Dependence and phase changes in random m-ary search trees

Hua-Huai Chern
Department of Computer Science
National Taiwan Ocean University
Keelung 202
Taiwan

Michael Fuchs*
Department of Applied Mathematics
National Chiao Tung University
Hsinchu 300
Taiwan

Hsien-Kuei Hwang†
Institute of Statistical Science
Academia Sinica
Taipei 115
Taiwan

Ralph Neininger‡
Institute for Mathematics
Goethe University
60054 Frankfurt a.M.
Germany

January 30, 2016

Abstract

We study the joint asymptotic behavior of the space requirement and the total path length (either summing over all root-key distances or over all root-node distances) in random m-ary search trees. The covariance turns out to exhibit a change of asymptotic behavior: it is essentially linear when $3 \leqslant m \leqslant 13$ but becomes of higher order when $m \geqslant 14$. Surprisingly, the corresponding asymptotic correlation coefficient tends to zero when $3 \leqslant m \leqslant 26$ but is periodically oscillating for larger m. Such a less anticipated phenomenon is not exceptional and we extend the results in two directions: one for more general shape parameters, and the other for other classes of random log-trees such as fringe-balanced binary search trees and quadtrees. The methods of proof combine asymptotic transfer for the underlying recurrence relations with the contraction method.

AMS 2010 subject classifications. Primary 60F05, 68Q25; secondary 68P05, 60C05, 05A16.

Key words. m-ary search tree, correlation, dependence, recurrence relations, fringe-balanced binary search tree, quadtree, asymptotic analysis, limit law, asymptotic transfer, contraction method.

*Partially supported by the Ministry of Science and Technology, Taiwan under the grant MOST-103-2115-M-009-007-MY2.

†This author's research stay at J. W. Goethe-Universität was partially supported by the Simons Foundation and by the Mathematisches Forschungsinstitut Oberwolfach.

‡Supported by DFG grant NE 828/2-1.

1 Introduction

The m-ary search trees are a class of data structures introduced by Muntz and Uzgalis [35] in 1971 in computer algorithms to support efficient searching and sorting of data; see the next section for more details. When constructed from a random permutation of n elements, the space requirement (total number of nodes to store the input) S_n of such *random m-ary search trees* ($m \geqslant 3$) is known to exhibit a *phase change phenomenon*: its distribution is asymptotically Gaussian for large n when the branching factor m satisfies $3 \leqslant m \leqslant 26$ but does not approach a limit law when $m \geqslant 27$; see [8, 22, 30, 31] and the references therein. On the other hand, it is also known that the total key path length K_n (the sum over all distances from the root to any *key*) does not change its limiting behavior when m varies, and tends asymptotically, after properly centered and normalized, to a limit law for each $m \geqslant 3$. Another closely related shape measure, the total node path length N_n (summing over all distances from the root to any *node*) also follows asymptotically a very similar behavior.

Our motivating question was "how does K_n or N_n depend on S_n?" Surprisingly, despite the strong dependence of the definition of N_n on S_n (see (2)), we show that the correlation coefficient $\rho(S_n, N_n)$ satisfies

$$\rho(S_n, N_n) \sim \begin{cases} 0, & \text{if } 3 \leqslant m \leqslant 26; \\ F_\rho(\beta \log n), & \text{if } m \geqslant 27, \end{cases} \tag{1}$$

where $F_\rho(t)$ is a 2π-periodic function and $\beta = \beta_m$ is a structural constant depending on m. The same type of results also holds for $\rho(S_n, K_n)$. In words, *N_n and S_n are asymptotically uncorrelated for $3 \leqslant m \leqslant 26$ and their correlation fluctuates (between -1 and 1) for $m \geqslant 27$*; see Figure 1 for an illustration.

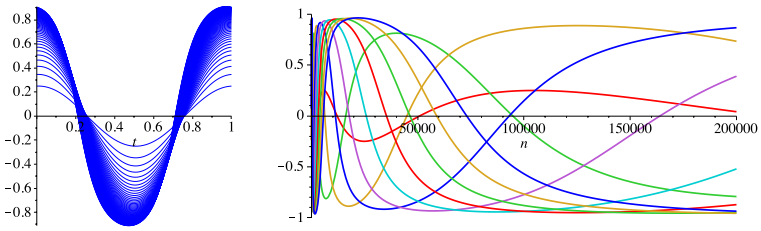

Figure 1: *The periodic functions $F_\rho(2\pi t)$ for $m = 27, \ldots, 100$ (left) and $F_\rho(\beta \log n)$ for $m = 27, 54, \ldots, 270$ (right).*

One reason why the above result (1) may seem less or even counter-intuitive is because of the seemingly strong dependence of N_n on S_n in the recursive equations satisfied by both random variables

$$\begin{cases} S_n \overset{d}{=} S_{I_1}^{(1)} + \cdots + S_{I_m}^{(m)} + 1, \\ N_n \overset{d}{=} N_{I_1}^{(1)} + \cdots + N_{I_m}^{(m)} + S_{I_1}^{(1)} + \cdots + S_{I_m}^{(m)}, \end{cases} \tag{2}$$

where the $(S_i^{(r)}, N_i^{(r)})$'s are independent copies of (S_i, N_i), respectively, also independent of $(I_1, \ldots, I_m)$, and

$$\mathbb{P}(I_1 = i_1, \ldots, I_m = i_m) = \frac{1}{\binom{n}{m-1}}, \tag{3}$$

when $i_1, \dots, i_m \geqslant 0$ and $i_1 + \cdots + i_m = n - m + 1$. Intuitively, we expect, from the above relations, that the node path length N_n would have a strong correlation with S_n.

While one might ascribe this seemingly less intuitive result to the possibly nonlinear dependence between N_n and S_n, we enhance such an uncorrelation by a stronger joint limit law for (S_n, N_n) for $3 \leqslant m \leqslant 26$, which further accents the asymptotic independence between N_n and S_n; for $m \geqslant 27$, they are asymptotically dependent and we will derive a precise characterization of their joint asymptotic distributions. See Section 4 for a more precise description of the joint asymptotic behaviors of (S_n, N_n) and (S_n, K_n).

Let α denote the real part of the second largest zero (in real parts) of the indicial equation $\Lambda(z) = 0$, where

$$\Lambda(z) = z(z+1)\cdots(z+m-2) - m!. \tag{4}$$

Then $\alpha < 1$ for $m < 14$ and $1 < \alpha < \frac{3}{2}$ for $14 \leqslant m \leqslant 26$; see Table 1. The main reason that

m	3	4	5	6	7	8	9	10
α	-3	-2.5	-1.5	-0.768	-0.260	0.101	0.366	0.568
m	11	12	13	14	15	16	17	18
α	0.726	0.852	0.955	1.040	1.112	1.173	1.226	1.272
m	19	20	21	22	23	24	25	26
α	1.313	1.348	1.380	1.409	1.435	1.458	1.479	1.499

Table 1: *Approximate numerical values of* $\alpha = \alpha_m$ *for* $3 \leqslant m \leqslant 26$.

$\rho(S_n, N_n) \to 0$ for $3 \leqslant m \leqslant 26$ is roughly that their covariance is of order $\max\{n \log n, n^\alpha\}$ (see Theorem 2.3 below), while the standard deviations for S_n and N_n are of orders $\sqrt{n}$ and n, respectively. So that

$$\rho(S_n, N_n) = \begin{cases} O\left(n^{-\frac{1}{2}} \log n\right), & \text{if } 3 \leqslant m \leqslant 13; \\ O\left(n^{-\frac{3}{2}+\alpha}\right), & \text{if } 14 \leqslant m \leqslant 26, \end{cases}$$

which tends to zero in both cases. Briefly, *the large quadratic variance of* N_n *is the major cause of the asymptotic independence between* S_n *and* N_n *for* $3 \leqslant m \leqslant 26$.

Such a change from being asymptotically independent to being asymptotically dependent under a varying structural parameter is not an exception. We will extend our study to fringe-balanced binary search trees and quadtrees; a typical related instance states that: *the number of comparisons (or exchanges) used by the median-of-*$(2t+1)$ *quicksort is asymptotically independent of the number of partitioning stages when* $0 \leqslant t \leqslant 58$*, but is asymptotically dependent for* $t \geqslant 59$.

2 M-ary search trees

We briefly introduce m-ary search trees in this section and then describe the random variables we are studying in this paper.

An *m-ary tree* is either empty or comprises of a single node called the root, together with an ordered m-tuple of subtrees, each of which is, by definition, an m-ary tree. Given a sequence

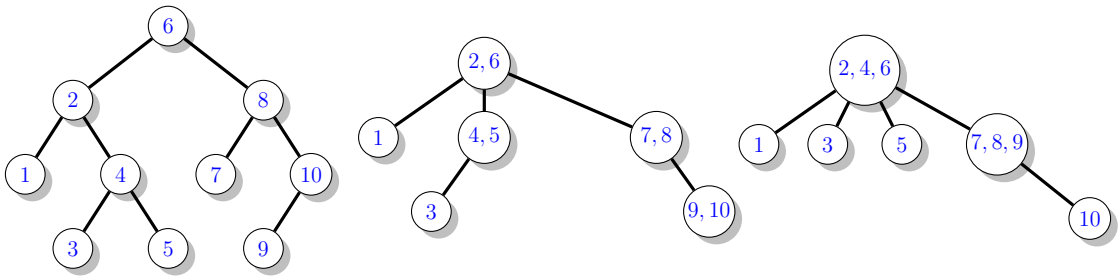

Figure 2: *Three m-ary search trees for the sequence $\{6, 2, 4, 8, 7, 1, 5, 3, 10, 9\}$: $m = 2$ (left), $m = 3$ (middle), and $m = 4$ (right).*

of numbers, say $\{x_1, \ldots, x_n\}$, we construct an m-ary search tree by the following procedure, $m \geqslant 2$. If $1 \leqslant n < m$, then all keys are stored in the root. If $n \geqslant m$ the first $m-1$ keys are sorted and stored in the root, the remaining keys are directed to the m subtrees, each corresponding to one of the m intervals formed by the $m-1$ sorted keys in the root node; see Figure 2 for an illustration (the rectangular nodes denote yet empty subtrees of full nodes). If the $m-1$ numbers in the root are $x_{j_1} < \cdots < x_{j_{m-1}}$, then the keys directed to the ith subtree all have their values lying between $x_{j_{i-1}}$ and x_{j_i}, where $x_{j_0} := 0$ and $x_{j_m} := n+1$. All subtrees are themselves m-ary search trees by definition. For more details, see Mahmoud [30].

While the practical usefulness of m-ary search trees is largely overshadowed by their balanced counterparts such as B-trees, they have been a source of many interesting phenomena, which are to some extent universal. The study of m-ary search trees is thus of fundamental and prototypical value. Furthermore, the close connection between m-ary search trees and generalized quicksort adds an extra dimension to the richness of diverse variations and their asymptotic behaviors.

2.1 Space requirement and total path lengths

Assume that the input sequence $\{x_1, \ldots, x_n\}$ is a random permutation, where all $n!$ permutations are equally likely. The resulting m-ary search tree constructed from the given sequence is then called a random m-ary search tree.

The major shape parameters of particular algorithmic interest include the depth, the height, the space requirement, the total path length, and the profile; see [11, 30] for more information. We are concerned in this paper with the following three random variables.

- S_n (space requirement): the total number of nodes used to store the input; the three trees in Figure 2 have S_{10} equal to $10, 6, 6$, respectively. If $m = 2$, then $S_n \equiv n$; if $m \geqslant 3$, we can compute S_n recursively by $S_0 = 0$, and

$$S_n \stackrel{d}{=} \begin{cases} 1, & \text{if } 1 \leqslant n < m, \\ S_{I_1}^{(1)} + \cdots + S_{I_m}^{(m)} + 1, & \text{if } n \geqslant m, \end{cases} \tag{5}$$

  where the $S_i^{(r)}$'s are independent copies of S_i, $1 \leqslant r \leqslant m$, $0 \leqslant i \leqslant n-m+1$, and independent of $(I_1, \ldots, I_m)$ defined in (3).

- K_n (key path length, KPL): the sum of the distance between the root and each key; for the trees in Figure 2, $K_{10} = \{19, 11, 8\}$, respectively. For $m \geqslant 2$, K_n satisfies the recurrence

$$K_n \stackrel{d}{=} \begin{cases} 0, & \text{if } n < m, \\ K_{I_1}^{(1)} + \cdots + K_{I_m}^{(m)} + n - m + 1, & \text{if } n \geqslant m, \end{cases} \tag{6}$$

  where the $K_i^{(r)}$'s are independent copies of K_i, $1 \leqslant r \leqslant m, 0 \leqslant i \leqslant n - m + 1$, independent of $(I_1, \ldots, I_m)$.

