PHASE CHANGES IN SUBTREE VARIETIES IN RANDOM RECURSIVE AND BINARY SEARCH TREES

QUNQIANG FENG*, HOSAM M. MAHMOUD†, AND ALOIS PANHOLZER‡

Abstract. We study the variety of subtrees lying on the fringe of recursive trees and binary search trees by analyzing the distributional behavior of $X_{n,k}$, which counts the number of subtrees of size k in a random tree of size n, with $k = k(n)$ dependent on n. Using analytic methods we can characterize for both tree families the phase change behavior of $X_{n,k}$ as follows. In the subcritical case, when $k(n)/\sqrt{n} \to 0$, we show that $X_{n,k}$ is (after normalization) asymptotically normally distributed, whereas in the supercritical case, when $k(n)/\sqrt{n} \to \infty$, $X_{n,k}$ converges to 0. In the critical case, when $k(n) = \Theta(\sqrt{n}\,)$, we show that if $k/\sqrt{n}$ approaches a limit then $X_{n,k}$ converges in distribution to a Poisson random variable, whereas if $k/\sqrt{n}$ does not approach a finite nonzero limit, the size oscillates and does not converge in distribution to any random variable. This provides for recursive trees and binary search trees an understanding of the complete spectrum of phases of $X_{n,k}$ and the gradual change from the subcritical to the supercritical phase.

Key words. random trees, recurrence, moments, Riccati equation

AMS subject classifications. 05C05, 60C05, 60F05

1. Introduction. The occurrence of patterns in random objects is an important area of modern research. The prime example is the interest one may have in the number of occurrences of a word in a given text, or the occurrence of words of a certain length in that text. Applications abound in linguistics where one wishes to analyze grammatical frequencies, or in genetics where one tries to identify genes in strands of DNA. The equivalent and equally important view in random trees is to find patterns (which are trees of a certain size or a certain shape) in a given tree generated randomly. This area has already attracted attention in the recent literature on random trees (see Flajolet, Gourdon, and Martínez, 1997, Chyzak, Drmota, Klausner and Kok, 2007+, and Feng, Mahmoud, and Su, 2007). In this paper we investigate this area for two of the most important random tree models. We look at the number of subtrees of a certain size on the fringe of random recursive trees, and random binary search trees.

The random recursive tree is a naturally growing structure that underlies many stochastic phenomena, such as contagion, and algorithms such as the Union-Find algorithm. For numerous applications of recursive trees we refer the reader to the survey in Smythe and Mahmoud (1994). The binary search tree is another naturally growing structure that underlies many algorithms, such as combinatorial sorting and searching algorithms and serves well as a data structure that supports fast retrieval of data. For numerous applications of binary search trees we refer the reader to Knuth (1998) or Mahmoud (2000).

2. Scope. We consider the variety of sizes that appear on the fringe of trees. For both classes of random recursive trees and binary search trees we analyze the random variable $X_{n,k}$, which counts the number of subtrees of size k in a random tree of size n, with both k fixed or $k = k(n)$ dependent on n. For k fixed it was

*Department of Statistics and Finance, University of Science and Technology of China, Hefei 230026, China (`fengqq@mail.ustc.edu.cn`).

†Department of Statistics, The George Washington University, Washington, D.C. 20052, U.S.A. (`hosam@gwu.edu`).

‡Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien, Wiedner Hauptstr. 8-10/104, 1040 Wien, Austria (`Alois.Panholzer@tuwien.ac.at`).

shown in Feng, Mahmoud, and Su (2007) for recursive trees by using the contraction method and modeling via Pólya urns, and in Flajolet, Gourdon, and Martínez (1997) for binary search trees by using analytic techniques, that $X_{n,k}$ is (after normalization) asymptotically normal distributed. Furthermore, in Feng, Mahmoud, and Su (2007) the three "phases" of $X_{n,k}$ could be identified for recursive tree by computing explicit formulas for the first two moments of $X_{n,k}$.

To show the characterization of the distributional behavior of $X_{n,k}$ for $k = k(n)$ dependent on n obtained in the present paper we utilize analytic methods based on functional equations for generating functions. Such methods proved to be very effective in combinatorial analysis. It is a well accepted folklore that when analytic methods are applicable, they provide fuller asymptotic expansion (hence better approximation) than purely probabilistic methods; contrast, for example, the very detailed asymptotic expansion in Panholzer and Prodinger (1998) with the probabilistic approach in Lent and Mahmoud (1996).

The result we establish for recursive trees is the following.

THEOREM 2.1. *Let $X_{n,k}$ be the number of subtrees of size k on the fringe of a random recursive tree of size n.*

(a) In the subcritical case, when $k/\sqrt{n} \to 0$,

$$\frac{X_{n,k} - \frac{n}{k(k+1)}}{\sqrt{\frac{(2k^2-1)n}{k(k+1)^2(2k+1)}}} \xrightarrow{D} \mathcal{N}(0,1).$$

(b) We are in the critical case when $k = \Theta(\sqrt{n})$. In the critical case, when $k/\sqrt{n} \to c > 0$,

$$X_{n,k} \xrightarrow{D} \mathrm{Poi}\Big(\frac{1}{c^2}\Big),$$

and if $k/\sqrt{n}$ does not converge to a limit, no limiting distribution exists for $X_{n,k}$.

(c) In the supercritical case, when $k/\sqrt{n} \to \infty$,

$$X_{n,k} \xrightarrow{D} 0.$$

The analytic method employed for recursive trees extends naturally to binary search trees, which we include in this paper, and possibly to other tree classes, too, which we leave for future research. The result we establish for binary search trees is the following.

THEOREM 2.2. *Let $X_{n,k}$ be the number of subtrees of size k on the fringe of a random binary search tree of size n.*

(a) In the subcritical case, when $k/\sqrt{n} \to 0$,

$$\frac{X_{n,k} - \frac{2n}{(k+1)(k+2)}}{\sqrt{\frac{2k(4k^2+5k-3)n}{(k+1)(k+2)^2(2k+1)(2k+3)}}} \xrightarrow{D} \mathcal{N}(0,1).$$

(b) *We are in the critical case when $k = \Theta(\sqrt{n}\,)$. In the critical case, when $k/\sqrt{n} \to c > 0$,*

$$X_{n,k} \xrightarrow{D} \mathrm{Poi}\Big(\frac{2}{c^2}\Big),$$

and if $k/\sqrt{n}$ does not converge to a limit, no limiting distribution exists for $X_{n,k}$.

(c) *In the supercritical case, when $k/\sqrt{n} \to \infty$,*

$$X_{n,k} \xrightarrow{D} 0.$$

We remark that $X_{n,k}$ can also be interpreted as the number of nodes in a random tree of size n, which have exactly k *descendants* (if we assume that a node itself is counted as its own descendant). From this point of view the "counterpart" of $X_{n,k}$ are the level polynomials $L_{n,k}$, i.e., the random variable that counts the number of nodes in a random tree of size n, which have exactly k *ascendants.* The distributional behavior $L_{n,k}$, i.e., the number of nodes at level k in a random tree of size n, has been analyzed recently for recursive trees and binary search trees by Fuchs, Hwang, and Neininger (2006).

The rest of the paper is organized as follows. In Section 3 we first give the definition of random recursive trees, and derive a partial differential equation for the number of subtrees of a given size (possibly dependent on n). This involves setting up a basic stochastic recurrence based on a decomposition property of random recursive trees. In Subsection 3.1 we show how to compute the exact moments. The critical and supercritical cases fall out directly from the raw moments, as shown in Subsections 3.2 and 3.3. However, the subcritical case needs centering (and scaling) to get a Gaussian limit. Therefore, to reach the limit distribution via moments, we need to have them centered. This is done in Subsection 3.4, where several technical subsections are dedicated to deriving estimates for the centered moments (Subsections 3.4.1–3.4.4).

Having seen the mechanics of the method applied to recursive trees, the presentation for binary search trees is brief, and sketched in sections modeled after the exposition of recursive trees: starting with the definition of random binary search trees and its partial differential equation in Section 4, then moving on to the exact moments (Section 4.1). Then we show how to compute the limit distributions in the critical and supercritical cases (Sections 4.2 and 4.3). The subcritical case is dealt with in Subsection 4.4, and the string of technical estimates is in Subsections 4.4.1–4.4.5). We conclude in Section 5 with a few remarks on the choice of an analytic methodology.

2.1. Notation. In the discourse we shall use the following notation. The symbol $[\![A]\!]$ denotes Iverson's bracket for the predicate A, i. e. $[\![A]\!] = 1$, if A is true and 0 otherwise. For any number x, and nonnegative integer m, the notation $x^{\underline{m}}$ stands for the falling factorial $x(x-1)\dots(x-m+1)$; we interpret $x^{\underline{0}}$ as 1. The notation $\left\{ {r \atop s} \right\}$ stands for the rth Stirling number (of the second kind) of order s. The operator $[z^n]$ extracts the nth coefficients when applied to a function, that is, $[z^n]f(z)$ is the coefficient of z^n in the power series expansion of $f(z)$.

For probabilistic convergence we use $\xrightarrow{D}$ to denote convergence in distribution. The standard random variables $\mathrm{Poi}(\lambda)$ (the Poisson distributed with parameter λ)

and $\mathcal{N}(\mu, \sigma^2)$ (the normally distributed with mean μ and variance σ^2) appear in the results as limiting random variables. Other standard nomenclature is employed, and we assume it is familiar to the reader, and needing no special mention here.

3. The variety of subtree sizes in a recursive tree. The *random recursive tree* is an outgrowth from a single node labeled 1. Progressively, nodes are added in stages: At the nth stage a node in the existing tree is chosen at random as a parent for the nth entrant (labeled n). In this context *random* means that all nodes in the tree of size $n-1$ are equally likely parents. According to this construction algorithm, the nodes along any root-to-leaf path carry increasing labels, hence these trees are also members of the class of *increasing trees*.

The model of randomness in the growth of random recursive trees induces a uniform distribution on the trees: All $(n-1)!$ recursive trees of size n are generated with equal probability. Many important properties of recursive trees have been analyzed from the convenient viewpoint of recurrence occurring naturally in the stochastic growth. Nonetheless, the uniform distribution of the trees also gave rise to analyses based on generating functions and analytic methods such as integral transforms (see Bergeron, Flajolet and Salvy (1992)).

Let $X_{n,k}$ be the number of nodes in a random recursive tree of size n that are the roots of subtrees of size k. (A subtree rooted at a node is the entire structure of descendant nodes from the given one.) Equivalently, $X_{n,k}$ is the number of subtrees on the fringe of the recursive tree with size k. Our count of the size of a subtree includes its root. Thus, for example, a leaf roots a subtree of size 1. Applications of $X_{n,k}$ include finding out the number of participants in a chain letter scheme, who will attain a certain profit (see Gastwirth and Bhattacharya, 1984), or the number of copies of a particular ancient text in a philological study (see Najock and Heyde, 1982).

