}}}, \dots, \frac{B^*_{n,q}}{n^{1-\frac{1}{k}}}\right)\right) \xrightarrow{d} \mathcal{L}\left((B_1, \dots B_q)\right),$$

where (see (1.4) and (1.3)) $B_j \stackrel{\text{def}}{=} (1 - U_j)\left(\prod_{1\leq s<j} U_s\right)^{1-\frac{1}{k}} S_j$. Thus, the convergence also holds without conditioning on $\mathcal{F}$. □

3.2 The leftovers

In this section, we show that for q large enough, $\sum_{s>q} B_s$, $\sum_{s>q} B^*_{n,s}/n^{1-\frac{1}{k}}$, and $\sum_{s>q} B_{n,s}/n^{1-\frac{1}{k}}$ are all negligible.

Lemma 3.5. *For all $k \in \{2, 3, \dots\}$, $\varepsilon > 0$ and $\delta > 0$, there exists $q \in \mathbb{N}$ and $n_0 \in \mathbb{N}$ such that for all $n > n_0$,*

$$\mathbf{P}\left\{\sum_{q<s} B_s \geq \varepsilon\right\} < \delta, \qquad \mathbf{P}\left\{\frac{\sum_{j>q} B_{n,j}}{n^{1-\frac{1}{k}}} \geq \varepsilon\right\} < \delta, \qquad \mathbf{P}\left\{\frac{\sum_{j>q} B_{n,j}}{n^{1-\frac{1}{k}}} \geq \varepsilon\right\} < \delta.$$

Proof. We only give the proof for $\sum_{s>q} B_s$, since the other two can be dealt essentially in the same way.

Let $U'_1, E'_1, U'_2, E'_2, \dots$ be independent random variables such that $U'_j \stackrel{\mathcal{L}}{=} \mathrm{Unif}[0,1]$ and $E'_j \stackrel{\mathcal{L}}{=} \mathrm{Exp}(1)$. By the definition of B_s (see (1.4) and (1.3)), we have

$$B_s \preceq B'_s \stackrel{\text{def}}{=} \left(\left(\prod_{1\leq j\leq s} U'_j\right)\left(k! \sum_{1\leq j\leq s} E'_s\right)\right)^{1/k},$$

i.e., B_s is stochastically dominated by B'_s. Thus, we can prove the lemma for B'_s instead. Let W_s and W'_s be independent $\mathrm{Gamma}(s)$ random variables. Then

$$-\log\left(\prod_{1\leq j\leq s} U'_j\right) \stackrel{\mathcal{L}}{=} W_s, \qquad \sum_{1\leq j\leq s} E'_j \stackrel{\mathcal{L}}{=} W'_s.$$

It is well known that $\mathbf{E}\left[(W_s - s)^4\right] = 3s^2 + 6s$ [19, pp. 339]. It follows from Markov's inequality that for $s \geq 1$,

$$\mathbf{P}\left\{|W_s - s| \geq \frac{s}{2}\right\} \leq \frac{\mathbf{E}\left[(W_s - s)^4\right]}{s^4/16} = \frac{3s^2 + 6s}{s^4/16} = \frac{9s^2}{s^4/16} \leq \frac{144}{s^2}.$$

Therefore

$$\mathbf{P}\left\{(B'_s)^k \geq k!\frac{3}{2}s\mathrm{e}^{-s/2}\right\} \leq \mathbf{P}\left\{\prod_{1\leq j\leq s} U'_j \geq \mathrm{e}^{-s/2}\right\} + \mathbf{P}\left\{\sum_{1\leq j\leq s} E'_j \geq \frac{3}{2}s\right\}$$
$$= \mathbf{P}\left\{W_s \leq \frac{s}{2}\right\} + \mathbf{P}\left\{W'_s \geq \frac{3s}{2}\right\} = O\left(\frac{1}{s^2}\right).$$

We are done since

$$\sum_{s>q}\frac{1}{s^2} = O(q^{-1}), \qquad \sum_{s>q}\left(k!\frac{3}{2}s\mathrm{e}^{-s/2}\right)^{\frac{1}{k}} = O\left(\mathrm{e}^{-\frac{q}{4k}}\right). \qquad \square$$

3.3 Finishing the proof Theorem of 1.5

By Theorem 3.5, the contribution of $\sum_{s>q} B_s$ and $\sum_{s>q} B^*_{n,s}/n^{1-\frac{1}{k}}$ in $\sum_{s>1} B_s$ and $\sum_{s>1} B^*_{n,s}/n^{1-\frac{1}{k}}$ respectively can be made arbitrarily small by choosing q large enough. Thus, it follows from Theorem 3.1 that $\mathcal{L}\left(\mathcal{X}^*_n/n^{1-\frac{1}{k}}\right) \xrightarrow{d} \mathcal{L}\left(\mathcal{B}_k\right)$ as we claimed.

Now we fill the gap between $\mathcal{X}^*_n$ and $\mathcal{X}_{n,1}$ by the following lemma, from which Theorem 1.5 follows immediately.

Lemma 3.6. *Let $k \in \{2, 3, \dots\}$. There exists a coupling such that*

$$\frac{\mathcal{X}^*_n - \mathcal{X}_{n,1}}{n^{1-\frac{1}{k}}} \xrightarrow{p} 0.$$

Proof. Recall that $(T^*_{i,j})_{i\geq 1, j\geq 1}$ are i.i.d. $\mathrm{Exp}(1)$ random variables that we used, together with $(P_{n,j}, R_{n,j})_{j\geq 0}$ to define $\mathcal{X}^*_n$. Now we modify $(T_{i,j})_{i\geq 1,j\geq 1}$ by letting $T_{i,j} = T^*_{i,j}$ for all $i \in \mathbb{N}$ and $j \notin \{P_{n,j}\}_{j\geq 0}$, unless there is a *discrepancy*, i.e., if for some $q \geq 1$,

$$P_{n,q-1} < j < P_{n,q}, \qquad \text{and} \qquad \sum_{i=1}^{k} T^*_{j,i} < R_{n,q}.$$

This is a valid coupling since it does not change the distribution of $(B_{n,j})_{j\geq 1}$.