- N_n (node path length, NPL): the sum of the distance between the root and each node; so that $N_{10} = \{19, 7, 6\}$ for the three trees in Figure 2. Obviously, $N_n = K_n$ when $m = 2$. When $m \geqslant 3$,

$$N_n \stackrel{d}{=} \begin{cases} 0, & \text{if } n < m, \\ N_{I_1}^{(1)} + \cdots + N_{I_m}^{(m)} + S_{I_1}^{(1)} + \cdots + S_{I_m}^{(m)}, & \text{if } n \geqslant m, \end{cases} \tag{7}$$

  where the $(N_i^{(r)}, S_i^{(r)})$'s are independent copies of (N_i, S_i), $1 \leqslant r \leqslant m, 0 \leqslant i \leqslant n - m + 1$, independent of $(I_1, \ldots, I_m)$.

While the first two random variables have been widely studied in the literature, NPL was only considered previously in [4, 21] in connection with the process of cutting trees. In addition to this, our interest was to understand the extent to which the asymptotic independence for small m between S_n and K_n subsists when the "toll function" changes from a linear function to a function that is random and may depend on S_n.

2.2 A summary of known results

Let $H_m := \sum_{1 \leqslant j \leqslant m} j^{-1}$. Knuth [27, §6.2.4] was the first to show that

$$\mathbb{E}(S_n) \sim \phi n, \quad \text{where} \quad \phi := \frac{1}{2(H_m - 1)}.$$

(see also [1]). Here ϕ denotes the "occupancy constant", which will appear all over our analysis. Mahmoud and Pittel [31] improved the result and derived an identity for $\mathbb{E}(S_n)$, which implies in particular that

$$\mathbb{E}(S_n) = \phi(n+1) - \frac{1}{m-1} + O\left(n^{\alpha-1}\right),$$

where α has the same meaning as in Introduction; see (4). They also discovered and proved the surprising result for the variance

$$\mathbb{V}(S_n) \sim \begin{cases} C_S n, & \text{if } 3 \leqslant m \leqslant 26; \\ F_1(\beta \log n) n^{2\alpha-2}, & \text{if } m \geqslant 27, \end{cases}$$

where C_S is a constant depending on m, F_1 is a π-periodic function given in (22), $\alpha + i\beta$ is the second largest zero (in real part) with $\beta > 0$ of the equation $\Lambda(z) = 0$ (see (4)), and $2\alpha - 2 > 1$ for $m \geqslant 27$. See also [9, 25, 33] for a closely related fragmentation model with the same asymptotic behavior. A central limit theorem for S_n was then proved for $3 \leqslant m \leqslant 26$ in

[28, 31]; see also [30] for more details. Their approach is based on an inductive approximation argument.

By the method of moments, two authors of this paper re-proved in [8] the central limit theorem for S_n when $3 \leqslant m \leqslant 26$; the same approach was also used to establish the nonexistence of a limit law for S_n due to inherent oscillations. Moreover, the convergence rates to the normal distribution were characterized in [22] by a refined method of moments, which undergo further change of behaviors.

Then several different approaches were developed in the literature for a deeper understanding of the "phase change" at $m = 26$; these include martingale [6], renewal theory [25], urn models [23, 32], contraction method [13, 39], method of moments [22], statistical physics [9, 33], etc.

On the other hand, the KPL for general $m \geqslant 2$ was first studied by Mahmoud [29] and he proved

$$\mathbb{E}(K_n) = 2\phi n \log n + c_1 n + o(n),$$

for some explicitly computable constant c_1. The variance was computed in [30, §3.5] and satisfies ($H_m^{(2)} := \sum_{1\leqslant j\leqslant m} j^{-2}$)

$$\mathbb{V}(K_n) \sim C_K n^2, \quad \text{where} \quad C_K = 4\phi^2 \left(\tfrac{(m+1)H_m^{(2)}-2}{m-1} - \tfrac{\pi^2}{6} \right). \tag{8}$$

The corresponding limit law was characterized in [38] by the contraction method

$$\frac{K_n - \mathbb{E}(K_n)}{n} \xrightarrow{d} K, \tag{9}$$

where K is given by the recursive distributional equation (41); see also [4, 34] for a general framework.

For NPL N_n, Broutin and Holmgren [4] proved that

$$\mathbb{E}(N_n) = 2\phi^2 n \log n + c_2 n + o(n),$$

for some constant c_2. We will give an alternative proof of this result below with tools from [8, 14]. Our approach makes the computation of c_2 feasible, but we have not carried it out since the value is irrelevant for this work (no value was given in [4]).

It should be mentioned that there is a large literature on K_n when $m = 2$ because it is identical to the comparison cost used by quicksort. Many fine results were obtained; see, for example, the recent papers [3, 12, 17, 20, 37, 41] and the references therein for more information.

2.3 Covariance, correlation, dependence and phase changes

We state in this section our results for the covariance and correlation between the space requirement and the total path lengths (KPL and NPL). The proofs and the tools needed will be given in the next sections.

Unlike the space requirement S_n whose variance changes its asymptotic behavior for $m \geqslant 27$, the covariance $\operatorname{Cov}(S_n, K_n)$ changes its asymptotic behavior at $m = 14$.

Theorem 2.1. *The covariance between S_n and K_n satisfies*

$$\mathrm{Cov}(S_n, K_n) \sim \begin{cases} C_R n, & \text{if } 3 \leqslant m \leqslant 13; \\ F_2\left(\beta \log n\right) n^{\alpha}, & \text{if } m \geqslant 14; \end{cases}$$

where C_R is a suitable constant and $F_2(z)$ is a 2π-periodic function given in (23) below.

This result has the following consequence.

Corollary 2.2. *The correlation coefficient between S_n and K_n satisfies*

$$\rho(S_n, K_n) \begin{cases} \to 0, & \text{if } 3 \leqslant m \leqslant 26; \\ \sim \dfrac{F_2\left(\beta \log n\right)}{\sqrt{C_K F_1(\beta \log n)}}, & \text{if } m \geqslant 27, \end{cases}$$

where $C_K > 0$ is given in (8).

See Figure 1 for two different plots for the periodic functions when $m \geqslant 27$.

The same consideration extends easily to clarify the correlation between space requirement and NPL.

Theorem 2.3. *The covariance between S_n and N_n satisfies*

$$\mathrm{Cov}(S_n, N_n) \sim \begin{cases} 2\phi C_S n \log n, & \text{if } 3 \leqslant m \leqslant 13; \\ \phi F_2\left(\beta \log n\right) n^{\alpha}, & \text{if } m \geqslant 14, \end{cases}$$

where C_S is as in Section 2.2. Moreover, the variance of N_n satisfies

$$\mathbb{V}(N_n) \sim \phi^2 C_K n^2.$$

Notice the appearance of an extra $\log n$ factor when $3 \leqslant m \leqslant 13$, which reflects the additional random effect introduced by the toll function in (7). These estimates imply the following consequence.

Corollary 2.4. *The correlation coefficient $\rho(S_n, N_n)$ satisfies*

$$\rho(S_n, N_n) \begin{cases} \to 0, & \text{if } 3 \leqslant m \leqslant 26; \\ \sim \rho(S_n, K_n) \sim \dfrac{F_2\left(\beta \log n\right)}{\sqrt{C_K F_1(\beta \log n)}}, & \text{if } m \geqslant 27. \end{cases}$$

The last relation suggests considering the correlation between K_n and N_n.

Corollary 2.5. *The random variable K_n is asymptotically linearly correlated to N_n*

$$\rho(K_n, N_n) \to 1.$$

Indeed, we will show that

$$\|N_n - \phi K_n - (\mathbb{E}(N_n - \phi K_n))\|_2 = o(n)$$

which then by Slutsky's theorem implies that

$$\left(\frac{K_n - \mathbb{E}(K_n)}{n}, \frac{N_n - \mathbb{E}(N_n)}{n}\right) \xrightarrow{d} (K, \phi K);$$

see (9), Section 4.3 and 4.4.

These results will be proved by working out the asymptotics of the corresponding recurrence relations, which all have the same form

$$a_n = m \sum_{0 \leqslant j \leqslant n-m+1} \pi_{n,j} a_j + b_n, \qquad (n \geqslant m-1),$$

where

$$\pi_{n,j} = \frac{\binom{n-1-j}{m-2}}{\binom{n}{m-1}} \qquad (0 \leqslant j \leqslant n-m+1)$$

is a probability distribution, and $\{b_n\}$ is a given sequence (referred to as the toll-function). For that asymptotic purpose, our key tools will rely on the *asymptotic transfer techniques* (see [8, 14]), which provide a direct asymptotic translation from the asymptotic behaviors of b_n to those of a_n. The remaining analysis will then consist of simplifying some multiple Dirichlet's integrals.

Since Pearson's product-moment correlation coefficient ρ is known to be poor in measuring nonlinear dependence between two random variables, we go further by considering the joint limit laws for (S_n, K_n) and (S_n, N_n), which exhibits a change of behavior depending on whether $3 \leqslant m \leqslant 26$ (convergent case) or $m \geqslant 27$ (periodic case): they are asymptotically independent in the former case but dependent in the latter.

Theorem 2.6. *Assume $3 \leqslant m \leqslant 26$. Let $(X_n)_n \in \{(K_n)_n, (N_n)_n\}$ and $Q_n = (X_n, S_n)$ denote the vector of KPL or NPL and the space requirement used by a random m-ary search tree. Then the convergence in distribution holds:*

$$\mathrm{Cov}(Q_n)^{-1/2}(Q_n - \mathbb{E}[Q_n]) \xrightarrow{d} (X, \mathscr{N}), \tag{10}$$

where $\mathscr{N}$ has the standard normal distribution and the limit law $(X, \mathscr{N})$ is described in Lemma 4.2; moreover, X and $\mathscr{N}$ are independent.

Theorem 2.7. *Assume $m \geqslant 27$. Let $(X_n)_n \in \{(K_n)_n, (N_n)_n\}$ and*

$$Y_n := \left(\frac{X_n - \mathbb{E}[X_n]}{\iota_X n}, \frac{S_n - \phi n}{n^{\alpha-1}}\right)$$

with $\iota_X = 1$ for $(X_n)_n = (N_n)_n$ and $\iota_X = \phi^{-1}$ for $(X_n)_n = (K_n)_n$. Then we have

$$\ell_2(Y_n, (X, \Re(n^{i\beta}\Lambda))) \to 0,$$

where β is as in Section 2.2 and (X, Λ) is a random vector whose distribution is specified as the unique fixed point solution appearing in Lemma 4.1 for the choice $\gamma = (0, \theta)$ (θ being defined below in (25)).

See Section 4 for a more precise formulation. The proof is based on the *contraction method* (see [36]) where we use the above moment asymptotics as input and combine well-known estimates within the minimal L_2-metric for the convergent case (as in [40]), and those with estimates for the periodic case (as in [13]). Similar proof techniques related to periodic distributional behaviors are also applied in [25, Theorem 1.3(iii)] and [26, Theorem 6.10]. If one is only interested in the asymptotic (univariate) distribution of the NPL N_n (the case of the KPL being known before), there are more direct proofs which we also discuss in Sections 4.3 and 4.4.

Our study on the dependence of random variables on random m-ary search trees can be extended in at least two directions by the same methods used in this paper, namely, asymptotic transfer techniques and the contraction method.

- *Extension to more general linear and $n\log n$ shape measures*: That the asymptotic covariance undergoes a phase change after $m = 13$ and the asymptotic correlation undergoes a phase change after $m = 26$ is not restricted to the space requirement and KPL or NPL. Indeed, we can replace the space requirement by many other linear shape measures such as the number of leaves, the number of nodes of a specified type, the number of occurrences of a fixed pattern, etc. (see [8] for more examples), and KPL or NPL by other shape measures with mean of order $n\log n$ such as summing over the root-node or root-key distance for certain specified nodes or patterns and weighted path length.

- *Extension to other random trees of logarithmic height*: the same change of asymptotic behaviors from being independent to being dependent under a varying structural parameter also occurs in other classes of random log-trees; we content ourselves with the brief discussion of two classes of random trees: *fringe-balanced binary search trees* and *quadtrees*. The behaviors will be however very different for the classes of trees where the underlying distribution of the subtree sizes are dictated by a binomial distribution, which will be examined elsewhere; see a companion paper [18] for more information.