We denote by $T_n = (n-1)!$ the number of different recursive trees of size n. Let $M_k(z, v)$ be the moment generating function

$$M_k(z,v) = \sum_{n\ge1}\sum_{m\ge0} P\{X_{n,k}=m\} T_n \frac{z^n}{n!} v^m = \sum_{n\ge1}\sum_{m\ge0} P\{X_{n,k}=m\} \frac{z^n}{n} v^m.$$

According to the recursive nature of these trees, we can read the probabilities that the subtrees have certain sizes right off the definition. For all $k \ge 1$, the probabilities $P\{X_{n,k} = m\}$ satisfy (for $n > k$)

$$\begin{aligned} P\{X_{n,k}=m\} = \sum_{r\ge1} \frac{1}{r!} &\sum_{\substack{n_1+\cdots+n_r=n-1\\ n_i\ge1,\ 1\le i\le r}} \binom{n-1}{n_1,\ldots,n_r} \frac{T_{n_1}\cdots T_{n_r}}{T_n} \\ &\times \sum_{\substack{m_1+\cdots+m_r=m\\ m_i\ge0,\ 1\le i\le r}} P\{X_{n_1,k}=m_1\}\cdots P\{X_{n_r,k}=m_r\}, \end{aligned}$$

with initial values $P\{X_{k,k} = 1\} = 1$, and $P\{X_{n,k} = 0\} = 1$, for $1 \le n < k$.

Multiply this recurrence by $T_n \frac{z^{n-1}}{(n-1)!} v^m$ and sum up over $n > k$ and $m \ge 0$ to obtain the following functional equation, for $k \ge 1$,

$$\frac{\partial}{\partial z} M_k(z,v) = e^{M_k(z,v)} + (v-1) z^{k-1}, \tag{3.1}$$

with initial condition $M_k(0, v) = 0$.

It can be easily checked that the function

$$M_k(z,v) = \frac{(v-1)z^k}{k} + \log\left(\frac{1}{1-\int_0^z e^{\frac{(v-1)t^k}{k}}\,dt}\right) \tag{3.2}$$

is a solution to the partial differential equation (3.1), that also satisfies the initial conditions. The solution (3.2) can be found by using the substitution $Q(z,v) := \exp(M_k(z,v))$ in (3.1), which leads to a solvable Riccati differential equation.

From the solution (3.2) we can obtain the rth factorial moments, which directly furnish a limiting distribution for the critical and supercritical cases.

3.1. Exact moments. To get the rth factorial moments we use the substitution $w := v-1$ in $M_k(z,v)$, and extract coefficients:

$$\mathbf{E}\big(X_{n,k}^{\underline{r}}\big) = \mathbf{E}\big(X_{n,k}(X_{n,k}-1)\cdots(X_{n,k}-r+1)\big) = nr!\,[z^n w^r]M_k(z,1+w).$$

In order to expand $M_k(z,1+w)$ around $w=0$ we consider

$$\begin{aligned}\log\left(\frac{1}{1-\int_0^z e^{\frac{wt^k}{k}}\,dt}\right) &= \log\left(\frac{1}{1-\int_0^z \sum_{j\ge 0}\frac{w^j t^{kj}}{j!\,k^j}dt}\right)\\ &= \log\frac{1}{1-z} + \log\left(\frac{1}{1-\frac{z}{1-z}\sum_{j\ge 1}\frac{w^j z^{kj}}{j!\,k^j(kj+1)}}\right),\end{aligned}$$

and obtain from (3.2):

$$M_k(z,1+w) = \log\frac{1}{1-z} + \frac{wz^k}{k} + \log\left(\frac{1}{1-\frac{z}{1-z}\sum_{j\ge 1}\frac{w^j z^{kj}}{j!\,k^j(kj+1)}}\right). \tag{3.3}$$

Next we extract coefficients and get for $r \ge 1$:

$$\begin{aligned}&[w^r]\log\left(\frac{1}{1-\frac{z}{1-z}\sum_{j\ge 1}\frac{w^j z^{kj}}{j!\,k^j(kj+1)}}\right)\\ &\quad= [w^r]\sum_{\ell\ge 1}\frac{\left(\frac{z}{1-z}\right)^\ell}{\ell}\left(\sum_{j\ge 1}\frac{w^j z^{kj}}{j!\,k^j(kj+1)}\right)^\ell\\ &\quad= \sum_{\ell\ge 1}\frac{\left(\frac{z}{1-z}\right)^\ell}{\ell}\sum_{\substack{j_1+\cdots+j_\ell=r\\ j_q\ge 1,\ 1\le q\le \ell}}\frac{z^{kr}}{k^r\prod_{i=1}^\ell j_i!\prod_{i=1}^\ell(j_ik+1)}\\ &\quad= \sum_{\ell=1}^r\frac{\left(\frac{z}{1-z}\right)^\ell}{\ell}\times\frac{z^{kr}}{k^r}\sum_{\substack{j_1+\cdots+j_\ell=r\\ j_q\ge 1,\ 1\le q\le \ell}}\frac{1}{\prod_{i=1}^\ell j_i!\prod_{i=1}^\ell(j_ik+1)}\\ &\quad= \frac{1}{r!}\sum_{\ell=1}^r\frac{\left(\frac{z}{1-z}\right)^\ell}{\ell}\times\frac{z^{kr}}{k^r}\sum_{\substack{j_1+\cdots+j_\ell=r\\ j_q\ge 1,\ 1\le q\le \ell}}\binom{r}{j_1,\ldots,j_\ell}\frac{1}{\prod_{i=1}^\ell(j_ik+1)}.\end{aligned}$$

This immediately leads from (3.3) to the following formula for the rth coefficients ($r \ge 1$):

$$
\begin{aligned}
[w^r]M_k(z,v) &= \frac{z^k}{k}[\![r=1]\!] \\
&+\frac{1}{r!}\sum_{\ell=1}^{r} \frac{\left(\frac{z}{1-z}\right)^{\ell}}{\ell} \times \frac{z^{kr}}{k^r} \sum_{\substack{j_1+\cdots+j_\ell=r \\ j_q\ge 1,\ 1\le q\le \ell}} \binom{r}{j_1,\ldots,j_\ell} \frac{1}{\prod_{i=1}^{\ell}(j_i k+1)}.
\end{aligned} \tag{3.4}
$$

The formula (3.4) gives

$$
\begin{aligned}
\mathbf{E}(X_{n,k}) &= n[z^n w]M_k(z,1+w) \\
&= n[z^n]\left(\frac{z^k}{k} + \frac{z^{k+1}}{k(k+1)(1-z)}\right) \\
&= n\Big(\frac{1}{k}[\![n=k]\!] + \frac{1}{k(k+1)}[\![n \ge k+1]\!]\Big),
\end{aligned}
$$

and thus the following result ensues for the expectation of $X_{n,k}$, as given in Feng, Mahmoud and Su (2007):

$$
\mathbf{E}(X_{n,k}) = \begin{cases} \frac{n}{k(k+1)}, & \text{for } n \ge k+1; \\ 1, & \text{for } n = k; \\ 0, & \text{for } 1 \le n < k. \end{cases} \tag{3.5}
$$

Owing to the explicit form of (3.4) we also get the following closed form solution for the rth factorial moments with $r \ge 2$:

$$
\begin{aligned}
\mathbf{E}\big(X_{n,k}^{\underline{r}}\big) &= nr![z^n w^r]M_k(z,v) \\
&= \frac{[\![n \ge kr+1]\!]n}{k^r} \sum_{\ell=1}^{r} \frac{\binom{n-kr-1}{\ell-1}}{\ell} \\
&\quad \times \sum_{\substack{j_1+\cdots+j_\ell=r \\ j_q\ge 1,\ 1\le q\le \ell}} \binom{r}{j_1,\ldots,j_\ell} \frac{1}{\prod_{i=1}^{\ell}(j_i k+1)}.
\end{aligned} \tag{3.6}
$$

As an example, we compute the second factorial moment:

$$
\mathbf{E}\big(X_{n,k}(X_{n,k}-1)\big) = \frac{n(n-2k-1)}{k^2(k+1)^2} - \frac{n}{k^2(2k+1)}, \qquad \text{for } n \ge 2k+1,
$$

and is 0 otherwise.

In view of the classical relation

$$
\mathbf{E}\big(X_{n,k}^{r}\big) = \sum_{\ell=1}^{r} \left\{ {r \atop \ell} \right\} \mathbf{E}\big(X_{n,k}^{\underline{\ell}}\big),
$$

one could also get (at least in principle) closed formulas for every ordinary rth moment.

3.2. The critical case. We are in the critical case when $k = \Theta(\sqrt{n}\,)$. There are two flavors within this case: one flavor is when $k/\sqrt{n}$ converges to a limit, such as for example the case $k = 3\lfloor\sqrt{n} + \log n\rfloor$, the other flavor is when k is asymptotic to $g(n)\sqrt{n}$ with $g(n)$ being a function of bounded variation, but oscillating and not converging to any limit, such as for example the case $k = \lfloor(2 + \sin n)\sqrt{n} + 6\rfloor$.

Firstly, we consider the critical case having the flavor $\frac{n}{k^2} \to \lambda$, for some $\lambda > 0$. It is evident from (3.5) that we have $\mathbf{E}(X_{n,k}) \to \lambda$. Next we consider a fixed $r \geq 2$. We split the sum appearing in (3.6) as follows:

$$\mathbf{E}\big(X_{n,k}^{\underline{r}}\big) = \underbrace{\frac{[\![n \geq kr+1]\!]n}{k^r} \times \frac{\binom{n-kr-1}{r-1}}{r} \sum_{\substack{j_1+\cdots+j_r=r\\ j_q\geq 1,\ 1\leq q\leq r}} \frac{\binom{r}{j_1,\dots,j_r}}{\prod_{i=1}^{r}(j_i k+1)}}_{=:A}$$
$$+ \underbrace{\frac{[\![n \geq kr+1]\!]n}{k^r} \sum_{\ell=1}^{r-1} \frac{\binom{n-kr-1}{\ell-1}}{\ell} \sum_{\substack{j_1+\cdots+j_\ell=r\\ j_q\geq 1,\ 1\leq q\leq \ell}} \frac{\binom{r}{j_1,\dots,j_\ell}}{\prod_{i=1}^{\ell}(j_i k+1)}}_{=:B}.$$

For large n we can drop Iverson's brackets in both A and B. Using

$$\sum_{r\geq 0} \sum_{\substack{j_1+\cdots+j_\ell=r\\ j_q\geq 1,\ 1\leq q\leq \ell}} \binom{r}{j_1,\dots,j_\ell} \frac{z^r}{r!} = (e^z-1)^\ell,$$

we obtain the inequality

$$\sum_{\substack{j_1+\cdots+j_\ell=r\\ j_q\geq 1,\ 1\leq q\leq \ell}} \binom{r}{j_1,\dots,j_\ell} = r![z^r](e^z-1)^\ell \leq r![z^r]e^{\ell z} = \ell^r,$$

and further by trivial estimates:

$$\begin{aligned}
B &\leq \sum_{\ell=1}^{r-1} \frac{n^\ell}{\ell!\,k^r k^\ell}\ell^r \\
&\leq \frac{r^{r-1}}{k} \sum_{\ell=1}^{r-1} \frac{1}{k^{r-1-\ell}} \Big(\frac{n}{k^2}\Big)^\ell \frac{1}{(\ell-1)!} \\
&\leq \frac{r^{r-1}}{k} \times \frac{n}{k^2} \sum_{\ell\geq 0} \frac{\big(\frac{n}{k^2}\big)^\ell}{\ell!} \\
&= \frac{r^{r-1}n}{k^3} e^{\frac{n}{k^2}},
\end{aligned}$$