Let $J_{n,q}$ denote the number of discrepancies between $P_{n,q-1}$ and $P_{n,q}$, i.e.,

$$J_{n,q} = \sum_{j\geq 1} [\![P_{n,q-1} < j < P_{n,q}]\!][\![R_{n,q} > \sum_{1\leq i\leq k} T^*_{i,j}]\!].$$

By the definition (3.1) and (3.3), with the above coupling, for all fixed $q \in \mathbb{N}$,

$$|\mathcal{X}_{n,1} - \mathcal{X}^*_n| \leq \sum_{1\leq j\leq q} J_{n,j} + 2\mathcal{X}_{n,k} + \sum_{j>q} B_{n,q} + \sum_{j>q} B^*_{n,q}. \tag{3.7}$$

By Theorem 1.1, we have $\mathcal{X}_{n,k}/n^{1-\frac{1}{k}} \xrightarrow{p} 0$. It follows from Theorem 3.5 that by choosing q large enough, the last two terms of the right-hand-side of (3.7) divided by $n^{1-\frac{1}{k}}$ are negligible. Thus, it suffices to show that $\sum_{1\leq j\leq q} J_{n,j}/n^{1-1/k} \xrightarrow{p} 0$.

Conditioning on $(R_{n,j}, P_{n,j})_{n\geq 1, j\geq 0}$,

$$J_{n,q} \stackrel{\mathcal{L}}{=} \mathrm{Bin}\left(P_{n,q-1} - P_{n,q} - 1, \mathbf{P}\left\{G_k < R_{n,q}\right\}\right),$$

where $G_k \stackrel{\mathcal{L}}{=} \mathrm{Gamma}(k)$. Therefore, it follows from the series expansion of the incomplete gamma function [24, 8.7.3] that

$$\mathbf{E}\left[J_{n,q} \mid (R_{n,j}, P_{n,j})_{n\geq 1, j\geq 0}\right] \leq (P_{n,q-1} - P_{n,q}) \cdot \left(1 - \frac{\Gamma(k, R_{n,q})}{\Gamma(k)}\right)$$

$$\leq P_{n,q-1} R_{n,q}^k = \frac{P_{n,q-1}}{(L_{n,q}^*)^k} (S_{n,q})^k \xrightarrow{d} \mathcal{L}\left(S_q^k\right),$$

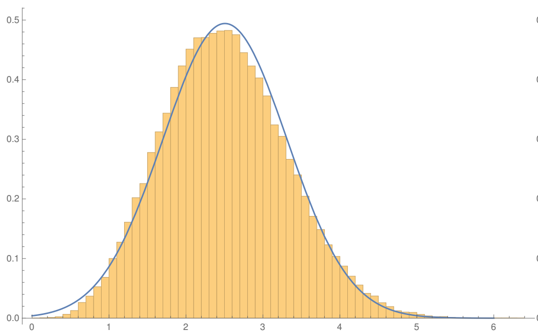
where the convergence follows from (3.5) and Theorem 3.3. By the definition (1.3), $S_q^k \preceq k! G_q$. Thus for all fixed $q \in \mathbb{N}$, $\sup_{n \geq 1} \mathbf{E}\left[J_{n,q}\right] < \infty$ and $\sum_{1 \leq i \leq q} J_{n,i}/n^{1-\frac{1}{k}} \xrightarrow{p} 0$. $\square$

3.4 The density of $\mathcal{B}_k$

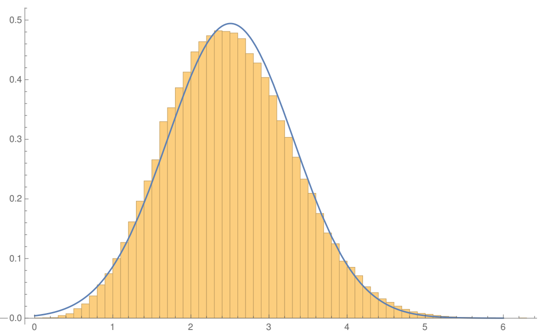
Lemma 3.7. *For all $k \in \{2, 3, \dots\}$ the random variable $\mathcal{B}_k$ defined in* (1.5) *has a density function.*

Proof. The random variable $\mathcal{B}_k$ can be written as $\sum_{1 \leq q} (a(q) + b(q) E_1)^{1/k}$, where $a(q)$ and $b(q)$ are non-negative functions of the random vector $(U_1, U_2, E_2, U_3, E_3, \dots)$. Conditioning on this vector, $\mathcal{B}_k$ has a density provided that $b(q) \neq 0$ for some q. Thus, a sufficient condition for $\mathcal{B}_k$ to have a density is that $\mathbf{P}\left\{b(1) = 0\right\} = 0$, which is obvious since $b(1) = (1 - U_1)^{1/k} k!$ $\square$

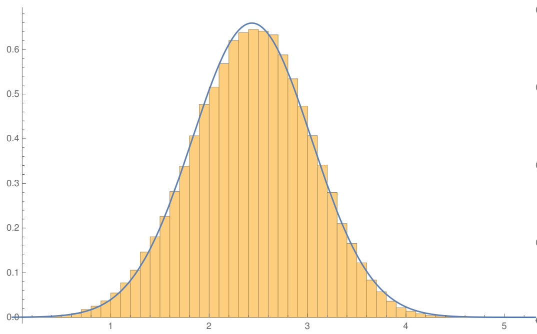
It is not easy to see what the density function of $\mathcal{B}_k$ should be like analytically. But through simulation, it is obvious that $\mathcal{B}_k$ has a density very close to that of the normal distribution $\mathcal{N}(\mathbf{E}\mathcal{B}_k, \sqrt{\mathrm{Var}\left(\mathcal{B}_k\right)})$, see Figure 3. Comparing Figure 3a with the simulation result for $\mathcal{X}_n$ with $k = 2$ shown in Figure 4, we see that $\mathcal{B}_k$ is indeed the limit distribution of $\mathcal{X}_n$.