This paper is organized as follows. We prove in the next section our results for the covariances and the correlations. These results are then used to study the bivariate distributional asymptotics in Section 4 by the multivariate contraction method (see [36]). Finally, in Section 5, we discuss the dependence and phase changes in fringe-balanced binary search trees and in quadtrees, where for the former, we study the joint behavior of the size and total path length, while for the latter (since the size is a constant) we consider the joint behavior of the number of leaves and total path length. Also we include a brief discussion for extending the study and results to other shape parameters in Section 5.

3 Correlation between space requirement and path lengths

We prove in this section Theorems 2.1 and 2.3 for the covariances $\mathrm{Cov}(S_n, K_n)$ and $\mathrm{Cov}(S_n, N_n)$, respectively.

3.1 Preliminaries and recurrences

We collect here the notations to be used in the proofs. Let $m \geqslant 2$ be a fixed integer. For $n \geqslant m$, denote by $I^{(n)} = (I_1^{(n)}, \dots, I_m^{(n)})$ the vector of the number of keys inserted in the m

ordered subtrees of the root in a random m-ary search tree with n keys. When the dependence on n is obvious, we write simply $(I_1, \dots, I_m)$. Generate independently n uniform random variables $U_1, \dots, U_n$ on $[0, 1]$. Store the first $m-1$ elements $U_1, \dots, U_{m-1}$ in the root-node of the tree. Then they decompose the unit interval $[0, 1]$ into spacings of lengths $V_1, \dots, V_m$, where $V_j = U_{(j)} - U_{(j-1)}$ for $j = 1, \dots, m$ with $U_{(0)} := 0, U_{(m)} := 1$ and $U_{(j)}$ for $j = 1, \dots, m-1$ are the order statistics of $U_1, \dots, U_{m-1}$. The uniform permutation model implies, that, conditional on $U_1, \dots, U_{m-1}$, the vector $I^{(n)}$ has the multinomial distribution with success probabilities $V_1, \dots, V_m$, namely, we have

$$(I_1, \dots, I_m) \stackrel{\mathrm{d}}{=} M(n - m + 1; V_1, \dots, V_m).$$

In particular, we have the convergence

$$\frac{I_r}{n} \longrightarrow V_r, \tag{11}$$

for all $r = 1, \dots, m$, where the convergence is in L_p for all $1 \leqslant p < \infty$. Note that we also have (3) for all m-tuples $i_1, \dots, i_m \geqslant 0$ with $i_1 + \cdots + i_m = n - m + 1$ and all $n \geqslant m$.

For each of the subtrees, the randomness (uniformity) is preserved; more precisely, conditional on the number of keys inserted in a subtree, each subtree has the same distribution as a random m-ary search tree of that number of keys in the uniform model. Moreover, conditional on $(I_1, \dots, I_m)$, the subtrees are independent. This can be seen by switching back to the ranks $\{1, \dots, n\}$ of the input elements, and then by checking that a uniform random permutation yields independent permutations on the respective ranges. This recursive structure of the random m-ary search tree implies the recursive relations for S_n, K_n and N_n given in (5)–(7), where the summands appearing on the right-hand sides, namely, $S_j^{(1)}, \dots, S_j^{(m)}$ and $K_j^{(1)}, \dots, K_j^{(m)}$ and $N_j^{(1)}, \dots, N_j^{(m)}$ have the same distributions as S_j and K_j and N_j, respectively. Furthermore, the triples $\left(\left(S_j^{(r)}\right)_{0 \leqslant j \leqslant n-m+1}, \left(K_j^{(r)}\right)_{0 \leqslant j \leqslant n-m+1}, \left(N_j^{(r)}\right)_{0 \leqslant j \leqslant n-m+1} \right)$ are independent for $r = 1, \dots, m$ and independent of $(I_1, \dots, I_m)$. Finally, the recursive structure of the m-ary search tree implies recurrences satisfied by their joint distributions. In particular, the pair $Q_n := (N_n, S_n)$ satisfies the recurrence

$$(Q_n)^{\mathrm{t}} \stackrel{d}{=} \sum_{1 \leqslant r \leqslant m} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \left(Q_{I_r}^{(r)}\right)^{\mathrm{t}} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad (n \geqslant m), \tag{12}$$

where, as in (5)–(7), the $Q_j^{(r)}$'s are distributed as Q_j for all $1 \leqslant r \leqslant m$ and $0 \leqslant j \leqslant n-m+1$, and the $\left(Q_j^{(r)}\right)_{0 \leqslant j \leqslant n-m+1}$ are independent for $r = 1, \dots, m$ and independent of $(I_1, \dots, I_n)$. The recurrence satisfied by the pair $Z_n := (K_n, S_n)$ is

$$(Z_n)^{\mathrm{t}} \stackrel{d}{=} \sum_{1 \leqslant r \leqslant m} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \left(Z_{I_r}^{(r)}\right)^{\mathrm{t}} + \begin{pmatrix} n - m + 1 \\ 1 \end{pmatrix}, \qquad (n \geqslant m), \tag{13}$$

with conditions on independence and identical distributions similar to (12).

3.2 Asymptotic transfer and Dirichlet integrals

Starting from the distributional recurrences (5) and (6), we see that all centered and non-centered moments satisfy the same recurrence of the following type

$$a_n = m \sum_{0\leqslant j\leqslant n-m+1} \pi_{n,j} a_j + b_n, \qquad \pi_{n,j} = \frac{\binom{n-1-j}{m-2}}{\binom{n}{m-1}}, \tag{14}$$

for $n \geqslant m-1$, where $\{b_n\}_{n\geqslant m-1}$ is a given sequence. The asymptotics of a_n can be systematically characterized by that of b_n through the use of the following transfer techniques; see Proposition 7 in [8] and Theorem 2.4 in [14] for details.

Proposition 3.1. *Assume that a_n satisfies (14) with finite initial conditions $a_0, \dots, a_{m-2}$. Define $b_n := a_n$ for $0 \leqslant n \leqslant m-2$.*

(i) Assume $b_n = c(n+1) + t_n$, where $c \in \mathbb{C}$. Then the conditions

$$t_n = o(n) \quad \textit{and} \quad \left| \sum_{n\geqslant 1} t_n n^{-2} \right| < \infty$$

are both necessary and sufficient for

$$a_n = 2c\phi n H_n + c' n + o(n),$$

where

$$c' = 2\phi \sum_{j\geqslant 0} \frac{t_j}{(j+1)(j+2)} + \frac{c}{2} - 2c\phi + 2c(H_m^{(2)} - 1)\phi^2;$$

(ii) if $b_n \sim c n^v$, where $v > 1$, then

$$a_n \sim \frac{c}{1 - \frac{m!\Gamma(v+1)}{\Gamma(v+m)}} n^v.$$

In particular, when $c = 0$ in (i), then we see that a_n is asymptotically linear

$$\frac{a_n}{n} \sim 2\phi \sum_{j\geqslant 0} \frac{b_j}{(j+1)(j+2)} \quad \text{iff} \quad b_n = o(n) \quad \text{and} \quad \left| \sum_{n\geqslant 1} b_n n^{-2} \right| < \infty.$$

We will be dealing with Dirichlet integrals of the following type

$$I(u,v) := \int_{\substack{x_1+\cdots+x_m=1 \\ 0\leqslant x_1,\dots,x_m\leqslant 1}} \left(\sum_{1\leqslant l\leqslant m} x_l^{u-1} \right) \left(\sum_{1\leqslant r\leqslant m} x_r^{v-1} \right) \mathrm{d}\mathbf{x}, \qquad (\Re(u), \Re(v) > 0).$$

Here $\mathrm{d}\mathbf{x}$ is an abbreviation for $\mathrm{d}x_1 \cdots \mathrm{d}x_{m-1}$. Such integrals have a closed-form expression.

Lemma 3.2. *For $m \geqslant 2$ and $\Re(u), \Re(v) > 0$,*

$$I(u,v) = \frac{m\Gamma(u+v-1) + m(m-1)\Gamma(u)\Gamma(v)}{\Gamma(u+v+m-2)}. \tag{15}$$

Proof. First, the claim is easily proved for $m = 2$. Assume $m \geqslant 3$. Then, by symmetry,

$$\begin{aligned} I(u,v) &= \int_{\substack{x_1+\cdots+x_m=1 \\ 0\leqslant x_1,\ldots,x_m\leqslant 1}} \left(m x_1^{u+v-2} + m(m-1)x_1^{u-1}x_2^{v-1}\right) \mathrm{d}\mathbf{x} \\ &= \frac{m}{(m-2)!}\int_0^1 x_1^{u+v-2}(1-x_1)^{m-2}\mathrm{d}x_1 \\ &\qquad + \frac{m(m-1)}{(m-3)!}\int_0^1\int_0^{1-x_1} x_1^{u-1}x_2^{v-1}(1-x_1-x_2)^{m-3}\mathrm{d}x_2\mathrm{d}x_1 \\ &= \frac{m\Gamma(u+v-1)}{\Gamma(u+v+m-2)} + \frac{m(m-1)\Gamma(u)\Gamma(v)}{\Gamma(u+v+m-2)}, \end{aligned}$$

which leads to (15). ∎

The following two identities will be needed below.

$$\begin{aligned} &\int_{\substack{x_1+\cdots+x_m=1 \\ 0\leqslant x_1,\ldots,x_m\leqslant 1}} \left(\sum_{1\leqslant l\leqslant m} x_l^{u-1}\right)\left(\sum_{1\leqslant r\leqslant m} x_r\log x_r\right)\mathrm{d}\mathbf{x} \\ &\quad = \frac{\partial}{\partial v} I(u,v)\bigg|_{v=2} \\ &\quad = \frac{m\Gamma(u)}{\Gamma(m+u)}\left(u\psi(u+1) + (m-1)(1-\gamma) - (m+u-1)\psi(m+u)\right), \end{aligned} \tag{16}$$

where ψ is the digamma function and γ is Euler's constant. Similarly,

$$\begin{aligned} \int_{\substack{x_1+\cdots+x_m=1 \\ 0\leqslant x_1,\ldots,x_m\leqslant 1}} \left(\sum_{1\leqslant r\leqslant m} x_r\log x_r\right)^2 \mathrm{d}\mathbf{x} &= \frac{\partial^2}{\partial u\partial v} I(u,v)\bigg|_{u=v=2} \\ &= H_m^{(2)} + \frac{4}{\phi^2} - \frac{2}{m+1} - \frac{(m-1)\pi^2}{6(m+1)}. \end{aligned} \tag{17}$$

3.3 Correlation between the space requirement and KPL

We are now ready to prove Theorem 2.1.

Expected values of S_n and K_n. For convenience, let $\mu_n := \mathbb{E}(S_n)$ and $\kappa_n := \mathbb{E}(K_n)$. Then, by (5) and (6), for $n \geqslant m-1$

$$\begin{aligned} \mu_n &= m\sum_{0\leqslant j\leqslant n-m+1} \pi_{n,j}\mu_j + 1, \\ \kappa_n &= m\sum_{0\leqslant j\leqslant n-m+1} \pi_{n,j}\kappa_j + n - m + 1, \end{aligned}$$

with the initial conditions $\mu_0 = \kappa_n = 0$ for $0 \leqslant n \leqslant m-2$ and $\mu_n = 1$ for $1 \leqslant n \leqslant m-2$.

By applying Proposition 3.1(i), we obtain

$$\mu_n \sim \phi n, \quad \text{and} \quad \kappa_n = 2\phi n\log n + cn + o(n), \tag{18}$$

for some constant c whose value matters less. The latter approximation is sufficient for all our purposes, but the former is not and we need the following stronger expansion (see [8, 31, 30])

$$\mu_n = \phi(n+1) - \frac{1}{m-1} + \sum_{2\leqslant k\leqslant 3} \frac{A_k}{\Gamma(\lambda_k)}\, n^{\lambda_k - 1} + o(n^{\alpha-1}), \tag{19}$$

where $\lambda_2 = \alpha + i\beta$ and $\lambda_3 := \alpha - i\beta$ and

$$A_k = \frac{1}{\lambda_k(\lambda_k-1)\sum_{0\leqslant j\leqslant m-2}\frac{1}{j+\lambda_k}}.$$

Note that for $3 \leqslant m \leqslant 13$ the constant term $-\frac{1}{m-1}$ (together with ϕ) is the second-order term on the right-hand side of (19), while for larger m, it is absorbed in the o-term.