Since $\frac{n}{k^2} \to \lambda$, we further have

$$B = O\Big(\frac{1}{k}\Big) = O\Big(\frac{1}{\sqrt{n}}\Big).$$

The expression A simplifies considerably, since the only possible composition of r in r integers $j_q \geq 1$ is obtained for $j_1 = \cdots = j_r = 1$; thus for large n we have

$$\begin{aligned} A &= \frac{n}{k^r} \times \frac{\binom{n-kr-1}{r-1}}{r} \times \frac{r!}{(k+1)^r} \\ &= \frac{n^r}{k^{2r}}\Big(1 + O\Big(\frac{k}{n}\Big) + O\Big(\frac{1}{k}\Big)\Big) \\ &= \Big(\frac{n}{k^2}\Big)^r \big(1 + O\Big(\frac{1}{\sqrt{n}}\Big)\big). \end{aligned}$$

Therefore, if $\frac{n}{k^2} \to \lambda$, we also get for every $r \geq 1$:

$$\mathbf{E}\big(X_{n,k}^{\underline{r}}\big) = \Big(\frac{n}{k^2}\Big)^r + O\Big(\frac{1}{\sqrt{n}}\Big) \to \lambda^r.$$

Since λ^r are the moments of $\mathrm{Poi}(\lambda)$, and convergence of all moments to a random variable uniquely characterized by its moments implies weak convergence, we get

$$X_{n,k} \xrightarrow{D} \mathrm{Poi}(\lambda).$$

This proves the convergence part of Theorem 2.1 (b), where we used the substitution $c := \frac{1}{\sqrt{\lambda}}$, and thus $\frac{k}{\sqrt{n}} \to c$. When k is of the order $\sqrt{n}$, but fluctuations persist, as in the case involving sinusoids mentioned at the beginning of this subsection, all the moments oscillate, and no limit distribution can exist.

3.3. The supercritical case. We consider the supercritical case in passing, as it does not require additional effort. Assume that $k := k_n$ grows with n such that $\frac{n}{k^2} = o(1)$.

Crude estimates for (3.6) are sufficient for our purpose. We get for $r \geq 2$:

$$\begin{aligned} \mathbf{E}\big(X_{n,k}^{\underline{r}}\big) &\leq \frac{n}{k^r} \sum_{\ell=1}^{r} \frac{n^{\ell-1}}{\ell!} \sum_{\substack{j_1+\cdots+j_\ell=r \\ j_q \geq 1,\ 1 \leq q \leq \ell}} \binom{r}{j_1, \ldots, j_\ell} \frac{1}{k^\ell} = \frac{1}{k^r} \sum_{\ell=1}^{r} \frac{n^\ell}{\ell!\, k^\ell} \ell^r \\ &\leq r^{r-1} \sum_{\ell=1}^{r} \frac{1}{k^{r-\ell}} \Big(\frac{n}{k^2}\Big)^\ell \frac{1}{(\ell-1)!} \leq r^{r-1}\, \frac{n}{k^2} \sum_{\ell \geq 0} \frac{\big(\frac{n}{k^2}\big)^\ell}{\ell!} \\ &= r^{r-1}\, \frac{n}{k^2} e^{\frac{n}{k^2}}. \end{aligned}$$

Since $\frac{n}{k^2} \to 0$, we have that for all $r \geq 1$: $\mathbf{E}\big(X_{n,k}^{\underline{r}}\big) \to 0$. Again we have convergence of all moments, and subsequently convergence of $X_{n,k}$ to a degenerate distribution with mass at 0. Part (c) of Theorem 2.1 is proved. Because the limit is constant, this convergence is in probability, too.

3.4. The subcritical case. For the critical and supercritical cases, it was sufficient to work with raw moments and establish their convergence to the moments of a known simple distribution. For the subcritical case, we need centering and scaling, too. We appeal to the method of recursive moments (see Chern, Hwang and Tsai, 2002), although here we are not forced to "pump out" the behaviour of the moments in an inductive way, but we are able to obtain exact (but rather involved) expression for the moments directly by extracting coefficients of the moment generating function of the centered random variable. Consider the centered random variable $\tilde{X}_{n,k} := X_{n,k} - \mathbf{E}(X_{n,k})$ and introduce the generating function

$$\tilde{M}_k(z,s) := \sum_{n\geq 1} \mathbf{E}\big(e^{\tilde{X}_{n,k}s}\big)\frac{z^n}{n} = \sum_{n\geq 1} e^{-\mathbf{E}(X_{n,k})s}\mathbf{E}\big(e^{X_{n,k}s}\big)\frac{z^n}{n}.$$

Using the explicit formula for $\mathbf{E}(X_{n,k})$ as given by (3.5) we obtain by routine manipulations

$$\tilde{M}_k(z,s) = M_k\big(e^{-\frac{s}{k(k+1)}}z, e^s\big) + \big(1 - e^{\frac{ks}{k+1}}\big)\frac{z^k}{k} + \sum_{1\leq n<k}\frac{z^n}{n} - \sum_{1\leq n<k}\frac{\big(e^{-\frac{s}{k(k+1)}}z\big)^n}{n}.$$

Plugging in the formula for $M_k(z,v)$ given by (3.2) we further obtain after simplifications:

$$\begin{aligned}\tilde{M}_k(z,s) &= \log\left(\frac{1}{1-\int_0^{e^{-\frac{s}{k(k+1)}}z} e^{\frac{(e^s-1)t^k}{k}}\,dt}\right)\\ &\quad+\sum_{1\leq n\leq k}\frac{z^n}{n} - \sum_{1\leq n\leq k}\frac{\big(e^{-\frac{s}{k(k+1)}}z\big)^n}{n}.\end{aligned}\tag{3.7}$$

In the following chain of technical subsections, we analyze the centered moments, identify the main contribution and put upper bounds on the negligible terms.

3.4.1. Expanding around $s = 0$**.** We first expand $\int_0^{e^{-\frac{s}{k(k+1)}}z} e^{\frac{(e^s-1)t^k}{k}}\,dt$ around $s = 0$ and $z = 1$:

$$\begin{aligned}&\int_0^{e^{-\frac{s}{k(k+1)}}z} e^{\frac{(e^s-1)t^k}{k}}\,dt\\
&= \int_0^{e^{-\frac{s}{k(k+1)}}z}\sum_{j\geq 0}\frac{(e^s-1)^j t^{kj}}{j!\,k^j}\,dt = \sum_{j\geq 0}\frac{(e^s-1)^j e^{-\frac{s(kj+1)}{k(k+1)}} z^{kj+1}}{j!\,(kj+1)k^j}\\
&= e^{-\frac{s}{k(k+1)}}z + \sum_{j\geq 1}\frac{(e^s-1)^j e^{-\frac{s(kj+1)}{k(k+1)}} z^{kj+1}}{j!\,(kj+1)k^j}\\
&= z + \sum_{j\geq 1}\left[\frac{(e^s-1)^j e^{-\frac{s(kj+1)}{k(k+1)}} z^{kj+1}}{j!\,(kj+1)k^j} + \frac{(-1)^j s^j z}{j!\,k^j(k+1)^j}\right]\\
&= z + \sum_{j\geq 1}\left[\frac{\sum_{m\geq j}\left\{ {m \atop j} \right\}\frac{j!\,s^m}{m!}}{j!\,(kj+1)k^j}\sum_{m\geq 0}\frac{(-1)^m(kj+1)^m s^m}{k^m(k+1)^m m!}z^{kj+1} + \frac{(-1)^j s^j z}{j!\,k^j(k+1)^j}\right]\\
&= z + \sum_{\ell\geq 1}s^\ell\Big(\sum_{j=1}^{\ell}\sum_{m=0}^{\ell-j}\frac{\left\{ {\ell-m \atop j} \right\}(-1)^m(kj+1)^m z^{kj+1}}{m!\,(\ell-m)!\,(kj+1)k^{j+m}(k+1)^m} + \frac{(-1)^\ell z}{\ell!\,k^\ell(k+1)^\ell}\Big)\\
&= z + \sum_{\ell\geq 1}s^\ell\Big(\sum_{j=1}^{\ell}\sum_{m=0}^{\ell-j}\frac{\left\{ {\ell-m \atop j} \right\}(-1)^m(kj+1)^m}{m!\,(\ell-m)!\,(kj+1)k^{j+m}(k+1)^m}\end{aligned}$$

$$\times \sum_{i=0}^{kj+1} (-1)^i \binom{kj+1}{i} (1-z)^i + \frac{(-1)^\ell}{\ell!\, k^\ell (k+1)^\ell} + \frac{(-1)^{\ell-1}}{\ell!\, k^\ell (k+1)^\ell}(1-z)\Big)$$

$$= z + \sum_{\ell \ge 1} s^\ell \Big(\sum_{i=0}^{\ell k+1} (-1)^i (1-z)^i \sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{\left\{ {\ell-m \atop j} \right\} (-1)^m (kj+1)^m \binom{kj+1}{i}}{m!\,(\ell-m)!\,(kj+1)k^{j+m}(k+1)^m}$$

$$+ \frac{(-1)^\ell}{\ell!\, k^\ell (k+1)^\ell} + \frac{(-1)^{\ell-1}}{\ell!\, k^\ell (k+1)^\ell}(1-z)\Big).$$

So, we obtain the following simple structure for the integral considered:

$$\int_0^{e^{-\frac{s}{k(k+1)}} z} e^{\frac{(e^s-1)t^k}{k}}\, dt = z + \sum_{\ell \ge 1} s^\ell \sum_{i=0}^{\ell k+1} c_{\ell,i}(k)\,(1-z)^i,$$

where the functions $c_{\ell,i}(k)$ are given by

$$c_{\ell,i}(k) = (-1)^i \sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{\left\{ {\ell-m \atop j} \right\} (-1)^m (kj+1)^m \binom{kj+1}{i}}{m!\,(\ell-m)!\,(kj+1)k^{j+m}(k+1)^m}$$

$$+ \frac{(-1)^\ell [\![i=0]\!]}{\ell!\, k^\ell (k+1)^\ell} + \frac{(-1)^{\ell-1} [\![i=1]\!]}{\ell!\, k^\ell (k+1)^\ell}. \tag{3.8}$$

Next we consider

$$\log \left(\frac{1}{1 - \int_0^{e^{-\frac{s}{k(k+1)}} z} e^{\frac{(e^s-1)t^k}{k}}\, dt} \right)$$

$$= \log \frac{1}{1-z} + \log \left(\frac{1}{1 - \frac{1}{1-z} \sum_{\ell \ge 1} s^\ell \sum_{i=0}^{\ell k+1} c_{\ell,i}(k)(1-z)^i} \right),$$

and extract coefficients. This gives for $r \ge 1$:

$$[s^r] \log \left(\frac{1}{1 - \int_0^{e^{-\frac{s}{k(k+1)}} z} e^{\frac{(e^s-1)t^k}{k}}\, dt} \right)$$

$$= [s^r] \log \left(\frac{1}{1 - \frac{1}{1-z} \sum_{\ell \ge 1} s^\ell \sum_{i=0}^{\ell k+1} c_{\ell,i}(k)(1-z)^i} \right)$$

$$= \sum_{m=1}^{r} \frac{1}{m} [s^r] \left(\frac{1}{1-z} \sum_{\ell \ge 1} s^\ell \sum_{i=0}^{\ell k+1} (1-z)^i c_{\ell,i}(k) \right)^m$$

$$= \sum_{m=1}^{r} \frac{1}{m(1-z)^m} \sum_{\substack{r_1+\cdots+r_m=r \\ r_q \ge 1,\ 1 \le q \le m}} \prod_{j=1}^{m} \sum_{i=0}^{r_j k+1} (1-z)^i c_{r_j,i}(k)$$

$$= \sum_{m=1}^{r} \frac{1}{m(1-z)^m} \sum_{t=0}^{rk+m} (1-z)^t \times \sum_{\substack{r_1+\cdots+r_m=r \\ r_q \ge 1,\ 1 \le q \le m}} \sum_{\substack{t_1+\cdots+t_m=t \\ 0 \le t_q \le r_q k+1 \\ 1 \le q \le m}} \prod_{j=1}^{m} c_{r_j,t_j}(k),$$

and thus

$$[s^r]\log\left(\frac{1}{1-\int_0^{e^{-\frac{s}{k(k+1)}}z} e^{\frac{(e^s-1)t^k}{k}}dt}\right) = \sum_{p=-r}^{rk} f_{r,p}(k)(1-z)^p,$$

with

$$f_{r,p}(k) := \sum_{m=\max\{1,-p\}}^{r} \frac{1}{m} \sum_{\substack{r_1+\cdots+r_m=r\\ r_q\ge 1,\ 1\le q\le m}} \sum_{\substack{t_1+\cdots+t_m=p+m\\ 0\le t_q\le r_q k+1,\ 1\le q\le m}} \prod_{j=1}^{m} c_{r_j,t_j}(k),$$

where the functions $c_{\ell,i}(k)$ are defined by (3.8).