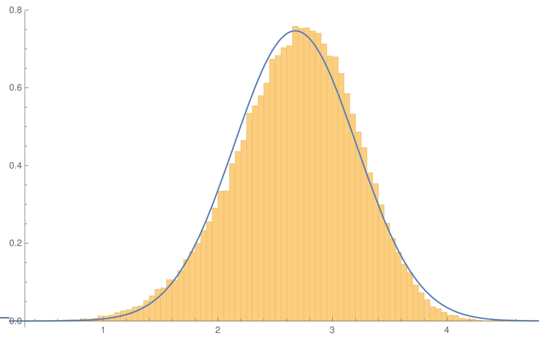
(a) $k = 2$



(b) $k = 3$



(c) $k = 4$



(d) $k = 5$

Figure 3: Histograms of 10^5 samples of $\mathcal{B}_k$ for $k = 2, \dots, 5$. The blue curves represent the density functions of $\mathcal{N}(\mathbf{E}\mathcal{B}_k, \sqrt{\mathrm{Var}\left(\mathcal{B}_k\right)})$.

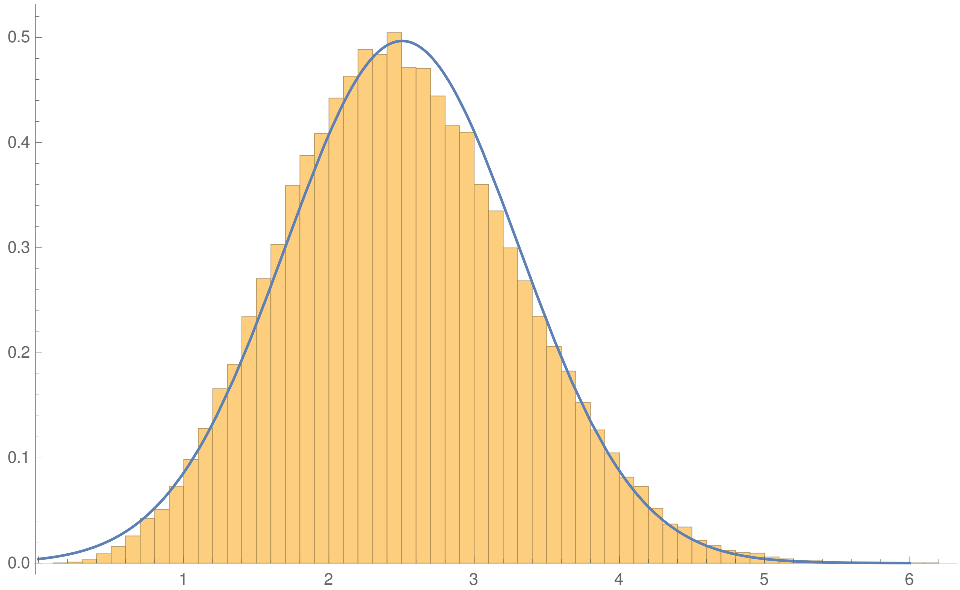


Figure 4: Simulation for $\mathcal{X}_n$ with $k = 2$, $n = 2^{17}$ and 60000 samples, after rescaled by $\sqrt{n}$. The blue curve represents the density function of a normal distribution with the empirical mean and variance.

4 Some extensions

4.1 A lower bound and an upper bound for general graphs

Let $\mathcal{G}_n$ be the set of rooted graphs with n nodes. It is obvious that $\mathbb{P}_n$ is the easiest to cut among all graphs in $\mathcal{G}_n$. We formalize this by the following lemma:

Lemma 4.1. *Let $k \in \mathbb{N}$. For all $\mathbb{G}_n \in \mathcal{G}_n$, $\mathcal{X}_n \stackrel{\text{def}}{=} \mathcal{K}(\mathbb{P}_n) \preceq \mathcal{K}(\mathbb{G}_n)$. Therefore,*

$$\min_{\mathbb{G}_n \in \mathcal{G}_n} \mathbf{E}\mathcal{K}(\mathbb{G}_n) \geq \mathbf{E}\mathcal{X}_n \sim \begin{cases} \dfrac{(k!)^{\frac{1}{k}}}{k-1} \Gamma\left(\dfrac{1}{k}\right) n^{1-\frac{1}{k}} & (k \geq 2), \\ \log n & (k = 1), \end{cases}$$

by Theorem 1.1.

The most resilient graph is obviously $\mathbb{K}_n$, the complete graph with n vertices. Thus, we have the following upper bound:

Lemma 4.2. *Let $k \in \mathbb{N}$.*

(i) Let $Y \stackrel{\mathcal{L}}{=} \mathrm{Gamma}(k)$, $Z \stackrel{\mathcal{L}}{=} \mathrm{Poi}(Y)$, and $W \stackrel{\mathcal{L}}{=} Z \wedge k$, i.e., $W \stackrel{\mathcal{L}}{=} \min\{Z, k\}$. Then

$$\mathcal{L}\left(\frac{\mathcal{K}(\mathbb{K}_n)}{n}\right) \xrightarrow{d} \mathcal{L}\left(\mathbf{E}\left[W|Y\right]\right) = \mathcal{L}\left(\frac{\Gamma(k+1, Y) - e^{-Y} Y^{k+1}}{k!} + k\right), \tag{4.1}$$

where $\Gamma(\ell, z)$ denotes the upper incomplete gamma function. Note that when $k = 1$, the right-hand-side is simply $\mathrm{Unif}[0, 1]$.

(ii) For all $\mathbb{G}_n \in \mathcal{G}_n$, $\mathcal{K}(\mathbb{G}_n) \preceq \mathcal{K}(\mathbb{K}_n)$. Therefore,

$$\max_{\mathbb{G}_n \in \mathcal{G}_n} \mathbf{E}\mathcal{K}(\mathbb{G}_n) \leq \mathbf{E}\mathcal{K}(\mathbb{K}_n) \sim k\left(1 - \frac{1}{2^{2k}}\binom{2k}{k}\right) n. \tag{4.2}$$

Proof. Let S_n be the tree of n nodes with one root and $n - 1$ leaves. Obviously $\mathcal{K}(\mathbb{K}_n) \stackrel{\mathcal{L}}{=} \mathcal{K}(S_n)$. Let Y be the time when the root is removed. Let $W_{1,n}, \ldots, W_{n-1,n}$

be the number of cuts each leaf receives by this time. Conditioning on the event $Y = y$, $W_{1,n}, \dots, W_{n-1,n}$ are i.i.d. with $W_{i,n} \overset{\mathcal{L}}{=} Z_i \wedge k$, where $Z_i \overset{\mathcal{L}}{=} \mathrm{Poi}(y)$. In other words, conditioning on $Y = y$, by the law of large numbers,

$$\frac{\mathcal{K}(S_n)}{n} = \frac{k + \sum_{i=1}^{n-1} W_{i,n}}{n} \overset{a.s.}{\to} \mathbf{E}\left[Z_1\right],$$

from which (4.1) and (4.2) follow immediately. □

4.2 Path-like graphs

If a graph $\mathbb{G}_n$ consists of only long paths, then the limit distribution $\mathcal{K}(\mathbb{G}_n)$ should be related to $\mathcal{B}_k$, the limit distribution of $\mathcal{K}(\mathbb{P}_n)/n^{1-\frac{1}{k}}$ (see Theorem 1.5). We give two simple examples with $k \in \{2, 3, \dots\}$.