Variance and covariance. To compute the asymptotics of the covariance, we first derive the corresponding recurrences and then apply Proposition 3.1 of asymptotic transfer.

First, let $\bar{S}_n = S_n - \mu_n$ and $\bar{K}_n = K_n - \kappa_n$. We consider the moment-generating function

$$\bar{P}_n(u,v) := \mathbb{E}\left(e^{\bar{S}_n u + \bar{K}_n v}\right).$$

Then, using (5) and (6), we obtain for $n \geqslant m-1$

$$\bar{P}_n(u,v) = \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} P_{j_1}(u,v)\cdots P_{j_m}(u,v) e^{\Delta_{\mathbf{j}} u + \nabla_{\mathbf{j}} v} \tag{20}$$

with the initial conditions $\bar{P}_n(u,v) = 1$ for $0 \leqslant n \leqslant m-2$. Here, $\mathbf{j} = (j_1,\dots,j_m)$ is a vector with $j_1,\dots,j_m \geqslant 0$ and $j_1+\cdots+j_m = n-m+1$ (we use this notation throughout),

$$\Delta_{\mathbf{j}} = 1 - \mu_n + \sum_{1\leqslant l\leqslant m} \mu_{j_l}, \quad \text{and} \quad \nabla_{\mathbf{j}} = n-m+1-\kappa_n + \sum_{1\leqslant l\leqslant m} \kappa_{j_l}. \tag{21}$$

Define

$$V_n^{[S]} = \mathbb{V}(S_n), \qquad V_n^{[SK]} = \mathrm{Cov}(S_n, K_n), \qquad V_n^{[K]} = \mathbb{V}(K_n).$$

Then, by taking derivatives in (20), we obtain

$$V_n^{[X]} = m \sum_{0\leqslant j\leqslant n-m+1} \pi_{n,j} V_j^{[X]} + b_n^{[X]}, \quad (X \in \{S, SK, K\}),$$

where

$$b_n^{[S]} = \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \Delta_{\mathbf{j}}^2, \quad b_n^{[SK]} = \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \Delta_{\mathbf{j}}\nabla_{\mathbf{j}}, \quad \text{and} \quad b_n^{[K]} = \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \nabla_{\mathbf{j}}^2.$$

We first derive uniform asymptotic approximations for $\Delta_{\mathbf{j}}$ and $\nabla_{\mathbf{j}}$.

Lemma 3.3. *Uniformly in* $\mathbf{j}$,

$$\Delta_{\mathbf{j}} = \sum_{2\leqslant k\leqslant 3} \frac{A_k}{\Gamma(\lambda_k)}\, n^{\lambda_k-1}\left(-1 + \sum_{1\leqslant r\leqslant m}\left(\frac{j_r}{n}\right)^{\lambda_k - 1}\right) + o(n^{\alpha-1}),$$

and

$$\nabla_{\mathbf{j}} = n\left(1 + 2\phi \sum_{1\leqslant r\leqslant m} \frac{j_r}{n}\log\frac{j_r}{n}\right) + o(n).$$

Proof. This follows from substituting the asymptotic approximations (18) and (19) into (21), and standard manipulations. ∎

Asymptotics of $V_n^{[S]}$. Although the asymptotic behaviors of the variance of S_n have been computed before, we re-derive them here by a different approach, which is easily amended for the calculation of other variances and covariances.

Consider first $3 \leqslant m \leqslant 26$. Then $\alpha < 3/2$. Moreover, from Lemma 3.3,

$$b_n^{[S]} = O(n^{2\alpha-2}) = O(n^{1-\varepsilon}),$$

for some $0 < \varepsilon < 0.00171$. Consequently, by applying Proposition 3.1(i),

$$V_n^{[S]} \sim C_S n,$$

for some constant C_S; see [8] for a more explicit expression and the proof that $C_S > 0$.

On other hand, if $m \geqslant 27$, since $\alpha > 3/2$, we then have, by Lemmas 3.2 and 3.3,

$$\begin{aligned} b_n^{[S]} &\sim \sum_{2\leqslant k_1,k_2\leqslant 3} \frac{(m-1)!A_{k_1}A_{k_2}n^{\lambda_{k_1}+\lambda_{k_2}-2}}{\Gamma(\lambda_{k_1})\Gamma(\lambda_{k_2})} \\ &\quad \times \int_{\substack{x_1+\cdots+x_m=1\\ 0\leqslant x_1,\ldots,x_m\leqslant 1}} \left(-1+\sum_{1\leqslant l\leqslant m} x_l^{\lambda_{k_1}-1}\right)\left(-1+\sum_{1\leqslant r\leqslant m} x_r^{\lambda_{k_2}-1}\right) \mathrm{d}\mathbf{x} \\ &\sim \sum_{2\leqslant k_1,k_2\leqslant 3} \frac{A_{k_1}A_{k_2}n^{\lambda_{k_1}+\lambda_{k_2}-2}}{\Gamma(\lambda_{k_1})\Gamma(\lambda_{k_2})}\left(1-\frac{m!\Gamma(\lambda_{k_1})}{\Gamma(\lambda_{k_1}+m-1)}-\frac{m!\Gamma(\lambda_{k_2})}{\Gamma(\lambda_{k_2}+m-1)}\right. \\ &\qquad \left. +\frac{m!\Gamma(\lambda_{k_1}+\lambda_{k_2}-1)}{\Gamma(\lambda_{k_1}+\lambda_{k_2}+m-2)}+\frac{m!(m-1)\Gamma(\lambda_{k_1})\Gamma(\lambda_{k_2})}{\Gamma(\lambda_{k_1}+\lambda_{k_2}+m-2)}\right). \end{aligned}$$

Note that

$$\frac{m!\Gamma(\lambda_{k_j})}{\Gamma(\lambda_{k_j}+m-1)} = 1, \qquad (2 \leqslant j \leqslant 3).$$

Applying Proposition 3.1(ii) term by term then gives

$$\begin{aligned} V_n^{[S]} &\sim \sum_{2\leqslant k_1,k_2\leqslant 3} \frac{A_{k_1}A_{k_2}n^{\lambda_{k_1}+\lambda_{k_2}-2}}{\Gamma(\lambda_{k_1})\Gamma(\lambda_{k_2})}\left(-1+\frac{m!(m-1)\Gamma(\lambda_{k_1})\Gamma(\lambda_{k_2})}{\Gamma(\lambda_{k_1}+\lambda_{k_2}+m-2)-m!\Gamma(\lambda_{k_1}+\lambda_{k_2}-1)}\right) \\ &=: F_1(\beta\log n)n^{2\alpha-2}, \end{aligned}$$

where

$$\begin{aligned} F_1(z) &:= 2\frac{|A_2|^2}{|\Gamma(\lambda_2)|^2}\left(-1+\frac{m!(m-1)|\Gamma(\lambda_2)|^2}{\Gamma(2\alpha+m-2)-m!\Gamma(2\alpha-1)}\right) \\ &\quad +2\Re\left(\frac{A_2^2e^{2iz}}{\Gamma(\lambda_2)^2}\left(-1+\frac{m!(m-1)\Gamma(\lambda_2)^2}{\Gamma(2\lambda_2+m-2)-m!\Gamma(2\lambda_2-1)}\right)\right). \end{aligned} \tag{22}$$

Asymptotics of $V_n^{[SK]}$. We now turn to $V_n^{[SK]}$. If $3 \leqslant m \leqslant 13$, then, by Lemma 3.3,

$$b_n^{[SK]} = O\left(n^{\alpha}\right),$$

where $\alpha < 1$. Consequently, by Proposition 3.1(i),

$$V_n^{[SK]} \sim C_R n,$$

for some constant C_R. For the remaining range where $m \geqslant 14$, we have $\alpha > 1$, and, by Lemma 3.3 and (16),

$$b_n^{[SK]} \sim \sum_{2\leqslant k\leqslant 3} \frac{(m-1)!A_k n^{\lambda_k}}{\Gamma(\lambda_k)} \int_{\substack{x_1+\cdots+x_m=1\\ 0\leqslant x_1,\ldots,x_m\leqslant 1}} \left(-1+\sum_{1\leqslant l\leqslant m} x_l^{\lambda_k-1}\right)\left(1+2\phi\sum_{1\leqslant r\leqslant m} x_r\log x_r\right)\mathrm{d}\mathbf{x}$$
$$\sim \sum_{2\leqslant k\leqslant 3} \frac{A_k n^{\lambda_k}}{\Gamma(\lambda_k)}\left(1-2\phi\frac{m!\Gamma(\lambda_k+1)}{\Gamma(\lambda_k+m)}\big\{m\psi(\lambda_k+m)-\psi(\lambda_k+1)-(m-1)(1-\gamma)\big\}\right).$$

Now, we apply Proposition 3.1(ii) and again after some straightforward simplifications

$$V_n^{[SK]} \sim F_2\left(\beta\log n\right) n^{\alpha},$$

where

$$F_2(z) := 2\phi\Re\left(\frac{(\lambda_2+m-1)A_2e^{iz}}{(m-1)\Gamma(\lambda_2)}\left(\frac{1}{2\phi}-\frac{\lambda_2}{\lambda_2+m-1}\big\{m\psi(\lambda_2+m)-\psi(\lambda_2+1)\right.\right.$$
$$\left.\left.-(m-1)(1-\gamma)\big\}\right)\right). \tag{23}$$

Asymptotics of $V_n^{[K]}$. In a similar manner, we obtain, by Lemma 3.3,

$$b_n^{[K]} \sim (m-1)!n^2 \int_{\substack{x_1+\cdots+x_m=1\\ 0\leqslant x_1,\ldots,x_m\leqslant 1}} \left(1+2\phi\sum_{1\leqslant l\leqslant m} x_l\log x_l\right)^2 \mathrm{d}\mathbf{x}$$
$$\sim 4\phi^2 n^2\left(H_m^{(2)}-\frac{2}{m+1}-\frac{\pi^2(m-1)}{6(m+1)}\right),$$

where the last line follows from applying (15), (16) and (17). Applying again Proposition 3.1(ii) gives

$$V_n^{[K]} \sim C_K n^2,$$

which completes the proof of Theorem 2.1. ∎

3.4 Correlation between space requirement and NPL

The calculations in this case are similar to those for $\rho(S_n, K_n)$, so we only sketch the major steps needed. Briefly, most asymptotic estimates differ either by a factor of the occupancy constant ϕ or its powers. The only exception is the additional factor $\log n$ appearing in the covariance $\mathrm{Cov}(S_n, N_n)$ (see (2.3)).

Let $\nu_n = \mathbb{E}(N_n)$. Then

$$\nu_n = m\sum_{0\leqslant j\leqslant n-m+1} \pi_{n,j}\nu_j + \mu_n - 1.$$

Consequently, by the asymptotic estimate (19) and by applying Proposition 3.1(i), we obtain

$$\nu_n = 2\phi^2 n\log n + c_2 n + o(n), \tag{24}$$

for some constant c_2.

Let $\bar{N}_n = N_n - \nu_n$. Then the moment-generating function $\bar{P}_n(u,v) := \mathbb{E}\big(e^{\bar{S}_n u + \bar{N}_n v}\big)$ satisfies for $n \geqslant m-1$

$$\bar{P}_n(u,v) = \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} P_{j_1}(u+v,v)\cdots P_{j_m}(u+v,v) e^{\Delta_{\mathbf{j}} u + \delta_{\mathbf{j}} v},$$

with the initial conditions $\bar{P}_n(u,v) = 1$ for $0 \leqslant n \leqslant m-2$ and

$$\delta_{\mathbf{j}} := -\nu_n + \sum_{1 \leqslant l \leqslant m} (\nu_{j_l} + \mu_{j_l}).$$

Now define

$$V_n^{[SN]} := \mathrm{Cov}(S_n, N_n) \qquad \text{and} \qquad V_n^{[N]} := \mathbb{V}(N_n).$$

Then

$$V_n^{[X]} = m \sum_{0 \leqslant l \leqslant n-m+1} \pi_{n,j} V_j^{[X]} + b_n^{[X]}, \qquad (X \in \{SN, N\}),$$

where

$$\begin{aligned} b_n^{[SN]} &= \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(V_j^{[S]} + \Delta_{\mathbf{j}} \delta_{\mathbf{j}} \right) \\ &= V_n^{[S]} + \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(\Delta_{\mathbf{j}} \delta_{\mathbf{j}} - \Delta_{\mathbf{j}}^2 \right) \\ b_n^{[N]} &= \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(V_j^{[S]} + 2V_j^{[SN]} + \delta_{\mathbf{j}}^2 \right) \\ &= V_n^{[S]} + 2V_n^{[SN]} + \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(\delta_{\mathbf{j}}^2 - 2\Delta_{\mathbf{j}} \delta_{\mathbf{j}} + \Delta_{\mathbf{j}}^2 \right). \end{aligned}$$

As in the case of KPL, the following uniform estimate is crucial in our analysis.