Since

$$[s^r]\sum_{1\le n\le k}\frac{\big(e^{-\frac{s}{k(k+1)}}z\big)^n}{n} = \sum_{1\le n\le k}\frac{z^n}{n}[s^r]e^{-\frac{sn}{k(k+1)}} = \sum_{1\le n\le k} z^n\frac{(-1)^r n^{r-1}}{r!\,k^r(k+1)^r},$$

we obtain for the coefficients at s^r of $\tilde{M}_k(z,s)$ as given by (3.7), for $r\ge 1$:

$$[s^r]\tilde{M}_k(z,s) = \sum_{p=-r}^{rk} f_{r,p}(k)(1-z)^p + \sum_{1\le n\le k} z^n\frac{(-1)^{r-1}n^{r-1}}{r!\,k^r(k+1)^r},$$

which can be written as

$$[s^r]\tilde{M}_k(z,s) = \sum_{p=1}^{r}\frac{1}{(1-z)^p}\tilde{f}_{r,p}(k) + \sum_{n=0}^{rk} g_{r,n}(k)z^n,$$

with

$$\tilde{f}_{r,p}(k) := f_{r,-p}(k) = \sum_{m=p}^{r}\frac{1}{m} \sum_{\substack{r_1+\cdots+r_m=r\\ r_q\ge 1,\ 1\le q\le m}} \sum_{\substack{t_1+\cdots+t_m=m-p\\ 0\le t_q\le r_q k+1,\ 1\le q\le m}} \prod_{j=1}^{m} c_{r_j,t_j}(k), \tag{3.9}$$

for $1\le p\le r$, and

$$g_{r,n}(k) := (-1)^n\sum_{p=n}^{rk}\binom{p}{n}f_{r,p}(k) + \frac{(-1)^{r-1}n^{r-1}}{r!k^r(k+1)^r}[\![1\le n\le k]\!].$$

This representation of $[s^r]\tilde{M}_k(z,s)$ immediately leads to the following explicit formula for the rth moments of $\tilde{X}_{n,k}$, which is valid for $n\ge rk+1$:

$$\mathbf{E}\big(\tilde{X}_{n,k}^r\big) = r!\,n[z^n s^r]\tilde{M}_k(z,s) = r!\,n\sum_{p=1}^{r}\tilde{f}_{r,p}(k)\binom{n+p-1}{p-1}, \tag{3.10}$$

where the functions $\tilde{f}_{r,p}(k)$ are given by (3.9).

We now examine the functions $c_{\ell,i}(k)$ as given by (3.8). Using the identity

$$\sum_{m=j}^{\ell}\binom{\ell}{m}\left\{ {m \atop j} \right\} = \left\{ {\ell+1 \atop j+1} \right\},$$

we get by very crude estimates:

$$
\begin{aligned}
&\left|(-1)^i \sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{\left\{ {\ell-m \atop j} \right\} (-1)^m (kj+1)^m \binom{kj+1}{i}}{m!\,(\ell-m)!\,(kj+1)k^{j+m}(k+1)^m}\right| \\
&\quad \le \sum_{j=1}^{\ell} \frac{\binom{kj+1}{i}}{(kj+1)k^j} \sum_{m=0}^{\ell-j} \frac{\left\{ {\ell-m \atop j} \right\}(kj+1)^m}{m!\,(\ell-m)!\,k^m(k+1)^m} \\
&\quad \le \sum_{j=1}^{\ell} \frac{k^i (j+1)^i}{\ell!\,i!\,k^{j+1}} \sum_{m=0}^{\ell-j} \frac{\left\{ {\ell-m \atop j} \right\}(j+1)^m \binom{\ell}{m}}{(k+1)^m} \\
&\quad \le \sum_{j=1}^{\ell} \frac{k^i (j+1)^i (j+1)^{\ell-j}}{\ell!\,i!\,k^{j+1}} \sum_{m=0}^{\ell-j} \binom{\ell}{m} \left\{ {\ell-m \atop j} \right\} \\
&\quad = \sum_{j=1}^{\ell} \frac{k^i (j+1)^i (j+1)^{\ell-j}}{\ell!\,i!\,k^{j+1}} \left\{ {\ell+1 \atop j+1} \right\} \\
&\quad \le \frac{k^i (\ell+1)^i (\ell+1)^{\ell-1}}{\ell!\,i!\,k^2} \sum_{j=2}^{\ell+1} \left\{ {\ell+1 \atop j} \right\} \qquad (3.11) \\
&\quad \le \frac{(\ell+1)^\ell B_{\ell+1}}{(\ell+1)!} \times \frac{(\ell+1)^i}{i!} k^{i-2},
\end{aligned}
$$

where B_ℓ denote the Bell numbers (number of partitions of a set with ℓ elements). Since

$$\left|\frac{(-1)^\ell}{k^\ell (k+1)^\ell \ell!}\right| \le \frac{1}{k^2} = \frac{1}{k^2}\left\{ {\ell+1 \atop 1} \right\},$$

we can add this summand in the preceding computations at step (3.11) and thus get the same bound. Therefore, we arrive at the following bound, which holds uniformly for all ℓ, i and k:

$$|c_{\ell,i}(k)| \le q_{\ell,i} k^{i-2}, \qquad \text{with} \qquad q_{\ell,i} := \frac{(\ell+1)^\ell B_{\ell+1}}{(\ell+1)!} \times \frac{(\ell+1)^i}{i!}.$$

The two functions that exert the most influence on the asymptotic distribution $c_{\ell,i}(k)$, and which can be computed easily from (3.8), are

$$
\begin{aligned}
c_{1,0}(k) &= \frac{1}{k(k+1)} - \frac{1}{k(k+1)} = 0, \\
c_{2,0}(k) &= \frac{\nu(k)}{2}, \qquad \text{with} \qquad \nu(k) := \frac{2k^2-1}{k(k+1)^2(2k+1)}.
\end{aligned}
$$

3.4.2. Estimates for $\tilde{f}_{r,p}(k)$. We next treat the functions $\tilde{f}_{r,p}(k)$ as given by (3.9). Since $1 \le p \le r$ holds in (3.10), this implies that in the compositions appearing in the definition of $\tilde{f}_{r,p}(k)$ we always have $r_q, t_q \le r$ for all $1 \le q \le m$. Thus, we obtain the following estimate, which holds for all $1 \le p \le r$ and all r_q, t_q:

$$|c_{r_j,t_j}(k)| \le q_{r_j,t_j} k^{t_j-2} \le c_r k^{t_j-2},$$

with the universal constant

$$c_r := \frac{(r+1)^{2r} B_{r+1}}{(r+1)!}. \tag{3.12}$$

This gives for $1 \le p \le r$ the estimate

$$\begin{aligned}
&\left| \sum_{\substack{r_1+\cdots+r_m=r \\ r_q\ge 1,\ 1\le q\le m}} \sum_{\substack{t_1+\cdots+t_m=m-p \\ 0\le t_q\le r_q k+1,\ 1\le q\le m}} \prod_{j=1}^{m} c_{r_j,t_j}(k) \right| \\
&\le \sum_{\substack{r_1+\cdots+r_m=r \\ r_q\ge 1,\ 1\le q\le m}} \sum_{\substack{t_1+\cdots+t_m=m-p \\ 0\le t_q\le r_q k+1,\ 1\le q\le m}} \prod_{j=1}^{m} c_r k^{t_j-2} \\
&= c_r^m k^{-p-m} \sum_{\substack{r_1+\cdots+r_m=r \\ r_q\ge 1,\ 1\le q\le m}} \sum_{\substack{t_1+\cdots+t_m=m-p \\ 0\le t_q\le r_q k+1,\ 1\le q\le m}} 1 \\
&\le c_r^m k^{-p-m} \sum_{\substack{r_1+\cdots+r_m=r \\ r_q\ge 1,\ 1\le q\le m}} \sum_{\substack{t_1+\cdots+t_m=m-p \\ t_q\ge 0,\ 1\le q\le m}} 1 \\
&= c_r^m k^{-p-m} [z^r] \left(\frac{z}{1-z}\right)^m [z^{m-p}] \left(\frac{1}{1-z}\right)^m \\
&= c_r^m k^{-p-m} \binom{r-1}{p-1} \binom{2m-p-1}{m-1} \\
&\le c_r^m \binom{2m-2}{m-1} \binom{r-1}{m-1} k^{-p-m},
\end{aligned}$$

and further

$$\begin{aligned}
\left|\tilde{f}_{r,p}(k)\right| &= \left| \sum_{m=p}^{r} \frac{1}{m} \sum_{\substack{r_1+\cdots+r_m=r \\ r_q\ge 1,\ 1\le q\le m}} \sum_{\substack{t_1+\cdots+t_m=m-p \\ 0\le t_q\le r_q k+1,\ 1\le q\le m}} \prod_{j=1}^{m} c_{r_j,t_j}(k) \right| \\
&\le \sum_{m=p}^{r} \frac{1}{m} c_r^m \binom{2m-2}{m-1} \binom{r-1}{m-1} k^{-p-m} \\
&\le \binom{2r-2}{r-1} (r-1)!\, c_r^r \sum_{m=p}^{r} \frac{1}{k^{p+m}} = \binom{2r-2}{r-1} (r-1)!\, c_r^r \frac{1}{k^{2p}} \sum_{q=0}^{r-p} \frac{1}{k^q} \\
&\le \binom{2r-2}{r-1} (r-1)!\, c_r^r \frac{1}{k^{2p}} r.
\end{aligned}$$

Thus, we obtain the following estimate, which holds for all $1 \le p \le r$ and $r, k \ge 1$:

$$\left|\tilde{f}_{r,p}(k)\right| \le \kappa_r \frac{1}{k^{2p}},$$

with

$$\kappa_r := \binom{2r-2}{r-1} r!\, c_r^r.$$

The constants c_r appearing here are given by (3.12).