Example 4.3 (Long path)**.** Let $(\mathbb{G}_n)_{n\ge1}$ be a sequence of rooted graphs such that $\mathbb{G}_n$ contains a path of length $m(n)$ starting from the root with $n - m(n) = o(n^{1-\frac{1}{k}})$. Since it takes at most $k(n - m(n))$ cuts to remove all the nodes outside the long path,

$$\mathcal{K}(P_{m(n)}) \preceq \mathcal{K}(\mathbb{G}_n) \preceq \mathcal{K}(P_{m(n)}) + ko\Big(n^{1-1/k}\Big).$$

Thus, by Theorem 4.1, this implies that $\mathcal{K}(\mathbb{G}_n)/n^{1-\frac{1}{k}}$ converges in distribution to $\mathcal{B}_k$.

Example 4.4 (Curtain)**.** Let $\ell \ge 2$ be a fixed integer. Let $\mathbb{T}_n^{(\ell)}$ be a graph that consists of only ℓ paths connected to the root, with the first $\ell - 1$ of them having length $\left\lceil \frac{n-1}{\ell} \right\rceil$. We call $\mathbb{T}_n^{(\ell)}$ an ℓ-curtain. It is easy to see that cutting $\mathbb{T}_n^{(\ell)}$ is very similar to cutting ℓ separated paths of length $\left\lceil \frac{n}{\ell} \right\rceil$. Therefore, we can show that

$$\mathcal{L}\left(\frac{\mathcal{K}(\mathbb{T}_n^{(\ell)})}{(n/\ell)^{\frac{1}{k}}}\right) \overset{d}{\to} \mathcal{L}\left(\sum_{j=1}^{\ell} \mathcal{B}_k^{[j]},\right)$$

where $\mathcal{B}_k^{[1]}, \dots, \mathcal{B}_k^{[\ell]}$ are i.i.d. copies of $\mathcal{B}_k$.

4.3 Deterministic and random trees

The approximation given in Theorem 2.1 can be used to compute the expectation of k-cut numbers in many deterministic or random trees. We give four examples: complete binary trees, split trees, random recursive trees, and Galton-Watson trees.

4.3.1 Complete binary trees

Let $\mathbb{T}_n^{\mathrm{bi}}$ be a complete binary tree of with $n = 2^{m+1} - 1$ nodes, i.e., its height is m. Recall that $I_{r,i+1}$ in Theorem 2.1 is the indicator that a node in $\mathbb{P}_n$ at depth i is an r-record. Since the probability of a node being an r-record only depends on its depth, it follows from Theorem 2.1 that

$$\mathbf{E}\mathcal{K}_r(\mathbb{T}_n^{\mathrm{bi}}) = \sum_{i=0}^{m} 2^i \mathbf{E} I_{r,i+1} \sim \frac{(k!)^{\frac{r}{k}}}{k} \frac{\Gamma\left(\frac{r}{k}\right)}{\Gamma(r)} \frac{2^{m+1}}{m^{\frac{r}{k}}}.$$

Thus, only the one-records matter as in the case of $\mathbb{P}_n$ and

$$\mathbf{E}\mathcal{K}(\mathbb{T}_n^{\mathrm{bi}}) \sim \mathbf{E}\mathcal{K}_1(\mathbb{T}_n^{\mathrm{bi}}) \sim \frac{(k!)^{\frac{1}{k}}\Gamma\left(\frac{1}{k}\right)}{k} \frac{2^{m+1}}{m^{\frac{1}{k}}} \sim \frac{(k!)^{\frac{1}{k}}\Gamma\left(\frac{1}{k}\right)}{k} \frac{n}{(\log_2 n)^{\frac{1}{k}}}.$$

The limit distribution of $\mathcal{K}(\mathbb{T}_n^{\mathrm{bi}})$ has been found in our follow-up paper [3].

4.3.2 Split trees

Split trees were first defined by Devroye [7] to encompass many families of trees that are frequently used in algorithm analysis, e.g., binary search trees and tries. Its exact construction is somewhat lengthy and we refer readers to either the original algorithmic definition in [14] or the more probabilistic version in [4, Section 2].

Very roughly speaking, $\mathbb{T}_n^{\mathrm{sp}}$ is constructed by first distributing randomly n balls among the nodes of an infinite b-ary tree and then removing all subtrees without balls. Each node in the infinite b-ary tree is given a random non-negative split vector $\mathcal{V} = (V_1, \dots, V_b)$, satisfying $\sum_{i=1}^{b} V_i = 1$, drawn independently from the same distribution. These vectors affect how balls are distributed.

In the study of split trees, the following condition of $\mathcal{V}$ is often assumed:

Condition 4.5. *The split vector $\mathcal{V}$ is permutation invariant. Moreover, $\mathbf{P}\{V_1 = 1\} = 0$, $\mathbf{P}\{V_1 = 0\} = 0$, and that $-\log(V_1)$ is non-lattice.*

Holmgren[14, Theorem 1.1] showed that , assuming condition 4.5, there exists a constant α such that $\mathbf{E}N \sim \alpha n$, where N is the random number of nodes in $\mathbb{T}_n^{\mathrm{sp}}$.