Lemma 3.4. *Uniformly in* $\mathbf{j}$*,*

$$\delta_{\mathbf{j}} = \phi n \left(1 + 2\phi \sum_{1 \leqslant l \leqslant m} \frac{j_l}{n} \log \frac{j_l}{n} \right) + o(n).$$

Proof. By the definition of $\delta_{\mathbf{j}}$ and the estimates (19) and (24). ∎

Note that the expansion differs from that for $\nabla_{\mathbf{j}}$ in Lemma 3.3 by an additional factor ϕ.

If $3 \leqslant m \leqslant 13$, then, by Lemmas 3.3 and 3.4,

$$b_n^{[SN]} = C_S n + O\left(n^{1-\varepsilon}\right),$$

for a sufficiently small $\varepsilon > 0$. Thus, by Proposition 3.1 (i),

$$V_n^{[SN]} \sim \frac{C_S n \log n}{H_m - 1}.$$

Assume now $m \geqslant 14$. Then, again from Lemma 3.3 and Lemma 3.4 together with the known asymptotics of $V_n^{[S]}$, we see that

$$b_n^{[SN]} \sim \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \Delta_{\mathbf{j}} \delta_{\mathbf{j}} \sim \frac{\phi}{\binom{n}{m-1}} \sum_{\mathbf{j}} \Delta_{\mathbf{j}} \nabla_{\mathbf{j}} \sim \phi b_n^{[SK]}.$$

Thus we deduce, as in the proof for $V_n^{[SK]}$,

$$V_n^{[SN]} \sim \phi V_n^{[SK]} \sim \phi F_2(\beta \log n) n^{\alpha}.$$

Similarly, we have

$$b_n^{[N]} \sim \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \delta_{\mathbf{j}}^2 \sim \frac{\phi^2}{\binom{n}{m-1}} \sum_{\mathbf{j}} \nabla_{\mathbf{j}}^2 \sim \phi^2 b_n^{[K]}.$$

Consequently,

$$V_n^{[N]} \sim \phi^2 V_n^{[K]} \sim \phi^2 C_K n^2.$$

This completes the proof of Theorem 2.3. ∎

4 Bivariate distributional asymptotics for space requirement and path lengths

In this section, we identify the asymptotic joint distributional behaviors of the pairs (N_n, S_n) and (K_n, S_n). Although the sequences (N_n) and (K_n) converge after normalization for all $m \geqslant 3$ with limit distributions depending on m, we split the analysis into two cases depending on $3 \leqslant m \leqslant 26$ or $m > 26$ due to the phase change in the limit behavior of S_n. We discuss the pair (N_n, S_n) in detail in Sections 4.1 and 4.2. (the corresponding analysis for (K_n, S_n) is similar and we will not give details). Moreover, in Section 4.3, we will show that the univariate limit random variables of the normalized sequences (N_n) and (K_n) do have the same distribution. We introduce the following notation

$$\mu(n) := \mu_n = \mathbb{E}[S_n] = \phi(n+1) + \Re(\theta n^{\lambda_2 - 1}) + o(1 \vee n^{\alpha - 1}), \tag{25}$$

where $\theta := 2A_2/\Gamma(\lambda_2)$; see (19). Similarly, write $\kappa(n) = \kappa_n = \mathbb{E}(K_n)$ and $\nu(n) = \nu_n = \mathbb{E}(N_n)$.

4.1 Node path length and space requirement. I. $\mathbf{m \geqslant 27}$

We give in this section the precise formulation of the periodic case $m \geqslant 27$ of Theorem 2.7.

Normalization. We first normalize the vector $Q_n = (N_n, S_n)$ as follows. Let $Y_0 := 0$ and

$$Y_n := \left(\frac{N_n - \mathbb{E}[N_n]}{n}, \frac{S_n - \phi n}{n^{\alpha - 1}} \right), \qquad (n \geqslant 1).$$

Then the recurrence (12) implies for $n \geqslant m-1$

$$(Y_n)^{\mathrm{t}} \stackrel{d}{=} \sum_{1\leqslant r\leqslant m} A_r^{(n)} \left(Y^{(r)}_{I_r^{(n)}}\right)^{\mathrm{t}} + b^{(n)}, \tag{26}$$

where

$$A_r^{(n)} := \begin{bmatrix} \dfrac{I_r^{(n)}}{n} & \dfrac{\left(I_r^{(n)}\right)^{\alpha-1}}{n} \\ 0 & \left(\dfrac{I_r^{(n)}}{n}\right)^{\alpha-1} \end{bmatrix}, \quad b^{(n)} := \begin{pmatrix} \dfrac{1}{n}\left(\displaystyle\sum_{1\leqslant r\leqslant m} \left(\nu\left(I_r^{(n)}\right) + \phi I_r^{(n)}\right) - \nu(n)\right) \\ -\phi\dfrac{m-1}{n^{\alpha-1}} \end{pmatrix},$$

with assumptions on independence and on identical distributions as in Section 3.1. The expansion (24) implies

$$\frac{1}{n}\left(\sum_{1\leqslant r\leqslant m} \left(\nu\left(I_r^{(n)}\right) + \phi I_r^{(n)}\right) - \nu(n)\right) = \phi + 2\phi^2 \sum_{1\leqslant r\leqslant m} \frac{I_r^{(n)}}{n} \log \frac{I_r^{(n)}}{n} + o(1).$$

Moreover, by (11), we obtain the L_2-convergence

$$\frac{I^{(n)}}{n} \xrightarrow{L_2} (V_1, \ldots, V_m) =: V. \tag{27}$$

This implies the L_2-convergences

$$\frac{1}{n}\left(\sum_{1\leqslant r\leqslant m} \left(\nu\left(I_r^{(n)}\right) + \phi I_r^{(n)}\right) - \nu(n)\right) \to \phi + 2\phi^2 \sum_{1\leqslant r\leqslant m} V_r \log V_r =: b_N, \tag{28}$$

and

$$b^{(n)} \to \begin{pmatrix} b_N \\ 0 \end{pmatrix}, \qquad A_r^{(n)} \to \begin{bmatrix} V_r & 0 \\ 0 & V_r^{\alpha-1} \end{bmatrix}. \tag{29}$$

For our limit result for $m \geqslant 27$, we first define a distribution which governs the asymptotics.

The limiting map. To describe the asymptotic behavior of Q_n, we use the following probability distribution on the space $\mathbb{R}\times\mathbb{C}$. Let $\mathcal{M}^{\mathbb{R}\times\mathbb{C}}$ denote the space of all distributions $\mathcal{L}(Z,W)$ on $\mathbb{R}\times\mathbb{C}$ and $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}$ the subspace of distributions with finite second moment, i.e., $\|(Z,W)\|_2 := (\mathbb{E}[Z^2] + \mathbb{E}[|W|^2])^{1/2} < \infty$. For $\gamma = (\gamma_1, \gamma_2) \in \mathbb{R}\times\mathbb{C}$, let

$$\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}(\gamma) := \left\{\mathcal{L}(Z,W) \in \mathcal{M}_2^{\mathbb{R}\times\mathbb{C}} \,\middle|\, \mathbb{E}[Z] = \gamma_1, \mathbb{E}[W] = \gamma_2\right\}.$$

We define the following map T_N on $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}$:

$$\begin{aligned} T_N : \mathcal{M}^{\mathbb{R}\times\mathbb{C}} &\to \mathcal{M}^{\mathbb{R}\times\mathbb{C}} \\ \mathcal{L}(Z,W) &\mapsto \mathcal{L}\left(\sum_{1\leqslant r\leqslant m} \begin{bmatrix} V_r & 0 \\ 0 & V_r^{\lambda_2-1} \end{bmatrix} \begin{pmatrix} Z^{(r)} \\ W^{(r)} \end{pmatrix} + \begin{pmatrix} b_N \\ 0 \end{pmatrix}\right), \end{aligned} \tag{30}$$

where $(Z^{(1)}, W^{(1)}), \dots, (Z^{(m)}, W^{(m)})$, V are independent, $(Z^{(r)}, W^{(r)})$ is distributed as (Z, W) for all $r = 1, \dots, m$ and b_N is defined in (28). The $\|\cdot\|_2$-norm induces the minimal L_2-metric ℓ_2 by

$$\ell_2(\mu, \nu) := \inf\{\|X - Y\|_2 \,:\, \mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu\}, \qquad (\mu, \nu \in \mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}).$$

Given random variables X, Y, write for simplicity $\ell_2(X, Y) = \ell_2(\mathcal{L}(X), \mathcal{L}(Y))$. For any distributions $\mu, \nu \in \mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}$, there exist optimal ℓ_2-couplings, i.e. random vectors Υ_1, Υ_2 in $\mathbb{R} \times \mathbb{C}$ with $\ell_2(\mu, \nu) = \|\Upsilon_1 - \Upsilon_2\|_2$.

Lemma 4.1. *Assume $m \geqslant 27$. For any $\gamma \in \mathbb{R} \times \mathbb{C}$, the restriction of the map T_N defined in (30) to $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}(\gamma)$ is a (strict) contraction with respect to ℓ_2, and has a unique fixed point in $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}(\gamma)$.*

Proof. Let $\gamma \in \mathbb{R} \times \mathbb{C}$ be arbitrary. For $\mu \in \mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}(\gamma)$, let Υ be a random variable with distribution $T(\mu)$. First, note that $\|\Upsilon\|_2 < \infty$ by independence and $\|b_N\|_2 < \infty$ (we even have $\|b_N\|_\infty < \infty$). To see that $\mathbb{E}[\Upsilon] = \gamma$, note that $\mathbb{E}[b_N] = 0$ and $\sum_{1\leqslant r\leqslant m} V_r = 1$ almost surely. Hence, we only need to show that $\mathbb{E}[V_1^{\lambda_2 - 1}] = 1/m$. Since V_1 has density $x \mapsto (m-1)(1-x)^{m-2}$ for $x \in [0, 1]$, we see that

$$\mathbb{E}\left[V_1^{\lambda_2-1}\right] = \int_0^1 (m-1)(1-x)^{m-2} x^{\lambda_2 - 1}\, \mathrm{d}x = (m-1)\frac{\Gamma(m-1)\Gamma(\lambda_2)}{\Gamma(m+\lambda_2-1)} = \frac{1}{m},$$

because $\Gamma(m + \lambda_2 - 1)/\Gamma(\lambda_2) = m!$. This implies that $\mathbb{E}[\Upsilon] = \gamma$, and thus $T(\mu) \in \mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}(\gamma)$. This in turn implies that the restriction of T to $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}(\gamma)$ maps into $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}(\gamma)$.

That the restriction of T to $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}(\gamma)$ is a contraction with respect to ℓ_2 follows from a standard calculation, e.g., with a slight modification as in [36, Lemma 3.1]. ∎

Proof of Theorem 2.7: NPL. Denote by $\mathcal{L}(X, \Lambda)$ the unique fixed point of the restriction of T_N to $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}((0, \theta))$, with θ defined in (25). By Lemma 4.1, the distribution $\mathcal{L}(X, \Lambda)$ as in the statement of the Theorem is well-defined. The fixed point property of (X, Λ) implies that

$$\begin{pmatrix} X \\ \Re(n^{i\beta}\Lambda) \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} \sum_{1\leqslant r\leqslant m} V_r X^{(r)} + b_N \\ \sum_{1\leqslant r\leqslant m} \Re(n^{i\beta} V_r^{\lambda_2 - 1} \Lambda^{(r)}) \end{pmatrix}, \tag{31}$$

where $(V_1, \dots, V_m), (X^{(1)}, \Lambda^{(1)}), \dots, (X^{(m)}, \Lambda^{(m)})$ are independent, and $(X^{(r)}, \Lambda^{(r)})$ are identically distributed as (X, Λ).