3.4.3. Cancelation of $\tilde{f}_{r,p}(k)$. We shall demonstrate that the functions $\tilde{f}_{r,p}(k)$ as defined by (3.9) satisfy $\tilde{f}_{r,p}(k) = 0$ for all $p \ge \lfloor \frac{r}{2} \rfloor + 1$. We will do this by simply showing that for every composition $m, r_1, \dots, r_m, t_1, \dots, t_m$ of r and $m - p$ respectively there exists a factor $c_{1,0}(k)$, which is zero, and thus the product $\prod_{j=1}^{m} c_{r_j,t_j}(k)$ vanishes.

Let us take $p \ge \lfloor \frac{r}{2} \rfloor + 1$. This implies then by (3.9) that also $m \ge \lfloor \frac{r}{2} \rfloor + 1$. We consider now an arbitrary but fixed composition $m, r_1, \dots, r_m, t_1, \dots, t_m$ with $r_q \ge 1$ and $t_q \ge 0$, such that $r_1 + \dots + r_m = r$ and $t_1 + \dots + t_m = m - p$. We define the sets $T = \{q : t_q = 0\}$ and $R = \{q : r_q = 1\}$. Then the relations

$$|T| \ge m - (m - p) = p, \qquad \text{and} \qquad |R| \ge 2m - r$$

hold. Otherwise, if $|T| \le p - 1$ this would imply $\sum_{q=1}^{m} t_q \ge m - p + 1 > m - p$, and if $|R| \le 2m - r - 1$ this would imply $\sum_{q=1}^{r} r_q \ge (2m - r - 1) + 2(m - (2m - r - 1)) = r + 1 > r$, which is a contradiction to the choice of r_q and t_q.

Next we consider the set $R \cap T = \{q : t_q = 0 \wedge r_q = 1\}$. Since $|R \cup T| \le m$, we get by a simple application of the inclusion-exclusion-principle for $p \ge \lfloor \frac{r}{2} \rfloor + 1$ the estimate

$$|R \cap T| = |R| + |T| - |R \cup T| \ge 2m - r + p - m = m + p - r \ge 2\Big\lfloor \frac{r}{2} \Big\rfloor - r + 2 \ge 1.$$

The latter equation guarantees that in every composition there exists a number q, such that $r_q = 1$ and $t_q = 0$, and consequently there is always a factor $c_{1,0}(k) = 0$ in every product $\prod_{j=1}^{m} c_{r_j,t_j}(k)$. Hence, we have

$$\tilde{f}_{r,p}(k) = 0, \qquad \text{for } p \ge \Big\lfloor \frac{r}{2} \Big\rfloor + 1.$$

We also consider the special case $r = 2d$ and $p = d$, i. e. we examine

$$\tilde{f}_{2d,d}(k) = \sum_{m=d}^{2d} \frac{1}{m} \sum_{\substack{r_1 + \dots + r_m = 2d \\ r_q \ge 1,\ 1 \le q \le m}} \sum_{\substack{t_1 + \dots + t_m = m - d \\ 0 \le t_q \le r_q k + 1,\ 1 \le q \le m}} \prod_{j=1}^{m} c_{r_j,t_j}(k).$$

If $m \ge d + 1$, we can show by arguments as above that there always exists a factor $c_{1,0}(k) = 0$ in every product $\prod_{j=1}^{m} c_{r_j,t_j}(k)$. Thus all summands with $m \ge d + 1$ are zero. This gives

$$\tilde{f}_{2d,d}(k) = \frac{1}{d} \sum_{\substack{r_1 + \dots + r_d = 2d \\ r_q \ge 1,\ 1 \le q \le d}} \sum_{\substack{t_1 + \dots + t_d = 0 \\ 0 \le t_q \le r_q k + 1,\ 1 \le q \le d}} \prod_{j=1}^{d} c_{r_j,t_j}(k) = \frac{1}{d} \sum_{\substack{r_1 + \dots + r_d = 2d \\ r_q \ge 1,\ 1 \le q \le d}} \prod_{j=1}^{d} c_{r_j,0}(k).$$

Now, since $c_{1,0}(k) = 0$, we see that the only non-zero contribution is obtained if $r_1 = \dots = r_d = 2$, which implies

$$\tilde{f}_{2d,d}(k) = \frac{1}{d} c_{2,0}^{d}(k) = \frac{\nu^d(k)}{d 2^d},$$

with $\nu(k) = \frac{2k^2 - 1}{k(k+1)^2(2k+1)}$.

3.4.4. Asymptotics of the centered moments. Summarizing the results of the previous subsections we have for $r \geq 1$ that for all $n \geq rk+1$ the rth centered moments are given by

$$\mathbf{E}\big(\tilde{X}_{n,k}^r\big) = r!\, n \sum_{p=1}^{\lfloor \frac{r}{2} \rfloor} \tilde{f}_{r,p}(k) \binom{n+p-1}{p-1},$$

with

$$|\tilde{f}_{r,p}(k)| \leq \frac{\kappa_r}{k^{2p}}, \qquad \text{and} \qquad \tilde{f}_{2d,d}(k) = \frac{\nu^d(k)}{d 2^d}.$$

We are now going to evaluate the centered moments asymptotically for the subcritical case, i. e. $\frac{k^2}{n} = o(1)$; thus we may always assume that we choose n large enough, such that $\frac{(d-1)^2}{n} \leq 1$ and $\frac{k^2}{n} \leq \frac{1}{2}$ are satisfied.

We consider the case $r = 2d$ with $d \geq 1$, which gives

$$\begin{aligned}
\mathbf{E}\big(\tilde{X}_{n,k}^{2d}\big) &= (2d)!\, \frac{\nu^d(k)}{d 2^d} \times \frac{(n+d-1)^{\underline{d-1}} n}{(d-1)!} + \sum_{p=1}^{d-1} (2d)!\, \tilde{f}_{2d,p}(k) \binom{n+p-1}{p-1} n \\
&= \frac{(2d)!}{d!\, 2^d} \nu^d(k) n^d \big(1 + R(n,k)\big),
\end{aligned}$$

with

$$R(n,k) := \prod_{j=1}^{d-1} \big(1 + \frac{j}{n}\big) - 1 + \sum_{p=1}^{d-1} \frac{2^d d!\, \tilde{f}_{2d,p}(k)}{\nu^d(k) n^{d-1}} \binom{n+p-1}{p-1}.$$

We obtain the simple estimate

$$\begin{aligned}
\left| \prod_{j=1}^{d-1} \Big(1 + \frac{j}{n}\Big) - 1 \right| &= \prod_{j=1}^{d-1} \Big(1 + \frac{j}{n}\Big) - 1 \leq \big(1 + \frac{d-1}{n}\big)^{d-1} - 1 \\
&\leq e^{\frac{(d-1)^2}{n}} - 1 \leq \frac{(d-1)^2}{n} e^{\frac{(d-1)^2}{n}} \\
&\leq \frac{(d-1)^2 e}{n}.
\end{aligned}$$

Using the trivial bounds

$$\frac{1}{12k^2} \leq \nu(k) \leq \frac{1}{k^2},$$

which hold for all $k \geq 1$, we obtain the following estimates:

$$\begin{aligned}
\left| \sum_{p=1}^{d-1} \frac{2^d d!\, \tilde{f}_{2d,p}(k) \binom{n+p-1}{p-1}}{\nu^d(k) n^{d-1}} \right| &\leq 2^d d!\, \kappa_{2d} \sum_{p=1}^{d-1} \frac{n^{p-1} \prod_{j=1}^{p-1} \big(1 + \frac{j}{n}\big)}{\nu^d(k) n^{d-1} k^{2p}} \\
&\leq 2^d d!\, \kappa_{2d} e^{\frac{(d-1)^2}{n}} \sum_{p=1}^{d-1} \frac{1}{n^{d-p} \nu^d(k) k^{2p}}
\end{aligned}$$

$$\begin{aligned}
&\le 2^d d!\,\kappa_{2d} e \sum_{p=1}^{d-1} \frac{12^d k^{2d}}{n^{d-p} k^{2p}} \\
&= (24)^d d!\,\kappa_{2d} e \frac{k^2}{n} \sum_{q=0}^{d-2} \Big(\frac{k^2}{n}\Big)^q \\
&\le (24)^d d!\,\kappa_{2d} e \frac{k^2}{n} \times \frac{1}{1-\frac{k^2}{n}} \\
&\le 2e(24)^d d!\,\kappa_{2d} \frac{k^2}{n}.
\end{aligned}$$

Combining these estimates yields in the subcritical case:

$$|R(n,k)| \le \frac{(d-1)^2 e}{n} + 2e(24)^d d!\,\kappa_{2d} \frac{k^2}{n} = o(1),$$

and further

$$\mathbf{E}\big(\tilde{X}_{n,k}^{2d}\big) \sim \frac{(2d)!}{2^d d!}\big(\nu(k)n\big)^d, \qquad \text{for } d \ge 1.$$

We have to consider also the case $r = 2d+1$ with $d \ge 0$, which gives

$$\begin{aligned}
\Big|\mathbf{E}\big(\tilde{X}_{n,k}^{2d+1}\big)\Big| &= \left|(2d+1)!\, n \sum_{p=1}^{d} \tilde{f}_{2d+1,p}(k) \binom{n+p-1}{p-1}\right| \\
&\le (2d+1)! \sum_{p=1}^{d} \frac{\kappa_{2d+1}}{k^{2p}} \frac{n^p}{(p-1)!} \prod_{j=1}^{p-1}\Big(1+\frac{j}{n}\Big) \\
&\le (2d+1)!\,\kappa_{2d+1} e^{\frac{(d-1)^2}{n}} \sum_{p=1}^{d} \Big(\frac{n}{k^2}\Big)^p \\
&\le (2d+1)!\,\kappa_{2d+1} e \Big(\frac{n}{k^2}\Big)^d \sum_{q=0}^{d-1} \Big(\frac{k^2}{n}\Big)^q \\
&\le (2d+1)!\,\kappa_{2d+1} e \Big(\frac{n}{k^2}\Big)^d \frac{1}{1-\frac{k^2}{n}} \\
&\le 2e\big[(2d+1)!\big]\kappa_{2d+1} \Big(\frac{n}{k^2}\Big)^d \\
&\le 2e(12)^d \big[(2d+1)!\big]\, \kappa_{2d+1} \big(\nu(k)n\big)^d,
\end{aligned}$$

and thus

$$\mathbf{E}\big(\tilde{X}_{n,k}^{2d+1}\big) = O\big((\nu(k)n)^d\big), \qquad \text{for } d \ge 0.$$

Therefore, considering

$$\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}} = \frac{X_{n,k} - \mathbf{E}(X_{n,k})}{\sqrt{\nu(k)n}},$$

with $\nu(k) = \frac{2k^2-1}{k(k+1)^2(2k+1)}$, we have for $k = o(\sqrt{n}\,)$:

$$\mathbf{E}\left(\Big(\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}}\Big)^{2d}\right) \to \frac{(2d)!}{d!\,2^d}, \qquad \text{for } d \ge 1,$$

and

$$\mathbf{E}\left(\Big(\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}}\Big)^{2d+1}\right) = O\left(\frac{1}{\sqrt{\nu(k)n}}\right) = O\big(\frac{k}{\sqrt{n}}\big) \to 0, \qquad \text{for } \ d \geq 0.$$

For the subcritical case the moments of $\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}}$ converge to the moments of a standard normal distribution, which proves the convergence

$$\frac{X_{n,k} - \frac{n}{k(k+1)}}{\sqrt{\frac{(2k^2-1)n}{k(k+1)^2(2k+1)}}} \xrightarrow{D} \mathcal{N}(0,1).$$

Part (a) of Theorem 2.1 has been established.