In the setup of split trees (and other random trees), we obtain $\mathcal{K}(\mathbb{T}_n^{\mathrm{sp}})$ by choosing a random split tree first and then carry out the k-cut process conditioning on the tree. Holmgren [13, Theorem 1.1] showed that condition 4.5 implies that $\mathcal{K}_k(\mathbb{T}_n^{\mathrm{sp}})$ converges to a weakly 1-stable distribution after normalization, and that $\mathbf{E}\mathcal{K}_k(\mathbb{T}_n^{\mathrm{sp}}) \sim \mu\alpha n/\log n$, where $\mu \stackrel{\mathrm{def}}{=} b\mathbf{E}[-V_1 \log V_1]$. We extend this result as follows:

Lemma 4.6. *Assuming condition 4.5, we have*

$$\mathbf{E}[\mathcal{K}_r(\mathbb{T}_n^{\mathrm{sp}})] \sim \frac{(k!\mu)^{\frac{r}{k}}}{k} \frac{\Gamma\left(\frac{r}{k}\right)}{\Gamma(r)} \frac{\alpha n}{(\log n)^{\frac{r}{k}}}, \qquad (1 \le r \le k),$$
$$\mathbf{E}[\mathcal{K}(\mathbb{T}_n^{\mathrm{sp}})] \sim (k!\mu)^{\frac{1}{k}} \frac{\Gamma\left(\frac{1}{k}\right)}{k} \frac{\alpha n}{(\log n)^{\frac{1}{k}}}.$$

Proof. We say a node v is *good* if it has depth $d(v)$ where $\left|d(v) - \frac{1}{\mu}\log n\right| \le \log^{0.6} n$, otherwise we say it is *bad*. Let $\mathbb{B}_n^{\mathrm{sp}}$ be the number of bad nodes in $\mathbb{T}_n^{\mathrm{sp}}$. By [14][Theorem 1.2], $\mathbf{E}\mathbb{B}_n^{\mathrm{sp}} = O\big(n/(\log n)^3\big)$. Thus, the number of r-records in bad nodes is negligible and it suffices to prove the lemma for good nodes. By Lemma 2.1 and the definition of good nodes, we have

$$\begin{aligned}\mathbf{E}[\mathcal{K}(\mathbb{T}_n^{\mathrm{sp}})|\mathbb{T}_n^{\mathrm{sp}}] =& (N - \mathbb{B}_n^{\mathrm{sp}}) \frac{(k!)^{\frac{r}{k}}}{k} \frac{\Gamma\left(\frac{r}{k}\right)}{\Gamma(r)} \left[\frac{\log n}{\mu} + O(\log^{0.6} n)\right]^{-\frac{r}{k}} (1 + O(\log^{-\frac{1}{2k}} n)) \\ =& (N - \mathbb{B}_n^{\mathrm{sp}}) \frac{(k!\mu)^{\frac{r}{k}}}{k} \frac{\Gamma\left(\frac{r}{k}\right)}{\Gamma(r)} \frac{1}{(\log n)^{\frac{r}{k}}} (1 + O(\log^{-\frac{1}{2k}} n)),\end{aligned}$$

from which the lemma follows by taking expectation and using that $\mathbf{E}N \sim \alpha n$. □

4.3.3 Random recursive trees

A random recursive tree $\mathbb{T}_n^{\mathrm{rr}}$ is random tree of n nodes constructed recursively as follows: let $\mathbb{T}_1^{\mathrm{rr}}$ be the tree of a single node labeled 1; given $\mathbb{T}_{n-1}^{\mathrm{rr}}$, choose a node in $\mathbb{T}_{n-1}^{\mathrm{rr}}$ uniformly at random and attach a node labeled n to the selected node as a child, which gives $\mathbb{T}_n^{\mathrm{rr}}$. Meir and Moon [22] introduced this model and showed that $\mathbf{E}\mathcal{K}_k(\mathbb{T}_n^{\mathrm{rr}}) \sim n/\log n$ and that $\mathcal{K}_k(\mathbb{T}_n^{\mathrm{rr}})$ concentrates around its mean. Drmota et al. [9] and subsequently Iksanov and Möhle [15] proved $\mathcal{K}(\mathbb{T}_n^{\mathrm{rr}})$ converges weakly to a stable law after proper shifting and normalization.

The intuition behind $\mathbf{E}\mathcal{K}_k(\mathbb{T}_n^{\mathrm{rr}}) \sim n/\log n$ is simply that almost all nodes in $\mathbb{T}_n^{\mathrm{rr}}$ are at depth around $\log n$. We say a node v in $\mathbb{T}_n^{\mathrm{rr}}$ is *good* if $|d(v) - \log(n)| \le \log(n)^{0.9}$; otherwise we say it is *bad*. The following lemma shows that there are very few bad nodes in expectation:

Lemma 4.7. *Let $\mathbb{B}_n^{\mathrm{rr}}$ be the number of bad nodes in $\mathbb{T}_n^{\mathrm{rr}}$, then $\mathbf{E}\mathbb{B}_n^{\mathrm{rr}} = O\big(n/\log(n)^3\big)$.*

Proof. Let $h(\mathbb{T}_n^{\mathrm{rr}})$ be the height of $\mathbb{T}_n^{\mathrm{rr}}$. By [8, 6.3.2]

$$\mathbf{P}\left\{|h(\mathbb{T}_n^{\mathrm{rr}}) - \mathrm{e}\log(n)| > \eta\right\} = O\big(\mathrm{e}^{-c\eta}\big),$$

for some constant c. Thus, we can choose some constant K large enough and ignore the nodes of depth greater than $K\log(n)$. Let $w_d(\mathbb{T}_n^{\mathrm{rr}})$ be the number of nodes at depth d in $\mathbb{T}_n^{\mathrm{rr}}$. By [11, Equation 3]

$$\mathbf{E}\left[w_d(\mathbb{T}_n^{\mathrm{rr}})\right] = \frac{\log(n)^d}{\Gamma(1 + d/\log(n))d!}\left(1 + O\big(\log(n)^{-1}\big)\right), \tag{4.3}$$

uniformly for all $n \ge 3$ and $1 \le d \le K\log(n)$, for all $K \ge 1$. Thus, the lemma follows by summing both sides of (4.3) over integers d in $[1, \log(n) - \log(n)^{0.9}] \cup [\log(n) + \log(n)^{0.9}, K\log(n)]$. □

Thus, by exactly the same argument of Theorem 4.6, we get:

Lemma 4.8. *We have*

$$\begin{aligned}
\mathbf{E}\left[\mathcal{K}_r(\mathbb{T}_n^{\mathrm{rr}})\right] &\sim \frac{(k!)^{\frac{r}{k}}}{k}\frac{\Gamma\left(\frac{r}{k}\right)}{\Gamma(r)}\frac{n}{(\log n)^{\frac{r}{k}}}, && (1 \le r \le k),\\
\mathbf{E}\left[\mathcal{K}(\mathbb{T}_n^{\mathrm{rr}})\right] &\sim (k!)^{\frac{1}{k}}\frac{\Gamma\left(\frac{1}{k}\right)}{k}\frac{n}{(\log n)^{\frac{1}{k}}}.
\end{aligned}$$

Remark 4.9. $\mathbb{T}_n^{\mathrm{sp}}$ and $\mathbb{T}_n^{\mathrm{rr}}$ are both of logarithmic height. Thus, the same method which we used for treating complete binary trees [3] should also work for them.