Define now three matrices

$$\widetilde{A}_r^{(n)} := \begin{bmatrix} \frac{I_r^{(n)}}{n} & 0 \\ 0 & \left(\frac{I_r^{(n)}}{n}\right)^{\alpha-1} \end{bmatrix}, \; B_r^{(n)} := \begin{bmatrix} V_r & 0 \\ 0 & n^{i\beta} V_r^{\lambda_2 - 1} \end{bmatrix}, \; C_r^{(n)} := \begin{bmatrix} \frac{I_r^{(n)}}{n} & 0 \\ 0 & \frac{\left(I_r^{(n)}\right)^{\lambda_2 - 1}}{n^{\alpha-1}} \end{bmatrix},$$

and write

$$\Delta(n) := \ell_2(Y_n, (X, \Re(n^{i\beta}\Lambda))).$$

To bound $\Delta(n)$, we use the following coupling between the $Y_j^{(r)}$'s appearing in the recurrence (26) and the quantities appearing on the right-hand side of (31). Note that for any pair of distributions on $\mathbb{R}^2$, there always exists an optimal ℓ_2-coupling. We first fix the random vectors $(X^{(1)}, \Lambda^{(1)}), \dots, (X^{(m)}, \Lambda^{(m)})$. Then, for each $j \geqslant 1$ and $r = 1, \dots, m$, we choose $Y_j^{(r)}$ as an optimal ℓ_2-coupling to $(X^{(r)}, \Re(j^{i\beta}\Lambda^{(r)}))$ on $\mathbb{R}^2$. This can be done such that the sequences

$$\left(Y_j^{(1)}, (X^{(1)}, \Re(j^{i\beta}\Lambda^{(1)}))\right)_{j\geqslant 1}, \dots, \left(Y_j^{(m)}, (X^{(m)}, \Re(j^{i\beta}\Lambda^{(m)}))\right)_{j\geqslant 1}$$

are independent and independent of $(I^{(n)}, V_1, \dots, V_m)$. Note that these couplings and independence assumptions do not violate equations (26) and (31). Hence, we obtain

$$\Delta(n) \leqslant \left\| \sum_{1\leqslant r\leqslant m} A_r^{(n)} \left(Y_{I_r^{(n)}}^{(r)}\right)^t + b^{(n)} - \Re\left(\sum_{1\leqslant r\leqslant m} B_r^{(n)} \begin{pmatrix} X^{(r)} \\ \Lambda^{(r)} \end{pmatrix} + \begin{pmatrix} b \\ 0 \end{pmatrix}\right)\right\|_2.$$

Using the triangle inequality and writing the components as $Y_n = (Y_{n,1}, Y_{n,2})$, we obtain

$$\begin{aligned}\Delta(n) \leqslant &\left\| \sum_{1\leqslant r\leqslant m} \left(\widetilde{A}_r^{(n)} \left(Y_{I_r^{(n)}}^{(r)}\right)^t - \Re\left(C_r^{(n)} \begin{pmatrix} X^{(r)} \\ \Lambda^{(r)} \end{pmatrix}\right)\right)\right\|_2 \\ &+ \left\| \sum_{1\leqslant r\leqslant m} \Re\left(C_r^{(n)} \begin{pmatrix} X^{(r)} \\ \Lambda^{(r)} \end{pmatrix}\right) - \Re\left(B_r^{(n)} \begin{pmatrix} X^{(r)} \\ \Lambda^{(r)} \end{pmatrix}\right)\right\|_2 \\ &+ \sum_{1\leqslant r\leqslant m} \left\| \frac{(I_r^{(n)})^{\alpha-1}}{n} Y_{I_r^{(n)},2}^{(r)} \right\|_2 + \left\| b^{(n)} - \begin{pmatrix} b \\ 0 \end{pmatrix}\right\|_2.\end{aligned}$$

The second and the fourth summand on the right-hand side tend to zero as $n \to \infty$ by (27) and (29). For the third summand, note that the asymptotic behavior of the normalized size $Y_{n,2}$ of m-ary search trees is covered by Theorem 1, eq. (2) in [8]. In particular, from that theorem we obtain $\sup_{n\geqslant 1} \|Y_{n,2}\|_2 < \infty$. Taking into account the prefactor $(I_r^{(n)})^{\alpha-1}/n$ and conditioning on $I_r^{(n)}$, we find that the third summand also tends to zero.

To bound the first summand in the latter display, we write, for $r = 1, \dots, m$ and $n \geqslant m-1$,

$$W_r^{(n)} := \widetilde{A}_r^{(n)} \left(Y_{I_r^{(n)}}^{(r)}\right)^t - \Re\left(C_r^{(n)} \begin{pmatrix} X^{(r)} \\ \Lambda^{(r)} \end{pmatrix}\right)$$

and denote the components of $W_r^{(n)}$ by $W_r^{(n)} = (W_{r,1}^{(n)}, W_{r,2}^{(n)})$. For $r = 1, \dots, m$, we have

$$\mathbb{E}\left\| \sum_{1\leqslant r\leqslant m} W_r^{(n)} \right\|^2 = \mathbb{E}\left[\sum_{1\leqslant r\leqslant m} \left\{(W_{r,1}^{(n)})^2 + (W_{r,2}^{(n)})^2\right\} + \sum_{r\neq s} \left\{W_{r,1}^{(n)} W_{s,1}^{(n)} + W_{r,2}^{(n)} W_{s,2}^{(n)}\right\}\right]. \tag{32}$$

We bound the three types of terms individually. First, for the dominant term

$$\begin{aligned}&\mathbb{E}\left[(W_{r,1}^{(n)})^2 + (W_{r,2}^{(n)})^2\right] \\ &= \mathbb{E}\left[\left(\frac{I_r^{(n)}}{n}\right)^2 \left(Y_{I_r^{(n)},1}^{(r)} - X^{(r)}\right)^2 + \left(\frac{I_r^{(n)}}{n}\right)^{2(\alpha-1)} \left(Y_{I_r^{(n)},2}^{(r)} - \Re\left((I_r^{(n)})^{i\beta}\Lambda^{(r)}\right)\right)^2\right]\end{aligned}$$

$$\leqslant \mathbb{E}\left[\left(\frac{I_r^{(n)}}{n}\right)^{2(\alpha-1)}\left(\left(Y^{(r)}_{I_r^{(n)},1}-X^{(r)}\right)^2+\left(Y^{(r)}_{I_r^{(n)},2}-\Re\left((I_r^{(n)})^{i\beta}\Lambda^{(r)}\right)\right)^2\right)\right]$$

where we used the inequality $(I_r^{(n)}/n)^2 \leqslant (I_r^{(n)}/n)^{2(\alpha-1)}$. Conditioning on $I_r^{(n)}$ and using that $Y_j^{(r)}$ and $(X^{(r)}, \Re(j^{i\beta}\Lambda^{(r)}))$ are optimal couplings, we obtain

$$\mathbb{E}\left[(W_{r,1}^{(n)})^2+(W_{r,2}^{(n)})^2\right] \leqslant \mathbb{E}\left[\left(\frac{I_r^{(n)}}{n}\right)^{2(\alpha-1)}\Delta^2(I_r^{(n)})\right].$$

For the cross-product terms in (32), assume $1 \leqslant r, s \leqslant m$ with $r \neq s$. Note that, by independence, we have $\mathbb{E}[W_{r,1}^{(n)}W_{s,1}^{(n)}] = 0$ conditioning on $I_r^{(n)}$ and $I_s^{(n)}$. From the expansion (25), we obtain

$$\mathbb{E}\left[Y_n\right] = \begin{pmatrix} 0 \\ \Re(\theta n^{i\beta}) + R(n) \end{pmatrix},$$

with a remainder $R(n) = o(1)$. By independence and $\mathbb{E}[\Lambda] = \theta$, we obtain $\mathbb{E}[W_{r,2}^{(n)}] = \mathbb{E}[(I_r^{(n)}/n)^{\alpha-1}R(I_r^{(n)})]$, and

$$\mathbb{E}[W_{r,2}^{(n)}W_{s,2}^{(n)}] = \mathbb{E}\left[\left(\frac{I_r^{(n)}}{n}\cdot\frac{I_s^{(n)}}{n}\right)^{\alpha-1}R(I_r^{(n)})R(I_s^{(n)})\right]$$

which tends to 0 by the dominated convergence theorem as $I_r^{(n)}, I_s^{(n)} \to 0$ in probability.

Hence, collecting all estimates, we obtain

$$\Delta(n) \leqslant \left(\mathbb{E}\left[\sum_{1\leqslant r\leqslant m}\left(\frac{I_r^{(n)}}{n}\right)^{2(\alpha-1)}\Delta^2(I_r^{(n)})\right]+o(1)\right)^{1/2}+o(1). \tag{33}$$

Now $\Delta(n) \to 0$ follows from a standard argument since we have

$$\lim_{n\to\infty}\sum_{1\leqslant r\leqslant m}\mathbb{E}\left[\left(\frac{I_r^{(n)}}{n}\right)^{2(\alpha-1)}\right] = \sum_{1\leqslant r\leqslant m}\mathbb{E}\left[V_r^{2(\alpha-1)}\right] = m^2B(m, 2\alpha-1) < 1;$$

(cf. the proof of Theorem 4.1 in [36]). This proves Theorem 2.7 for NPL.

4.2 Node path length and space requirement. II. $3 \leqslant m \leqslant 26$

We begin with the recurrence (12), and recall that, for $3 \leqslant m \leqslant 26$,

$$\mathbb{V}(S_n) \sim C_S n, \quad \mathbb{V}(N_n) \sim C_N n^2 \quad \text{with} \quad C_N = \phi^2 C_K;$$

see (18) and (24). There exists an $n_1 \geqslant 1$, such that for all $n \geqslant n_1$, the matrix $\mathrm{Cov}(Q_n)$ is positive definite. We normalize it by $\widetilde{Q}_n := Q_n$ for $0 \leqslant n < n_1$ and by

$$\left(\widetilde{Q}_n\right)^{\mathrm{t}} := \begin{bmatrix} (\sqrt{C_N}n)^{-1} & 0 \\ 0 & (C_S n)^{-1/2} \end{bmatrix}(Q_n - \mathbb{E}[Q_n])^{\mathrm{t}}, \qquad (n \geqslant n_1).$$

Then, by (12), $\widetilde{Q}_n$ satisfies the recurrence

$$\left(\widetilde{Q}_n\right)^{\mathrm{t}} \stackrel{d}{=} \sum_{1\leqslant r\leqslant m} D_r^{(n)} \left(\widetilde{Q}^{(r)}_{I_r^{(n)}}\right)^{\mathrm{t}} + \widetilde{b}_n, \qquad (n \geqslant m-1),$$

where (denoting by $F_{n,r}$ the event $F_{n,r} := \{I_r^{(n)} \geqslant n_1\}$ and $F_{n,r}^c$ its complement)

$$D_r^{(n)} = \begin{bmatrix} \left(\dfrac{I_r^{(n)}}{n}\right) \mathbf{1}_{F_{n,r}} + \dfrac{\mathbf{1}_{F_{n,r}^c}}{\sqrt{C_N}\, n} & \dfrac{\sqrt{C_S I_r^{(n)}}}{\sqrt{C_N}\, n} \mathbf{1}_{F_{n,r}} + \dfrac{\mathbf{1}_{F_{n,r}^c}}{\sqrt{C_N} n} \\ 0 & \dfrac{\sqrt{I_r^{(n)}}}{\sqrt{n}} \mathbf{1}_{F_{n,r}} + \dfrac{\mathbf{1}_{F_{n,r}^c}}{\sigma_Y \sqrt{n}} \end{bmatrix},$$

$$\widetilde{b}_n = \begin{pmatrix} \dfrac{1}{\sqrt{C_N}\, n} \left(\displaystyle\sum_{1\leqslant r\leqslant m} (\nu(I_r^{(n)}) + \mu(I_r^{(n)})) - \nu(n) \right) \\ \dfrac{1}{C_s n} \left(1 - \mu(n) + \displaystyle\sum_{1\leqslant r\leqslant m} \nu(I_r^{(n)}) \right) \end{pmatrix}, \tag{34}$$

with assumptions on independence and identical distributions as in (12). Note that the asymptotic expressions for the variances and covariance between N_n and S_n imply that

$$\mathrm{Cov}(\widetilde{Q}_n) = \mathrm{Id}_2 + o(1),$$

where Id_2 denotes the 2×2 identity matrix and the $o(1)$-term means that all four components of $\mathrm{Cov}(\widetilde{Q}_n)$ converge to the corresponding components of Id_2, each $o(1)$ in the four components being different in general. In particular, $\mathrm{Cov}(\widetilde{Q}_n)$ is a symmetric, positive definite matrix for all $n \geqslant n_1$. Let $R_n := \mathrm{Id}_2$ for $0 \leqslant n < n_1$ and $R_n := (\mathrm{Cov}(\widetilde{Q}_n))^{1/2}$ for $n \geqslant n_1$. Note that, by continuity, we have

$$R_n = \mathrm{Id}_2 + o(1), \quad R_n^{-1} = \mathrm{Id}_2 + o(1). \tag{35}$$

Now normalize $\widetilde{Q}_n$ by $Y_n := R_n^{-1}\widetilde{Q}_n$, for $n \geqslant 1$, so that $\mathrm{Cov}(Y_n) = \mathrm{Id}_2$ for $n \geqslant n_1$, and

$$(Y_n)^{\mathrm{t}} \stackrel{d}{=} \sum_{1\leqslant r\leqslant m} F_r^{(n)} \left(Y^{(n)}_{I_r^{(n)}}\right)^{\mathrm{t}} + b^{(n)}, \qquad (n \geqslant n_1), \tag{36}$$

where $F_r^{(n)} = R_n^{-1} D_r^{(n)} R_{I_r^{(n)}}$ and $b^{(n)} = R_n^{-1}\widetilde{b}_n$, with assumptions on independence and identical distributions as in (12). From (34), (35) and (27), we then obtain the convergences

$$F_r^{(n)} \to \begin{bmatrix} V_r & 0 \\ 0 & V_r^{1/2} \end{bmatrix} =: F_r^*, \quad b^{(n)} \to \begin{pmatrix} C_N^{-1/2} b_N \\ 0 \end{pmatrix} =: b_N^*, \tag{37}$$

which hold in L_p for any $1 \leqslant p < \infty$ (we will need $p = 3$ below).