4. The variety of subtree sizes in binary search trees. A binary search tree is constructed from the permutation $(\pi_1, \ldots, \pi_n)$ of the set $\{1, 2, \ldots, n\}$ by the following algorithm. The first element of the permutation is inserted in an empty tree, a root node is allocated for it. A subsequent element π_j (with $j \geq 2$) is directed to the left subtree if $\pi_j < \pi_1$, otherwise it is directed to the right subtree. In whichever subtree π_j goes, it is subjected to the same insertion algorithm recursively, until it is inserted in an empty subtree, in which case a node is allocated for it and linked appropriately as a left (right) child if its rank is less than (at least as much as) the value of the last node on the path.

Several models of randomness are in common use on binary trees. The uniform model in which all trees are equally likely has been proposed for applications in formal languages, compilers, computer algebra, etc. (see Kemp (1984)). However, for the searching and sorting algorithms alluded to the *random permutation model* is considered to be more appropriate. In this model of randomness we assume that the tree is built from permutations of $\{1, \ldots, n\}$, where a uniform probability model is imposed on the *permutations* instead of the trees. When all $n!$ permutations are equally likely or *random*, binary search trees are not equally likely. Several permutations give rise to the same tree, favoring shorter and well balanced trees rather than scrawny and tall shapes, which is a desirable property in searching and sorting algorithms (see Mahmoud (1992)). The term *random binary search tree* will refer to a binary search tree built from a random permutation. The random permutation model is not restrictive, as it covers a rather wide variety of instances, such as when the input is a sample drawn from *any* continuous probability distribution, and the construction algorithm is concerned only with the ranks of the keys, not their actual values.

Let $X_{n,k}$ be the random variable that counts the number of subtrees of size k on the fringe of a random binary search tree of size n, and let $M_k(z, v)$ be its generating function

$$M_k(z, v) = \sum_{n\geq 1} \sum_{m\geq 0} P\{X_{n,k} = m\} z^n v^m.$$

The binary dichotomy of the ranks in the permutation (relative to the first element) preserves the probabilistic structure in the conditionally independent subtrees. For all $n > k \geq 1$ the probabilities $P\{X_{n,k} = m\}$ satisfy the following recurrence:

$$P\{X_{n,k} = m\} = \frac{1}{n} \sum_{\substack{n_1+n_2=n-1 \\ m_1+m_2=m \\ m_1,m_2\geq 0, n_1,n_2\geq 0}} P\{X_{n_1,k} = m_1\} P\{X_{n_2,k} = m_2\},$$

with initial values $P\{X_{k,k} = 1\} = 1$, and $P\{X_{n,k} = 0\} = 1$, for $1 \le n < k$. After multiplication with $nz^{n-1}v^m$ and summing up over $n > k$ and $m \ge 0$, we obtain a functional equation for the generating function and find

$$\frac{\partial}{\partial z} M_k(z,v) = \bigl(1 + M_k(z,v)\bigr)^2 + (v-1)kz^{k-1},$$

for $k \ge 1$, with initial condition $M_k(0,v) = 0$.

Substituting $Q_k(z,v) := M_k(z,v) + 1$ we obtain the Riccati differential equation

$$\frac{\partial}{\partial z} Q_k(z,v) = Q_k^2(z,v) + (v-1)kz^{k-1}, \qquad Q_k(0,v) = 1.$$

The solution to this differential equation already appears in Flajolet, Gordon and Martínez (1997). Using this solution we obtain a representation for our generating function:

$$\begin{aligned} M_k(z,v) &= -1 \\ &+ \frac{1 + \sum_{j\ge1} \delta_j(k)(v-1)^j z^{(k+1)j} - \sum_{j\ge1} \gamma_j(k)(v-1)^j z^{(k+1)j-1}}{1 - z - \sum_{j\ge1} \beta_j(k)(v-1)^j z^{(k+1)j+1} + \sum_{j\ge1} \alpha_j(k)(v-1)^j z^{(k+1)j}}, \end{aligned} \tag{4.1}$$

with the functions

$$\alpha_j(k) = \xi_{j,-1,0}(k), \qquad \beta_j(k) = \xi_{j,1,0}(k), \tag{4.2}$$

$$\gamma_j(k) = \xi_{j,-1,-1}(k), \qquad \delta_j(k) = \xi_{j,1,1}(k), \tag{4.3}$$

where

$$\xi_{j,m,s}(k) = \frac{(-1)^j k^j \bigl((k+1)j + [\![s=1]\!]\bigr)^s}{\prod_{i=1}^{j} [(ik+i)(ik+i+m)]}.$$

4.1. The fac to rial moments for biñary search trees. From the solution (4.1) we can obtain the rth factorial moments, which lead directly to the limiting distribution for the critical and supercritical cases.

To get the rth factorial moments we use the substitution $w := v - 1$ in $M_k(z,v)$ and expand this function in powers of w. Straightforward computations leads then to the following expression:

$$\begin{aligned} M_k(z, 1+w) = -1 + \frac{1}{1-z}\Bigl(1 + \sum_{j\ge1} w^j \bigl(\delta_j(k) z^{(k+1)j} - \gamma_j(k) z^{(k+1)j-1}\bigr)\Bigr) \times \\ \times \sum_{\ell\ge0} \Bigl(\frac{z}{1-z} \sum_{j\ge1} w^j \bigl(\beta_j(k) z^{(k+1)j} - \alpha_j(k) z^{(k+1)j-1}\bigr)\Bigr)^\ell. \end{aligned}$$

Extract coefficients of w^r in $M_k(z, 1+w)$ eventually gives for $r \ge 1$:

$$\begin{aligned} &[w^r] M_k(z, 1+w) \\ &= \frac{z^{(k+1)r}}{(1-z)^{r+1}} \bigl(\beta_1(k) z - \alpha_1(k)\bigr)^r \end{aligned}$$

$$+z^{(k+1)r}\sum_{\ell=1}^{r-1}\frac{1}{(1-z)^{\ell+1}}\sum_{\substack{j_1+\cdots+j_\ell=r\\ j_i\geq 1}}\prod_{i=1}^{\ell}\big(\beta_{j_i}(k)z-\alpha_{j_i}(k)\big)$$
$$+z^{(k+1)r-1}\sum_{j=1}^{r-1}\big(\delta_j(k)z-\gamma_j(k)\big)\sum_{\ell=1}^{r-j}\frac{1}{(1-z)^{\ell+1}}$$
$$\times\sum_{\substack{j_1+\cdots+j_\ell=r-j\\ j_i\geq 1}}\prod_{i=1}^{\ell}\big(\beta_{j_i}(k)z-\alpha_{j_i}(k)\big)+\frac{z^{(k+1)r-1}}{1-z}\big(\delta_r(k)z-\gamma_r(k)\big).$$

Using

$$\big(\beta_1(k)z-\alpha_1(k)\big)^r=\big(\beta_1(k)-\alpha_1(k)\big)^r+\big(\beta_1(k)-\alpha_1(k)\big)^r$$
$$\times\sum_{m=1}^{r}\binom{r}{m}(-1)^m\Big(\frac{\beta_1(k)}{\beta_1(k)-\alpha_1(k)}\Big)^m(1-z)^m,$$

and analogous expansions, and $[z^n]\frac{1}{(1-z)^{\alpha+1}}=\binom{n+\alpha}{n}$, we can expand the above expression also around $z=1$ and get the following formula for the rth factorial moment of $X_{n,k}$, which is valid for all $n\geq(k+1)r$ and $r\geq 1$:

$$\mathbf{E}\big(X_{n,k}^{\underline{r}}\big)$$
$$=r!\,[z^nw^r]M_k(z,1+w)$$
$$=r!\binom{n-kr}{r}\Big(\frac{2}{(k+1)(k+2)}\Big)^r$$
$$+r!\Big(\frac{1}{(k+1)(k+2)}\Big)^r\sum_{p=1}^{r}\binom{n-(k+1)r+p-1}{p-1}\binom{r}{p-1}\Big(\frac{k}{2}\Big)^{r+1-p}$$
$$+r!\sum_{p=1}^{r}\binom{n-(k+1)r+p-1}{p-1}\sum_{\ell=\max(1,p-1)}^{r-1}(-1)^{\ell+1-p}$$
$$\times\sum_{\substack{j_1+\cdots+j_\ell=r\\ j_i\geq 1}}\prod_{i=1}^{\ell}\big(\beta_{j_i}(k)-\alpha_{j_i}(k)\big)$$
$$\times\sum_{1\leq i_1<\cdots<i_{\ell+1-p}\leq\ell}\prod_{q=1}^{\ell+1-p}\Big(\frac{\beta_{j_{i_q}}(k)}{\beta_{j_{i_q}}(k)-\alpha_{j_{i_q}}(k)}\Big)$$
$$+\,r!\sum_{p=1}^{r}\binom{n-(k+1)r+p}{p-1}\sum_{j=1}^{\min(r+1-p,r-1)}\big(\delta_j(k)-\gamma_j(k)\big)$$
$$\times\sum_{\ell=\max(1,p-1)}^{r-j}(-1)^{\ell+1-p}\sum_{\substack{j_1+\cdots+j_\ell=r-j\\ j_i\geq 1}}\prod_{i=1}^{\ell}\big(\beta_{j_i}(k)-\alpha_{j_i}(k)\big)$$
$$\times\sum_{1\leq i_1<\cdots<i_{\ell+1-p}\leq\ell}\prod_{q=1}^{\ell+1-p}\Big(\frac{\beta_{j_{i_q}}(k)}{\beta_{j_{i_q}}(k)-\alpha_{j_{i_q}}(k)}\Big)$$
$$-r!\sum_{p=1}^{r-1}\binom{n-(k+1)r+p}{p-1}\sum_{j=1}^{\min(r-p,r-1)}\delta_j(k)$$

$$\times \sum_{\ell=\max(1,p)}^{r-j} (-1)^{\ell-p} \sum_{\substack{j_1+\cdots+j_\ell=r-j\\ j_i\ge 1}} \prod_{i=1}^{\ell} \big(\beta_{j_i}(k)-\alpha_{j_i}(k)\big)$$

$$\times \sum_{1\le i_1<\cdots<i_{\ell-p}\le \ell} \prod_{q=1}^{\ell-p} \Big(\frac{\beta_{j_{i_q}}(k)}{\beta_{j_{i_q}}(k)-\alpha_{j_{i_q}}(k)}\Big) + r!\,\big(\delta_r(k)-\gamma_r(k)\big). \tag{4.4}$$

This yields in particular to results for the expectation and the variance:

$$\mathbf{E}(X_{n,k}) = \frac{2(n+1)}{(k+1)(k+2)}, \qquad \text{for } n\ge k+1, \tag{4.5}$$
$$\mathbf{V}(X_{n,k}) = \frac{2k(4k^2+5k-3)(n+1)}{(k+1)(k+2)^2(2k+1)(2k+3)}, \qquad \text{for } \ n\ge 2(k+1).$$

4.2. The critical case. We consider the critical case and assume that $k:=k_n$ grows with n such that $\frac{n}{k^2}\to\lambda$, for some $\lambda>0$.