4.3.4 Conditional Galton-Watson trees

A Galton-Watson tree $\mathbb{T}^{\mathrm{gw}}$ is a random tree that starts with the root node and recursively attaches a random number of children to each node in the tree, where the numbers of children are drawn independently from the same distribution $\mathcal{L}(\xi)$ (the offspring distribution). A conditional Galton-Watson tree $\mathbb{T}_n^{\mathrm{gw}}$ is $\mathbb{T}^{\mathrm{gw}}$ restricted to size n. See [18] for a comprehensive survey of conditional Galton-Watson trees.

Janson [17, Theorem 1.6] showed that $\mathcal{K}_k(\mathbb{T}_n^{\mathrm{gw}})/\sqrt{n}$ converges weakly to a Rayleigh distribution and the convergence is also in all moments if ξ has a finite exponential moment. In particular

$$\frac{\mathbf{E}\mathcal{K}_k(\mathbb{T}_n^{\mathrm{gw}})}{\sqrt{n}} \to \mathbf{E}\left[\int_0^1 \left(\frac{2e(t)}{\sigma}\right)^{-1} \mathrm{d}t\right] = \sigma\sqrt{\frac{\pi}{2}},$$

where $e(t)$ denotes a normalized Brownian excursion and $\sigma^2 = \mathrm{Var}\,(\xi)$. It is straight forward to adapt the method in [17] to get the first moment of $\mathbf{E}\mathcal{K}_r(\mathbb{T}_n^{\mathrm{gw}})$. (Though higher moments and the limit distribution seems to be elusive.) We formulate this as lemma and refer the reader to [17] for details.

Lemma 4.10. *Assume that $\mathbf{E}\left[\xi^3\right] < \infty$. Then for $r \in \{1, \ldots, k\}$,*

$$\frac{\mathbf{E}\mathcal{K}_r(\mathbb{T}_n^{\mathrm{gw}})}{n^{1-\frac{r}{2k}}} \to \frac{(k!)^{\frac{r}{k}}}{k}\frac{\Gamma(\frac{r}{k})}{\Gamma(r)}\mathbf{E}\left[\int_0^1 \left(\frac{2e(t)}{\sigma}\right)^{-\frac{r}{k}} \mathrm{d}t\right] = \frac{(k!)^{\frac{r}{k}}}{k}\frac{\Gamma(\frac{r}{k})\Gamma\left(1 - \frac{r}{2k}\right)}{\Gamma(r)}\left(\frac{\sigma}{\sqrt{2}}\right)^{\frac{r}{k}}.$$

As a result,

$$\mathbf{E}\mathcal{K}(\mathbb{T}_n^{\mathrm{gw}}) \sim \mathbf{E}\mathcal{K}_1(\mathbb{T}_n^{\mathrm{gw}}) \sim \frac{(k!)^{\frac{1}{k}}}{k}\Gamma\left(\frac{1}{k}\right)\Gamma\left(1-\frac{1}{2k}\right)\left(\frac{\sigma}{\sqrt{2}}\right)^{\frac{1}{k}} n^{1-\frac{1}{2k}}.$$

5 Some auxiliary results

Lemma 5.1. *Let* $G_k \overset{\mathcal{L}}{=} \mathrm{Gamma}(k)$*. Let* $\alpha \overset{\mathrm{def}}{=} \frac{1}{2}\left(\frac{1}{k}+\frac{1}{k+1}\right)$ *and* $x_0 \overset{\mathrm{def}}{=} m^{-\alpha}$*. Then uniformly for all* $x \in [0, x_0]$*,*

$$\mathbf{P}\left\{G_k > x\right\}^m = \left(\frac{\Gamma\left(k,x\right)}{\Gamma(k)}\right)^m = \left(1+O\left(m^{-\frac{1}{2k}}\right)\right)\exp\left(-\frac{mx^k}{k!}\right),$$

where $\Gamma(\ell, z)$ *denotes the upper incomplete gamma function.*

Proof. By the density function of gamma distributions, $\mathbf{P}\left\{G_k > x\right\} = \Gamma\left(k,x\right)/\Gamma(k)$. It then follows from the series expansion of the incomplete gamma function [24, 8.7.3], that uniformly for all $x \le x_0$,

$$\left(\frac{\Gamma\left(k,x\right)}{\Gamma(k)}\right)^m = \left(1-\frac{x^k}{k!}+O\left(x_0^{k+1}\right)\right)^m = \left(1+O\left(m^{-\frac{1}{2k}}\right)\right)\exp\left(-\frac{mx^k}{k!}\right), \tag{5.1}$$

where we use that $-\alpha(k+1)+1 = -\frac{1}{2k}$. □

Lemma 5.2. *Let* $G_k \overset{\mathcal{L}}{=} \mathrm{Gamma}(k)$*. Let* $a \ge 0$ *and* $b \ge 1$ *be fixed. Then uniformly for* $m \ge 1$*,*

$$\int_0^\infty x^{b-1}\mathrm{e}^{-ax}\mathbf{P}\left\{G_k > x\right\}^m \,\mathrm{d}x = \left(1+O\left(m^{-\frac{1}{2k}}\right)\right)\frac{(k!)^{\frac{b}{k}}}{k}\Gamma\left(\frac{b}{k}\right)m^{-\frac{b}{k}}. \tag{5.2}$$