The limiting map. To describe the asymptotic behavior of Q_n, we use the following probability distribution on the space $\mathbb{R}^2$. In accordance with the notation in [39], we denote by $\mathcal{M}^2$ the space of all probability distributions on $\mathbb{R}^2$, by $\mathcal{M}_3^2$ the subspace of all $\mathcal{L}(Z) \in \mathcal{M}^2$ with $\|Z\|_3 < \infty$, and furthermore

$$\mathcal{M}_3^2(0,\mathrm{Id}_2) := \left\{ \mathcal{L}(Z) \in \mathcal{M}_3^2 \,\middle|\, \mathbb{E}[Z] = 0, \mathrm{Cov}(Z) = \mathrm{Id}_2 \right\}.$$

Define the map T'_N on $\mathcal{M}^2$:

$$\begin{aligned} T'_N : \mathcal{M}^2 &\to \mathcal{M}^2, \\ \mathcal{L}(Z) &\mapsto \mathcal{L}\left(\sum_{1 \leqslant r \leqslant m} F_r^* Z^{(r)} + b_N^* \right), \end{aligned} \tag{38}$$

where $Z^{(1)}, \ldots, Z^{(m)}$, $(F_1^*, \ldots, F_m^*, b_N^*)$ are independent and $Z^{(r)}$ is distributed as Z for all $r = 1, \ldots, m$. Here F_r^* and b_N^* are defined in (37).

Lemma 4.2. *The restriction of T'_N in (38) to $\mathcal{M}_3^2(0,\mathrm{Id}_2)$ has a unique fixed point $\mathcal{L}(X', \Lambda')$ which is a product measure, i.e., its components X' and Λ' are independent.*

Proof. We check first that the restriction of T'_N to $\mathcal{M}_3^2(0,\mathrm{Id}_2)$ maps into $\mathcal{M}_3^2(0,\mathrm{Id}_2)$:

- For any $\mu \in \mathcal{M}_3^2(0,\mathrm{Id}_2)$, we see, by independence and $\|b_N\|_3 < \infty$, that $T'_N(\mu) \in \mathcal{M}_3^2$.
- For the mean of $T'_N(\mu)$, we have, from $\mathbb{E}[b_N] = 0$, that $T'_N(\mu)$ is centered.
- For the covariance of $T'_N(\mu)$, we obtain (see also [39, Lemma 3.2]) the matrix

$$\mathbb{E}\begin{bmatrix} b_N^2/C_N & 0 \\ 0 & 0 \end{bmatrix} + m\mathbb{E}\begin{bmatrix} V_1^2 & 0 \\ 0 & V_1 \end{bmatrix} = \mathrm{Id}_2. \tag{39}$$

Thus $T'_N(\mu) \in \mathcal{M}_3^2(0,\mathrm{Id}_2)$. By Lemma 3.3 in [39], the existence of a unique fixed point $\mathcal{L}(X', \Lambda')$ follows from the inequality

$$m\mathbb{E}\,\|F_1^*\|_{\mathrm{op}}^3 = m\mathbb{E}[V_1^{3/2}] < 1.$$

Alternatively, Theorem 5.1 in [11] (or Lemma 3.1 in [39] as well) implies the existence of a unique fixed point $\mathcal{L}(X', \Lambda')$ in $\mathcal{M}_3^2(0,\mathrm{Id}_2)$.

To show that $\mathcal{L}(X', \Lambda')$ is a product measure we recall that the existence of the unique fixed point that we just obtained is based on the fact that the restriction of T'_N to $\mathcal{M}_3^2(0,\mathrm{Id}_2)$ is a contraction with respect to a complete metric on $\mathcal{M}_3^2(0,\mathrm{Id}_2)$. We do not introduce this metric, the Zolotarev metric ζ_3, here, since we do not require the special description of ζ_3. For more information on ζ_3, in particular the completeness of the metric space $(\mathcal{M}_3^2(0,\mathrm{Id}_2), \zeta_3)$, see [11].

We denote the space of probability measures on $\mathbb{R}$ by $\mathcal{M}$ and

$$\mathcal{M}_3(0,1) := \left\{ \mathcal{L}(Z) \in \mathcal{M} \,\middle|\, \mathbb{E}[|Z|^3] < \infty, \mathbb{E}[Z] = 0, \mathbb{V}(Z) = 1 \right\}.$$

Furthermore, the product of probability measures ν_1 and ν_2 on $\mathbb{R}$ by $\nu_1 \otimes \nu_2$. Consider the space

$$\mathcal{G} := \{ \nu_1 \otimes \mathcal{N}(0,1) \,|\, \nu_1 \in \mathcal{M}_3(0,1) \}.$$

Then $\mathcal{G} \subset \mathcal{M}_3^2(0,\mathrm{Id}_2)$.

To show that $(\mathcal{G}, \zeta_3)$ is a closed subspace of $(\mathcal{M}_3^2(0, \mathrm{Id}_2), \zeta_3)$, let $(\mu_n \otimes \mathcal{N}(0,1))_{n\geqslant 1}$ be a sequence in $\mathcal{G}$ that converges in $(\mathcal{M}_3^2(0, \mathrm{Id}_2), \zeta_3)$, say to $\mathcal{L}(Y_1, Y_2)$. Since ζ_3-convergence implies weak convergence, we first obtain that Y_2 is standard normally distributed. Clearly, we have $\mathcal{L}(Y_1) \in \mathcal{M}_3(0,1)$. Since a weak limit of product measures is a product measure (see e.g. [2, Theorem 2.8(ii)]), $\mathcal{L}(Y_1, Y_2)$ is a product measure. Now $(\mathcal{G}, \zeta_3)$ as a closed subspace of the complete space $(\mathcal{M}_3^2(0, \mathrm{Id}_2), \zeta_3)$ is complete.

We next show that the restriction of T'_N to $\mathcal{G}$ maps to $\mathcal{G}$. Note that only here do we use the fact that the second component in the definition of $\mathcal{G}$ is a normal distribution; see (40) below. For $\mu = \mu_1 \otimes \mathcal{N}(0,1) \in \mathcal{G}$, the covariance matrix of $T'_N(\mu) =: \mathcal{L}(Y_1, Y_2)$ is Id_2 by (39). Since Y_2 is distributed as $\sum_{1\leqslant r\leqslant m} V_r^{1/2} N_r$, where the N_j's are independent normals and independent of $(V_1, \dots, V_m)$, we see that $\mathcal{L}(Y_2) = \mathcal{N}(0,1)$. Thus it remains to show that, for $T'_N(\mu) \in \mathcal{G}$, the components Y_1 and Y_2 are independent. Let $A, B \subset \mathbb{R}$ be measurable and $(Y_1^{(1)}, Y_2^{(1)}), \dots, (Y_1^{(m)}, Y_2^{(m)})$ be independent random vectors that are independent of $(V_1, \dots, V_m)$ and identically distributed as μ. Then, denoting the distribution of $V = (V_1, \dots, V_m)$ by Υ and, for $v = (v_1, \dots, v_m)$, writing $t_N(v) := C_N^{-1/2}(\varphi_m + 2\varphi_m^2 \sum_{1\leqslant r\leqslant m} v_r \log v_r)$, we have

$$
\begin{aligned}
\mathbb{P}(Y_1 \in A, Y_2 \in B) &= \mathbb{P}\left(\sum_{1\leqslant r\leqslant m} V_r Y_r^{(1)} + t_N(V) \in A, \sum_{1\leqslant r\leqslant m} V_r^{1/2} Y_r^{(2)} \in B \right) \\
&= \int \mathbb{P}\left(\sum_{1\leqslant r\leqslant m} v_r Y_r^{(1)} + t_N(v) \in A, \sum_{1\leqslant r\leqslant m} v_r^{1/2} Y_r^{(2)} \in B \right) \mathrm{d}\Upsilon(v) \\
&= \int \mathbb{P}\left(\sum_{1\leqslant r\leqslant m} v_r Y_r^{(1)} + t_N(v) \in A \right) \mathbb{P}\left(\sum_{1\leqslant r\leqslant m} v_r^{1/2} Y_r^{(2)} \in B \right) \mathrm{d}\Upsilon(v) \\
&= \int \mathbb{P}\left(\sum_{1\leqslant r\leqslant m} v_r Y_r^{(1)} + t_N(v) \in A \right) \mathcal{N}(0,1)(B) \mathrm{d}\Upsilon(v) \qquad (40) \\
&= \mathbb{P}(Y_1 \in A)\mathbb{P}(Y_2 \in B).
\end{aligned}
$$

We then deduce that $T'_N(\mu) \in \mathcal{G}$ and T'_N maps $\mathcal{G}$ to $\mathcal{G}$.

Finally, Banach's fixed point theorem implies that the restriction of T'_N to $\mathcal{G}$ has a unique fixed point. Since $\mathcal{G} \subset \mathcal{M}_3^2(0, \mathrm{Id}_2)$, we find $\mathcal{L}(X', \Lambda') \in \mathcal{G}$. Consequently, X' and Λ' are independent. ∎

Proof of Theorem 2.6: NPL. The proof of Theorem 2.6 relies on Theorem 4.1 in [39]. The parameter d there is taken to be the dimension $d = 2$ here, and we choose the parameter $s = 3$. Note that the normalization in (10) is as required in [39, eq. (22)] and is identical to the normalization leading to the Y_n in (36). We need to check the conditions (24)–(26) in [39]. Condition (24) in our case is, with $F_r^{(n)}$ and $b^{(n)}$ as in (37),

$$
(F_1^{(n)}, \dots, F_m^{(n)}, b^{(n)}) \to (F_1^*, \dots, F_m^*, b_N^*)
$$

in L_3. This is satisfied by (37). Condition (25) in our case is also satisfied because

$$
\sum_{1\leqslant r\leqslant m} \|F_r^*\|_{\mathrm{op}}^3 = m\mathbb{E}[V_1^{3/2}] < 1.
$$

Finally, condition (25) is, for all $r = 1, \dots, m$ and all $\ell \in \mathbb{N}$,

$$\mathbb{E}\left[\mathbf{1}_{\{I_r^{(n)} \leqslant \ell\} \cup \{I_r^{(n)} = n\}} \|F_r^{(n)}\|_{\mathrm{op}}^3\right] \to 0.$$

Since $\|F_r^{(n)}\|_{\mathrm{op}}$ are uniformly bounded random variables, this condition is equivalent to

$$\mathbb{P}\left(I_r^{(n)} \leqslant \ell\right) \to 0,$$

which is satisfied in view of (27). Hence, Theorem 4.1 in [39] applies and implies the convergence $\mathrm{Cov}(Q_n)^{-1/2}(Q_n - \mathbb{E}[Q_n]) \to (X', \Lambda')$ in the metric ζ_3, which implies the stated convergence in distribution.