Finding the limiting distribution requires some estimates on the functions appearing in (4.4). Using the definition of $\alpha_j(k)$, $\beta_j(k)$, $\gamma_j(k)$ and $\delta_j(k)$ given by (4.2)–(4.3) the following estimates, which hold for all $k\ge 2$ and $j\ge 1$, are not hard to show:

$$|\beta_j(k)-\alpha_j(k)| \le \frac{2}{k^{j+1}}, \qquad \left|\frac{\beta_j(k)}{\beta_j(k)-\alpha_j(k)}\right| \le 4^j k, \tag{4.6}$$

$$|\delta_j(k)-\gamma_j(k)| \le \frac{3}{k^j}, \qquad |\delta_j(k)| \le \frac{1}{k^{j-1}}. \tag{4.7}$$

We can now deduce the asymptotic behavior of the rth factorial moments of $X_{n,k}$ by inspecting the summands of (4.4). Considering a fixed $r\ge 1$ we trivially obtain for the first summand:

$$r!\binom{n-kr}{r}\Big(\frac{2}{(k+1)(k+2)}\Big)^r = \Big(\frac{2n}{k^2}\Big)^r\Big(1+O\Big(\frac{1}{k}\Big)+\Big(\frac{k}{n}\Big)\Big)$$
$$= (2\lambda)^r\Big(1+O\Big(\frac{1}{\sqrt{n}}\Big)\Big).$$

Let us denote by B the remaining summands of (4.4). They can then be estimated for fixed $r\ge 1$ by using (4.6) and (4.7), which finally gives

$$|B| = O\Big(\frac{1}{k}\Big) = O\Big(\frac{1}{\sqrt{n}}\Big).$$

Summarizing, we obtain for all $r\ge 1$ and sequences (n,k) such that $\frac{n}{k^2}\to\lambda$ the following asymptotic expansion:

$$\mathbf{E}\big(X_{n,k}^{\underline{r}}\big) = (2\lambda)^r\Big(1+O\Big(\frac{1}{\sqrt{n}}\Big)\Big).$$

This shows convergence in distribution of $X_{n,k}$ to a Poisson random variable with parameter 2λ. Thus we proved the convergence asserted in Theorem 2.2 Part (b), where we used the substitution $c:=\frac{1}{\sqrt{\lambda}}$, and thus $\frac{k}{\sqrt{n}}\to c$. When we have a critical case with $k/\sqrt{n}$ not converging to a limit, all factorial (and ordinary) moments oscillate, and no limit distribution exists for $X_{n,k}$.

4.3. The supercritical case. Again by using the estimates (4.6) and (4.7) one can show $\mathbf{E}(X_{n,k}) \to 0$, for $\frac{n}{k^2} = o(1)$ under the natural restriction $n > k$. Thus, we have convergence of $X_{n,k}$ to a degenerate distribution with mass at 0.

4.4. The subcritical case. We consider the centered random variables $\tilde{X}_{n,k} := X_{n,k} - \mathbf{E}(X_{n,k})$ and introduce the generating function

$$\tilde{M}_k(z,s) := \sum_{n\geq 1} \mathbf{E}\big(e^{\tilde{X}_{n,k}s}\big) z^n = \sum_{n\geq 1} e^{-\mathbf{E}(X_{n,k})s}\mathbf{E}\big(e^{X_{n,k}s}\big) z^n.$$

Using the explicit formula for $\mathbf{E}(X_{n,k})$ as given by (4.5) we obtain

$$\begin{aligned}\tilde{M}_k(z,s) &= e^{-\frac{2s}{(k+1)(k+2)}} M_k\big(e^{-\frac{2s}{(k+1)(k+2)}} z, e^s\big) + (1 - e^{\frac{ks}{k+2}})z^k \\ &\quad + \sum_{1\leq n<k} z^n - \sum_{1\leq n<k} e^{-\frac{2(n+1)s}{(k+1)(k+2)}} z^n,\end{aligned}$$

which further leads by using (4.1) and (4.2)–(4.3) to:

$$\begin{aligned}\tilde{M}_k(z,s) = \quad & \frac{e^{-\frac{2s}{(k+1)(k+2)}}U}{V} - e^{-\frac{2s}{(k+1)(k+2)}} + (1 - e^{\frac{ks}{k+2}})z^k \\ & + \sum_{1\leq n<k} z^n - \sum_{1\leq n<k} e^{-\frac{2(n+1)s}{(k+1)(k+2)}} z^n,\end{aligned}$$

with the expressions

$$\begin{aligned}U &:= 1 + \sum_{j\geq 1} \delta_j(k)(e^s-1)^j e^{-\frac{2js}{k+2}} z^{(k+1)j} \\ &\quad - \sum_{j\geq 1} \gamma_j(k)(e^s-1)^j e^{-\frac{2((k+1)j-1)s}{(k+1)(k+2)}} z^{(k+1)j-1}, \\ V &:= 1 - e^{-\frac{2s}{(k+1)(k+2)}} z - \sum_{j\geq 1} \beta_j(k)(e^s-1)^j e^{-\frac{2((k+1)j+1)s}{(k+1)(k+2)}} z^{(k+1)j+1} \\ &\quad + \sum_{j\geq 1} \alpha_j(k)(e^s-1)^j e^{-\frac{2js}{k+2}} z^{(k+1)j}.\end{aligned}$$

4.4.1. Expanding around $s = 0$. We expand $\tilde{M}_k(z,s)$ around $s = 0$ and $z = 1$. It is straightforward, though rather laborious, to come up with the following representation:

$$M_k\big(e^{-\frac{2s}{(k+1)(k+2)}} z, e^s\big) = -1 + \frac{\sum_{\ell\geq 0} s^l \sum_{i=0}^{\ell(k+1)} d_{\ell,i}(k)(1-z)^i}{1 - z - \sum_{\ell\geq 1} s^\ell \sum_{i=0}^{\ell(k+1)+1} c_{\ell,i}(k)(1-z)^i},$$

where the functions $d_{\ell,i}(k)$ and $c_{\ell,i}(k)$ are given as follows:

$$\begin{aligned}d_{\ell,i}(k) &= (-1)^i \sum_{j=1}^{\ell} \sum_{m=0}^{\ell-j} \frac{j! \left\{ {\ell-m \atop j} \right\} (-1)^m}{m!\,(\ell-m)!\,(k+2)^m} \\ &\quad \times \bigg(\binom{(k+1)j-1}{i} \Big((2j)^m \delta_j(k) - \frac{(2((k+1)j-1))^m \gamma_j(k)}{(k+1)^m} \Big)\end{aligned}$$

$$+\binom{(k+1)j-1}{i-1}(2j)^m\delta_j(k)[\![i\geq 1]\!]\Big), \qquad \text{for } \ell\geq 1, \tag{4.8}$$

$$d_{0,0}(k)=1,$$

$$c_{\ell,i}(k)=(-1)^i\sum_{j=1}^{\ell}\sum_{m=0}^{\ell-j}\frac{\left\{{l-m \atop j}\right\}j!(-1)^m}{m!\,(\ell-m)!\,(k+2)^m}$$
$$\times\Bigg(\binom{(k+1)j}{i}\Big(\frac{(2((k+1)j+1))^m\beta_j(k)}{(k+1)^m}-(2j)^m\alpha_j(k)\Big)$$
$$+\binom{(k+1)j}{i-1}\frac{(2((k+1)j+1))^m\beta_j(k)}{(k+1)^m}[\![i\geq 1]\!]\Bigg)$$
$$+\frac{(-1)^\ell 2^\ell}{\ell!(k+1)^\ell(k+2)^\ell}[\![i=0]\!]-\frac{(-1)^\ell 2^\ell}{\ell!(k+1)^\ell(k+2)^\ell}[\![i=1]\!]. \tag{4.9}$$

Thus, $\tilde{M}_k(z,s)$ can be written as

$$\tilde{M}_k(z,s)=\frac{e^{-\frac{2s}{(k+1)(k+2)}}\left(\sum_{\ell\geq 0}s^\ell\sum_{i=0}^{\ell(k+1)}d_{\ell,i}(k)(1-z)^i\right)}{1-z-\sum_{\ell\geq 1}s^\ell\sum_{i=0}^{\ell(k+1)+1}c_{\ell,i}(k)(1-z)^i}-e^{-\frac{2s}{(k+1)(k+2)}}$$
$$+(1-e^{\frac{ks}{k+2}})z^k+\sum_{1\leq n<k}z^n-\sum_{1\leq n<k}e^{-\frac{2(n+1)s}{(k+1)(k+2)}}z^n. \tag{4.10}$$

The only interesting part in equation (4.10) is the first summand, since the remaining summands do not give a contribution for $n>k$ (the remainder is a polynomial in z of degree k). The first summand of (4.10) can be expanded around $s=0$ and $z=1$ and finally leads for $r\geq 1$ to the following representation:

$$[s^r]\frac{e^{-\frac{2s}{(k+1)(k+2)}}\left(\sum_{\ell\geq 0}s^\ell\sum_{i=0}^{\ell(k+1)}d_{\ell,i}(k)(1-z)^i\right)}{1-z-\sum_{\ell\geq 1}s^\ell\sum_{i=0}^{\ell(k+1)+1}c_{\ell,i}(k)(1-z)^i}$$
$$=\sum_{p=0}^{r}\frac{1}{(1-z)^{p+1}}f_{r,p}(k)+\sum_{p=0}^{r-1}\frac{1}{(1-z)^{p+1}}g_{r,p}(k)+\sum_{p=0}^{r(k+1)-1}h_{r,p}(k)z^p,$$

with

$$f_{r,p}(k)=\sum_{m=p}^{r}\sum_{\substack{r_1+\cdots+r_m=r\\ r_q\geq 1}}\sum_{\substack{t_1+\cdots+t_m=m-p\\ 0\leq t_q\leq r_q(k+1)+1}}\prod_{j=1}^{m}c_{r_j,t_j}(k), \tag{4.11}$$

$$g_{r,p}(k)=\sum_{c=1}^{r}\sum_{a=p}^{r-c}\Big(\sum_{m=a}^{r-c}\sum_{\substack{r_1+\cdots+r_m=r-c\\ r_q\geq 1}}\sum_{\substack{t_1+\cdots+t_m=m-a\\ 0\leq t_q\leq r_q(k+1)+1}}\prod_{j=1}^{m}c_{r_j,t_j}(k)\Big)$$
$$\times\left(\sum_{b=\lceil\frac{a-p}{k+1}\rceil}^{c}\frac{(-1)^{c-b}2^{c-b}d_{b,a-p}(k)}{(k+1)^{c-b}(k+2)^{c-b}(c-b)!}\right), \tag{4.12}$$

and certain functions $h_{r,p}(k)$, which are of course irrelevant for our purpose.