Proof. By Theorem 5.1, the left-hand-side of (5.2) equals

$$\int_0^{x_0} x^{b-1}\mathrm{e}^{-ax}\left(\frac{\Gamma\left(k,x\right)}{\Gamma(k)}\right)^m \mathrm{d}x + \int_{x_0}^\infty x^{b-1}\mathrm{e}^{-ax}\left(\frac{\Gamma\left(k,x\right)}{\Gamma(k)}\right)^m \mathrm{d}x \overset{\mathrm{def}}{=} A_1 + A_2,$$

where $x_0 = m^{-\alpha}$ and $\alpha = \frac{1}{2}\left(\frac{1}{k}+\frac{1}{k+1}\right)$. Then

$$\begin{aligned} A_1 &= \left(1+O\left(m^{-\frac{1}{2k}}\right)\right)\int_0^{x_0} x^{b-1}\mathrm{e}^{-ax}\exp\left(-\frac{mx^k}{k!}\right)\mathrm{d}x \\ &= \left(1+O\left(m^{-\frac{1}{2k}}\right)\right)\frac{(k!)^{\frac{b}{k}}}{k}\left(\Gamma\left(\frac{b}{k}\right)-\Gamma\left(\frac{b}{k},w_0\right)\right)m^{-\frac{b}{k}}, \end{aligned}$$

where $w_0 = \frac{mx_0^k}{k!} = \Theta(m^{\frac{1}{2k(k+1)}})$. By the upper bound given in [24, 8.11.i], $\Gamma\left(\frac{b}{k}, w_0\right) = O\left(\mathrm{e}^{-\frac{w_0}{2}}\right)$, which is exponentially small and can be neglected. Using (5.1), one can verify that $A_2 = O\left(\mathrm{e}^{-\frac{w_0}{2}}\right)$ which can also be neglected. □

Lemma 5.3. *For* $a > 0$*,* $b > 0$ *and* $k \ge 2$*,*

$$\begin{aligned} \xi_k(a,b) &\overset{\mathrm{def}}{=} \int_0^\infty\int_y^\infty \mathrm{e}^{-ax^k/k!-by^k/k!}\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{\Gamma\left(\frac{2}{k}\right)}{k}\left(\frac{k!}{a}\right)^{\frac{2}{k}} F\left(\frac{2}{k},\frac{1}{k};1+\frac{1}{k};-\frac{b}{a}\right), \end{aligned} \tag{5.3}$$

where F denotes the hypergeometric function. In particular,

$$\xi_2(a,b) = \arctan\left(\sqrt{\frac{b}{a}}\right)(ab)^{-\frac{1}{2}}. \tag{5.4}$$

Proof. Changing to polar system by $x = r\cos(\theta)$ and $y = r\sin(\theta)$,

$$\begin{aligned}\xi_k(a,b) &= \int_0^{\pi/4}\int_0^{\infty}\exp\left[-r^k\left(a\frac{\cos(\theta)^k}{k!} + b\frac{\sin(\theta)^k}{k!}\right)\right] r\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \int_0^{\pi/4}\left(a\frac{\cos(\theta)^k}{k!} + b\frac{\sin(\theta)^k}{k!}\right)^{-\frac{2}{k}}\frac{\Gamma(\frac{2}{k})}{k}\,\mathrm{d}\theta \\ &= \frac{\Gamma(\frac{2}{k})}{k}\left(\frac{k!}{a}\right)^{\frac{2}{k}}\int_0^{\pi/4}\left(1+\frac{b}{a}\tan(\theta)^k\right)^{-\frac{2}{k}}\,\mathrm{d}\theta \\ &= \frac{\Gamma(\frac{2}{k})}{k^2}\left(\frac{k!}{a}\right)^{\frac{2}{k}}\int_0^{1}u^{\frac{1}{k}-1}\left(1+\frac{b}{a}u\right)^{-\frac{2}{k}}\,\mathrm{d}u,\end{aligned}$$

which equals the right-hand-side of (5.3) by [24, 15.6.1]. For (5.4), see [24, 15.4.3]. ☐

Lemma 5.4. *For $a > 0$, $b > 0$ and $k \ge 2$,*

$$(a+b)^{-\frac{2}{k}} \le \frac{k}{\Gamma\left(\frac{2}{k}\right)(k!)^{\frac{2}{k}}}\xi_k(a,b) \le a^{-\frac{2}{k}} + b^{-\frac{2}{k}}. \tag{5.5}$$

Moreover, $\xi_k(a,b)$ is monotonically decreasing in both a and b.

Proof. Let

$$\xi_k^*(a,b) \stackrel{\text{def}}{=} \frac{k}{\Gamma\left(\frac{2}{k}\right)(k!)^{\frac{2}{k}}}\xi_k(a,b) = (a+b)^{-\frac{2}{k}} F\left(\frac{2}{k},1;1+\frac{1}{k};\frac{b}{a+b}\right), \tag{5.6}$$

where we use [24, 15.8.1]. Let $\alpha_1 = \frac{k}{k+1}$. By [20, cor. 2], for $x \in (0,1)$,

$$(1-\alpha_1 x)^{-\frac{2}{k}} \le F\left(\frac{2}{k},1;1+\frac{1}{k};x\right) \le 1-\alpha_1+\alpha_1(1-x)^{-\frac{2}{k}}.$$

This together with (5.6) give us (5.5).

For monotonicity, using the derivative formula [24, 15.5.1], it is easy to verify that for $a > 0$ and $b > 0$ $\frac{\partial}{\partial a}\xi_k^*(a,b) < 0$ and $\frac{\partial}{\partial b}\xi_k^*(a,b) < 0$. ☐

Lemma 5.5. *For $k \ge 2$, let*

$$\lambda_k \stackrel{\text{def}}{=} \int_0^1\int_0^{1-s}\xi_k(s,t)\,\mathrm{d}t\,\mathrm{d}s.$$

Then

$$\lambda_k = \begin{cases} \dfrac{\pi\cot\left(\frac{\pi}{k}\right)\Gamma\left(\frac{2}{k}\right)(k!)^{\frac{2}{k}}}{2(k-2)(k-1)} & k > 2, \\ \dfrac{\pi^2}{4} & k = 2. \end{cases}$$

Proof. When $k = 2$, applying (5.4) and changing to the polar system by letting $s = (r\cot(\theta))^2$ and $t = (r\sin(\theta))^2$, we get

$$\lambda_2 = \int_0^1 \int_0^{1-s} \frac{\arctan\left(\sqrt{\frac{t}{s}}\right)}{\sqrt{st}} = \int_0^{\frac{\pi}{2}} \int_0^1 4r\theta \, \mathrm{d}r \, \mathrm{d}\theta = \frac{\pi^2}{4}.$$

For $k \geq 3$, by Theorem 5.3, it suffices to show that

$$\int_0^1 s^{-\frac{2}{k}} \int_0^{1-s} F\left(\frac{2}{k}, \frac{1}{k}; 1 + \frac{1}{k}; -\frac{t}{s}\right) \mathrm{d}t \, \mathrm{d}s = \frac{k\pi \cot\left(\frac{\pi}{k}\right)}{2(k-2)(k-1)}, \tag{5.7}$$

which is easily verifiable using Mathematica. A human proof can be derived using the series expansion of hypergeometric functions [24, 15.6.1]. □

Remark 5.6. In an attempt to prove Theorem 5.5, we discovered the following identity

$$\int_0^\infty (w+1)^{\frac{2}{k}-2} F\left(\frac{2}{k}, \frac{1}{k}; 1 + \frac{1}{k}; -w\right) \mathrm{d}w = \frac{\pi \cot\left(\frac{\pi}{k}\right)}{k-2}, \qquad (k \geq 3),$$

which we have not found in the literature. The proof follows from changing to polar system in the left-hand-side of (5.7) by letting $s = (r\cos(\theta))^k$ and $t = (r\sin(\theta))^k$.

References

[1] Louigi Addario-Berry, Nicolas Broutin, and Cecilia Holmgren, *Cutting down trees with a Markov chainsaw*, Ann. Appl. Probab. **24** (2014), no. 6, 2297–2339. MR-3262504

[2] M. Ahsanullah, *Record values–theory and applications*, University Press of America, Inc., Lanham, MD, 2004.

[3] Xing Shi Cai and Cecilia Holmgren, *Cutting resilient networks – complete binary trees*, arXiv e-prints (2018).

[4] Xing Shi Cai, Cecilia Holmgren, Svante Janson, Tony Johansson, and Fiona Skerman, *Inversions in Split Trees and Conditional Galton–Watson Trees*, Combinatorics, Probability and Computing **28** (2019), no. 3, 335–364 (en).

[5] K. N. Chandler, *The distribution and frequency of record values*, J. Roy. Statist. Soc. Ser. B. **14** (1952), 220–228. MR-0053463

[6] D. Dagon, G. Gu, C. P. Lee, and W. Lee, *A taxonomy of botnet structures*, Twenty-Third Annual Computer Security Applications Conference (ACSAC 2007), 2007, pp. 325–339.

[7] Luc Devroye, *Universal limit laws for depths in random trees*, SIAM J. Comput. **28** (1999), no. 2, 409–432. MR-1634354

[8] Michael Drmota, *Random Trees: An Interplay between Combinatorics and Probability*, Springer-Verlag.

[9] Michael Drmota, Alex Iksanov, Martin Moehle, and Uwe Roesler, *A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree*, Random Structures Algorithms **34** (2009), no. 3, 319–336. MR-2504401

[10] Rick Durrett, *Probability: Theory and examples*, fourth ed., Cambridge Series in Statistical and Probabilistic Mathematics, vol. 31, Cambridge University Press, Cambridge, 2010. MR-2722836

[11] Michael Fuchs, Hsien-Kuei Hwang, and Ralph Neininger, *Profiles of random trees: limit theorems for random recursive trees and binary search trees*, Algorithmica **46** (2006), no. 3-4, 367–407. MR-2291961

[12] Cecilia Holmgren, *Random records and cuttings in binary search trees*, Combin. Probab. Comput. **19** (2010), no. 3, 391–424. MR-2607374

[13] Cecilia Holmgren, *A weakly 1-stable distribution for the number of random records and cuttings in split trees*, Adv. in Appl. Probab. **43** (2011), no. 1, 151–177. MR-2761152

[14] Cecilia Holmgren, *Novel characteristic of split trees by use of renewal theory*, Electron. J. Probab. **17** (2012), no. 5, 27. MR-2878784

[15] Alex Iksanov and Martin Möhle, *A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree*, Electron. Comm. Probab. **12** (2007), 28–35. MR-2407414

[16] Svante Janson, *Random records and cuttings in complete binary trees*, Mathematics and Computer Science. III, Trends Math., Birkhäuser, Basel, 2004, pp. 241–253. MR-2090513

[17] Svante Janson, *Random cutting and records in deterministic and random trees*, Random Structures Algorithms **29** (2006), no. 2, 139–179. MR-2245498

[18] Svante Janson, *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation*, Probab. Surv. **9** (2012), 103–252. MR-2908619

[19] Norman L. Johnson, Samuel Kotz, and N. Balakrishnan, *Continuous univariate distributions. Vol. 1*, second ed., Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, Inc., New York, 1994, A Wiley-Interscience Publication. MR-1299979

[20] D. B. Karp, *Representations and inequalities for generalized hypergeometric functions*, J. Math. Sci. (N.Y.) **207** (2015), no. 6, 885–897. MR-3374217

[21] A. Meir and J. W. Moon, *Cutting down random trees*, J. Austral. Math. Soc. **11** (1970), 313–324. MR-0284370

[22] A. Meir and J.W. Moon, *Cutting down recursive trees*, Mathematical Biosciences **21** (1974), no. 3, 173 – 181.

[23] Michael Molloy and Bruce Reed, *Graph colouring and the probabilistic method*, Algorithms and Combinatorics, vol. 23, Springer-Verlag, Berlin, 2002. MR-1869439

[24] Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark (eds.), *NIST handbook of mathematical functions*, U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC; Cambridge University Press, Cambridge, 2010. MR-2723248

[25] Alfréd Rényi, *Théorie des éléments saillants d'une suite d'observations*, Ann. Fac. Sci. Univ. Clermont-Ferrand No. **8** (1962), 7–13. MR-0286162

Acknowledgments. We thank Henning Sulzbach and Svante Janson for helpful discussions.