This leads also to the required representation

$$[s^r]\tilde{M}_k(z,s)=\sum_{p=0}^{r}\frac{1}{(1-z)^{p+1}}f_{r,p}(k)+\sum_{p=0}^{r-1}\frac{1}{(1-z)^{p+1}}g_{r,p}(k)+\sum_{p=0}^{r(k+1)-1}\tilde{h}_{r,p}(k)z^p,$$

where the functions $f_{r,p}(k)$ and $g_{r,p}(k)$ are defined by (4.11) and (4.12). The functions $\tilde{h}_{r,p}(k)$ are irrelevant and thus not given explicitly.

The representation in the last display gives the following explicit formula for the rth moments of $\tilde{X}_{n,k}$, which is valid for $n \ge r(k+1)$:

$$\mathbf{E}\big(\tilde{X}_{n,k}^r\big) = r!\,[z^n s^r]\tilde{M}_k(z,s) = r! \sum_{p=0}^{r} \binom{n+p}{p} f_{r,p}(k) + r! \sum_{p=0}^{r} \binom{n+p}{p} g_{r,p}(k),$$

where the functions $f_{r,p}(k)$ and $g_{r,p}(k)$ are given by (4.11) and (4.12).

4.4.2. Estimates for $c_{\ell,i}(k)$ and $d_{\ell,i}(k)$. One obtains easily the following estimates of the functions $c_{\ell,i}(k)$ and $d_{\ell,i}(k)$ given by (4.8) and (4.9):

$$|d_{\ell,i}(k)| \le 2\ell^2(2\ell)^{2\ell+1}B_\ell(2\ell)^i k^{i-1}, \qquad \text{for } \ell \ge 1, i \ge 0, k \ge 1,$$
$$|c_{\ell,i}(k)| \le B_\ell 2^{2\ell+3}\ell^{\ell+3}(2\ell)^i k^{i-2}, \qquad \text{for } \ell \ge 1, i \ge 0, k \ge 1.$$

Moreover, we have the following values:

$$c_{1,0}(k) = 0, \qquad \text{and} \qquad c_{2,0}(k) = \frac{\nu(k)}{2},$$

with

$$\nu(k) := \frac{2k(4k^2+5k-3)}{(k+1)(k+2)^2(2k+1)(2k+3)}.$$

4.4.3. Estimates for $f_{r,p}(k)$ and $g_{r,p}(k)$. It is possible to give suitable estimates on the growth of $f_{r,p}(k)$ and $g_{r,p}(k)$: There exist constants κ_r and η_r (depending only on r), such that

$$|f_{r,p}(k)| \le \kappa_r \frac{1}{k^{2p}}, \qquad \text{and} \qquad |g_{r,p}(k)| \le \eta_r \frac{1}{k^{2p+2}},$$

for all $0 \le p \le r$ (respectively $0 \le p \le r-1$) and $k \ge 1$.

For instance one can choose the constants

$$\kappa_r = \frac{(2r-1)!\,(r+1)}{r!} c_r^r, \qquad \eta_r = 2^{r+2}\binom{2r-1}{r-1} r!\, r c_r^r\, d_r,$$

with

$$c_r := B_r 2^{2r+3} r^{r+3}(2r)^r, \qquad \text{and} \qquad d_r := 2B_r r^r (2r)^{3r+1}.$$

4.4.4. Cancelation of $\tilde{f}_{r,p}(k)$. Using the same "combinatorial argumentation" as made for recursive trees we can show:

$$f_{r,p}(k) = 0, \qquad \text{for } p \ge \left\lfloor \frac{r}{2} \right\rfloor + 1;$$
$$g_{r,p}(k) = 0, \qquad \text{for } p \ge \left\lceil \frac{r}{2} \right\rceil.$$

Furthermore, one can show

$$f_{2d,d}(k) = c_{2,0}^d(k) = \frac{\nu^d(k)}{2^d}.$$

4.4.5. Asymptotics of the centered moments. In view of the development in subsection 4.4.3 we obtain the explicit formulas for the rth moments of $X_{n,k}$:

$$\begin{aligned}\mathbf{E}\big(\tilde{X}_{n,k}^{2d}\big) &= (2d)!\left(\binom{n+d}{d} f_{2d,d}(k) \right.\\ &\qquad \left. + \sum_{p=0}^{d-1}\binom{n+p}{p}\big(f_{2d,p}(k)+g_{2d,p}(k)\big)\right), \qquad d \ge 1,\\ \mathbf{E}\big(\tilde{X}_{n,k}^{2d+1}\big) &= (2d+1)!\sum_{p=0}^{d}\binom{n+p}{p}\big(f_{2d+1,p}(k)+g_{2d+1,p}(k)\big), \qquad d \ge 0.\end{aligned}$$

Using the previous estimates we can easily show that, for $\frac{k^2}{n} \to 0$,

$$\begin{aligned}\mathbf{E}\big(\tilde{X}_{n,k}^{2d}\big) &\to \frac{(2d)!}{d!\,2^d} n^d \nu^d(k),\\ \mathbf{E}\big(\tilde{X}_{n,k}^{2d+1}\big) &= O\Big(\Big(\frac{n}{k^2}\Big)^d\Big) = O\big(n^d\nu^d(k)\big).\end{aligned}$$

Therefore, considering

$$\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}} = \frac{X_{n,k} - \mathbf{E}(X_{n,k})}{\sqrt{\nu(k)n}},$$

with $\nu(k) = \frac{2k(4k^2+5k-3)}{(k+1)(k+2)^2(2k+1)(2k+3)}$, we have for $k = o(\sqrt{n})$:

$$\begin{aligned}\mathbf{E}\left(\Big(\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}}\Big)^{2d}\right) &\to \frac{(2d)!}{d!\,2^d}, \qquad \text{for } d \ge 1,\\ \mathbf{E}\left(\Big(\frac{\tilde{X}_{n,k}}{\sqrt{\nu(k)n}}\Big)^{2d+1}\right) &= O\Big(\frac{k}{\sqrt{n}}\Big) \to 0, \qquad \text{for } d \ge 0.\end{aligned}$$

Whence, for the subcritical case we have

$$\frac{X_{n,k} - \frac{2n}{(k+1)(k+2)}}{\sqrt{\frac{2k(4k^2+5k-3)n}{(k+1)(k+2)^2(2k+1)(2k+3)}}} \xrightarrow{D} \mathcal{N}(0,1),$$

completing the proof for the subcritical case (Part (a) of Theorem 2.2).

5. Concluding remarks. Several methods have recently come into the repertoire of the analysis of algorithms, and there is a question of the choice when it comes to an analysis like the one we conducted in this investigation. For example, the contraction method (introduced in Rösler, 1991), Pólya urns (Johnson and Kotz, 1977), the Chen-Stein method (Barbour, Holst, and Janson ,1992), methods for additive functions in subtrees (Devroye, 1991 and Devroye, 2003) are all attractive for this kind of analysis. These methods have been used successfully in the analysis of algorithms, (see for example Neininger, 2002, or Rösler and Rüschendorf, 2001 for applications of the contraction method). However, Experience shows that several of these methods will work with ease in some of the phases, but not systematically across all the phases. For instance, in our multi-phase problem Pólya urns work well in the very low range (small fixed k), as was attempted in Feng, Mahmoud and Su (2007). The difficulty in this method is that one needs a detailed description of the fringe in terms of urns that are different from each other for each k, and that become more complex as k increases. As for the contraction method , again it can be used in the case of small fixed k, but does not lend itself easily to $k = k_n \to \infty$ in the subcritical case when k increases but at rate slower than $\sqrt{n}$. The difficulty stems from complicated toll functions. It is not clear how to apply the contraction method in the critical case. The same applies to the other methods above. It is our experience that the analytic methods, based on generating functions, are systematic enough to cover all the ranges.

Acknowledgment. The work of the first author (Q. Feng) is supported by the National Natural Science Foundation of China (Grant No. 10371117) and the Special Foundation of USTC. The work of the third author (A. Panholzer) is supported by the Austrian Science Foundation FWF (Grant No. S9608-N13).

REFERENCES

[1] Barbour, A., Lars Holst, L. and Janson S. (1992). *Poisson Approximation*. Oxford University Press.

[2] Bergeron, F. Flajolet, P. and Salvy, B. (1992). Varieties of increasing trees. *Lecture Notes in Computer Science*, **581**, 24–48. Springer, Berlin.

[3] Chern, H., Hwang, H. and Tsai, T. (2002). An asymptotic theory for Cauchy-Euler differential equations with applications to the analysis of algorithms. *Journal of Algorithms*, **44**, 177–225.

[4] Chyzak, F., Drmota, M., Klausner, T. and Kok, G. (2007+). The distribution of patterns in random trees, *Combinatorics, Probability and Computing*, to appear.

[5] Devroye, L. (1991). Limit laws for local counters in random binary search trees. *Random Structures and Algorithms*, **3**, 303–315.

[6] Devroye, L. (2003). Limit laws for sums of functions of subtrees of random binary search trees. *SIAM Journal on Computing*, **32**, 152–171

[7] Feng, Q., Mahmoud, H. and Su, C. (2007). On the variety of subtrees in a random recursive tree. Technical Report, The George Washington University.

[8] Flajolet, P., Gourdon, X. and Martínez, C. (1997). Patterns in random binary search trees. *Random Structures and Algorithms*, **11**, 223–244.

[9] Fuchs, M., Hwang, H.-K. and Neininger, R. (2006). Profiles of random trees: Limit theorems for random recursive trees and binary search trees. *Algorithmica*, **46**, 367–407.

[10] Gastwirth, J. and Bhattacharya, P. (1984). Two probability models of pyramids or chain letter schemes demonstrating that their promotional claims are unreliable. *Operations Research*, **32**, 527–536.

[11] Johnson, N. and Kotz, S. (1977). *Urn models and Their Applications*. Wiley, New York.

[12] Kemp, R. (1984). *Fundamentals of the Average Case Analysis of Particular Algorithms*. Wiley, New York.

[13] Knuth, D. (1998). *The Art of Computer Programming*, Vol. 3: *Sorting and Searching*, 2nd ed. Addison-Wesley, Reading, Massachusetts.

[14] Lent, J. and Mahmoud, H. (1996). Average-case analysis of multiple Quickselect: An algorithm for finding order statistics. *Statistics and Probability Letters*, **28**, 299–310.

[15] Mahmoud, H. (1992). *Evolution of Random Search Trees*. Wiley, New York.

[16] Mahmoud, H. (2000). *Sorting: A Distribution Theory*. Wiley, New York.

[17] Najock, D. and Heyde, C. (1982). On the number of terminal vertices in certain random trees with an application to stemma construction in philology. *Journal of Applied Probability*, **19**, 675–680.

[18] Neininger, R. (2002). The Wiener index of random trees. *Combinatorics, Probability and Computing*, **11**, 587–597.

[19] Rösler, U. (1991). A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl.*, **25**, 85–100.

[20] Rösler, U. and Rüschendorf, L. (2001). The contraction method for recursive algorithms. *Algorithmica*, **29**, 3–33.

[21] Panholzer, A. and Prodinger, H. (1998). A generating functions approach for the analysis of

grand averages for multiple QUICKSELECT. *Random Structures and Algorithms*, **13**, 189–209.

[22] Smythe, R. and Mahmoud, H. (1994). A survey of recursive trees. *Theorya Imovirnosty ta Matemika Statystika*, **51**, 1–29 (in Ukrainian); an English translation appears in *Theory of Probability and Its Applications* (1996).