Note that the components of T'_N imply univariate recursive distributional equations for $\mathcal{L}(\Lambda')$ and $\mathcal{L}(X')$:

$$\Lambda' \stackrel{d}{=} \sum_{1 \leqslant r \leqslant m} \sqrt{V_r}\Lambda'^{(r)},$$
$$X' \stackrel{d}{=} \sum_{1 \leqslant r \leqslant m} V_r X'^{(r)} + C_N^{-1/2} b_N,$$

with conditions on independence and identical distributions corresponding to the definition of T'_N. Moreover, both equations are subject to the constraints of zero mean, unit variance and bounded third absolute moment. The solution for $\mathcal{L}(\Lambda')$ is given by the standard normal distribution, and a comparison of the equation for $\mathcal{L}(X')$ with (30) shows that X' is identically distributed as $C_N^{-1/2}X$ with X as in Theorem 2.7.

4.3 Limit law for NPL

From the previous two subsections, we see that the limit law of $(N_n - \mathbb{E}(N_n))/n$ is the unique solution, subject to zero mean and finite variance, of the recursive distributional equation

$$X \stackrel{d}{=} \sum_{1 \leqslant r \leqslant m} V_r X^{(r)} + \phi + 2\phi^2 \sum_{1 \leqslant r \leqslant m} V_r \log V_r,$$

where $X^{(1)}, \dots, X^{(m)}, V$ are independent and the $X^{(r)}$ have the same distribution as X.

Moreover, it is well-known that the limit law of $(K_n - \mathbb{E}(K_n))/n$, which we denote by $\mathcal{L}(K)$ in Section 2.2, is the unique solution, again subject to zero mean and finite variance, of

$$X \stackrel{d}{=} \sum_{1 \leqslant r \leqslant m} V_r X^{(r)} + 1 + 2\phi \sum_{1 \leqslant r \leqslant m} V_r \log V_r, \tag{41}$$

where the meaning of the notations is as above.

Comparing these two distributional recurrences, we see that the solution to the first one is $\mathcal{L}(\phi K)$. Thus, we have

$$\frac{N_n - \mathbb{E}(N_n)}{n} \xrightarrow{d} \phi K,$$

i.e., the limit law of K_n and N_n are up to a constant identical. In fact, if one is only interested in this result, then one does not need the analysis in the last two subsections but there are simpler approaches, as we discussed below.

4.4 Short proofs for the limit law of N_n

In this section, we discuss different means of proving directly the limit law for NPL without the detour via the bivariate setting from Sections 4.1 and 4.2.

Limit law for NPL by the contraction method. A first alternative approach to the limit law for NPL uses the contraction method and "over-normalizing" in recurrence (12). More precisely, normalize with an $\alpha < \alpha' < 1$ by

$$\mathcal{R}_n := \begin{bmatrix} n-1 & 0 \\ 0 & n^{-\alpha'} \end{bmatrix} \begin{pmatrix} N_n - \mathbb{E}[N_n] \\ S_n - \mathbb{E}[S_n] \end{pmatrix}, \qquad (n \geqslant 1).$$

Now the recurrence (12) leads to the limit equation

$$(\mathcal{R})^{\mathrm{t}} \stackrel{d}{=} \sum_{1 \leqslant r \leqslant m} \begin{bmatrix} V_r & 0 \\ 0 & V_r^{\alpha'} \end{bmatrix} \left(\mathcal{R}^{(r)}\right)^{\mathrm{t}} + \begin{pmatrix} b_N \\ 0 \end{pmatrix}, \tag{42}$$

with conditions on independence and identical distributions as in (30). Theorem 4.1 in [36] directly applies and implies that $\mathcal{R}_n \to \mathcal{R}$ in distribution and with second (mixed) moments, where $\mathcal{R}$ is the unique fixed point subject to zero mean and finite second moment of the recursive distributional equation (42). By substituting into (42), we see that $(\phi K, 0)$ has the distribution of $\mathcal{R}$, which implies that

$$\frac{N_n - \mathbb{E}[N_n]}{n} \xrightarrow{d} \phi K.$$

Univariate limit law for NPL via Slutsky's theorem. Another approach is to apply Slutsky's theorem. For that purpose, we consider the moment generating function

$$\bar{P}_n(u, v, w) = \mathbb{E}\left(e^{\bar{S}_n u + \bar{K}_n v + \bar{N}_n w}\right).$$

Then $\bar{P}_n$ satisfies the recurrence

$$\bar{P}_n(u, v, w) = \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \bar{P}_{j_1}(u + w, v, w) \cdots \bar{P}_{j_l}(u + w, v, w) e^{\Delta_{\mathbf{j}} u + \nabla_{\mathbf{j}} v + \delta_{\mathbf{j}} w},$$

with the initial conditions $P_n(u, v, w) = 1$ for $0 \leqslant n \leqslant m - 2$. Now define

$$V_n^{[KN]} := \mathrm{Cov}(K_n, N_n).$$

Then

$$V_n^{[KN]} = m \sum_{0 \leqslant j \leqslant n-m+1} \pi_{n,j} V_j^{[KN]} + b_n^{[KN]},$$

where

$$b_n^{[KN]} = \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(V_j^{[SK]} + \nabla_{\mathbf{j}} \delta_{\mathbf{j}}\right) = V_n^{[SK]} + \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(\nabla_{\mathbf{j}} \delta_{\mathbf{j}} - \Delta_{\mathbf{j}} \nabla_{\mathbf{j}}\right).$$

Observe that Lemma 3.3 and Lemma 3.3, together with the asymptotics of $V_n^{[SK]}$, imply that

$$b_n^{[KN]} \sim \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \nabla_{\mathbf{j}} \delta_{\mathbf{j}} \sim \frac{\phi}{\binom{n}{m-1}} \sum_{\mathbf{j}} \nabla_{\mathbf{j}}^2 \sim \phi b_n^{[K]}.$$

Consequently, by the same method of proofs used in Section 3, we see that

$$V_n^{[KN]} \sim \phi C_K n^2.$$

Now consider the difference

$$\begin{aligned} \mathbb{E}(\phi \bar{K}_n - \bar{N}_n)^2 &= \phi^2 V_n^{[K]} - 2\phi V_n^{[KN]} + V_n^{[N]} \\ &\sim \phi^2 C_K n^2 - 2\phi^2 C_K n^2 + \phi^2 C_K n^2 \\ &= o(n^2). \end{aligned}$$

Consequently, by Chebyshev's inequality, we obtain the convergence in probability

$$\frac{\phi \bar{K}_n - \bar{N}_n}{n} \xrightarrow{\mathbb{P}} 0.$$

From this, the claimed result follows from Slutsky's theorem and the limit law for KPL.

Note that this argument in addition gives the following consequence.

Corollary 4.3. *The correlation coefficient between K_n and N_n tends asymptotically to one*

$$\rho(K_n, N_n) \to 1.$$

Identical limit random variables. To the pair (N_n, K_n), we could as well apply the contraction method, and prove that the normalization $(N_n - \mathbb{E}(N_n))/n, (K_n - \mathbb{E}(K_n))/n)$ converges to a limit given by

$$(\mathcal{P})^{\mathrm{t}} \stackrel{d}{=} \sum_{1 \leqslant r \leqslant m} \begin{bmatrix} V_r & 0 \\ 0 & V_r \end{bmatrix} \left(\mathcal{P}^{(r)}\right)^{\mathrm{t}} + \begin{pmatrix} \phi b_K \\ b_K \end{pmatrix},$$

with conditions on independence and identical distributions as in (30) and subject to zero mean and finite second moment. By plugging in, we find that $(\phi K, K)$ has the limit distribution. This re-derives Corollary 4.3 and shows that the limit random variables (up to scaling) are even almost surely identical. It seems reasonable to conjecture that the sequences

$$\left(\frac{N_n - \mathbb{E}[N_n]}{\phi n}\right)_{n \geqslant 1}, \qquad \left(\frac{K_n - \mathbb{E}[K_n]}{n}\right)_{n \geqslant 1}$$

both convergence almost surely to the same random variable with the distribution of K. This requires the m-ary search trees to grow as a combinatorial Markov chain, which canonically is obtained by building up the tree from i.i.d. uniformly on $[0, 1]$ distributed data. For the notion of a combinatorial Markov chain and related results on binary search trees, see Grübel [19].

5 Extensions

The dependence and phase changes we established above for space requirement and path lengths in random m-ary search trees are not confined to these shape parameters, neither are they specific to m-ary search trees. The same study (including the same methods of proof) can be carried out for other shape parameters and other classes of random trees. We consider first random median-of-$(2t+1)$ search trees in this section, where we discuss the joint asymptotics of size (defined as the number of nodes with at least $2t$ descendants) and total key path length (which is also the major cost measure for Quicksort using the median-of-$(2t+1)$ technique). Random quadtrees will be also briefly discussed. Then we consider another line of extension, namely, to other shape parameters in these trees. Since the technicalities follow more or less the same pattern, we skip all proofs.

5.1 Random fringe-balanced binary search trees

Fringe-balanced binary search trees (FBBSTs) are binary search trees ($m = 2$) with local re-organizations for all subtrees of size exactly $2t+1$ into more balanced ones. In terms of quicksort, the corresponding tree structures choose at each partitioning stage the median of a sample of $2t+1$ elements to partition the elements into smaller and larger groups. For a precise description and other connections, see [8, 10]. The number of nodes $\mathcal{S}_n$ with at least $2t$ descendants (or the number of median-partitioning stages) and the total path length of these nodes (TPL; KPL=NPL for binary search trees) $\mathcal{X}_n$ of a random FBBST constructed from a random permutation of n elements satisfy the following distributional recurrence ($\mathcal{Q}_n := (\mathcal{X}_n, \mathcal{S}_n)$)

$$(\mathcal{Q}_n)^{\mathrm{t}} \stackrel{d}{=} \left(\mathcal{Q}^{(1)}_{I'_n}\right)^{\mathrm{t}} + \left(\mathcal{Q}^{(2)}_{n-1-I'_n}\right)^{\mathrm{t}} + \binom{n-1}{1}, \qquad (n \geqslant 2t+1),$$

with conditions on independence and identical distributions as in (12) and the initial conditions $\mathcal{S}_0 = \cdots = \mathcal{S}_{2t} = \mathcal{X}_0 = \cdots = \mathcal{X}_{2t} = 0$. Here

$$\mathbb{P}(I'_n = j) = \frac{\binom{j-1}{t}\binom{n-j}{t}}{\binom{n}{2t+1}} \qquad (t \leqslant j \leqslant n-1-t).$$

We start with the mean. First, for $\mathcal{S}_n$, it was proved in [8] that

$$\mathbb{E}(\mathcal{S}_n) = C_1(n+1) - 1 + \sum_{2 \leqslant k \leqslant 3} \frac{C_k}{\Gamma(\varrho_k)} n^{\varrho_k - 1} + o(n^{\alpha_t - 1}) \tag{43}$$

where

$$C_k = \frac{t!}{2(\varrho_k - 1)\varrho_k \cdots (\varrho_k + t - 1) \sum_{t \leqslant j \leqslant 2t} \frac{1}{j+\varrho_k}} \qquad (k = 1, \ldots, t+1),$$

with $\varrho_1 = 2 > \Re(\varrho_2) = \Re(\varrho_3) = \alpha_t > \Re(\varrho_4) \geqslant \cdots \geqslant \Re(\varrho_{t+1})$ being the zeros of the indicial equation

$$(z+t)\cdots(z+2t) - \frac{2(2t+1)!}{t!}.$$

In particular,

$$C_1 = \phi_t := \frac{1}{2(t+1)(H_{2t+2} - H_{t+1})}.$$

Moreover, using the transfer theorems from [8], we obtain, for the mean of $\mathcal{X}_n$,

$$\mathbb{E}(\mathcal{X}_n) = \frac{1}{H_{2t+2} - H_{t+1}}\, n \log n + c_t n + o(n),$$

for some constant c_t. The same method of proofs (asymptotic transfer and the approach used in Section 3.3) also leads to asymptotic estimates for the variances and the covariance between $\mathcal{X}_n$ and $\mathcal{S}_n$.

Theorem 5.1. *The variance of the number of non-leaf nodes $\mathcal{S}_n$ and that of the TPL $\mathcal{X}_n$ in a random FBBST, and their covariance satisfy*

$$\begin{aligned}
\mathbb{V}(\mathcal{S}_n) &\sim \begin{cases} D_S n, & \text{if } 1 \leqslant t \leqslant 58; \\ G_1(\beta_t \log n) n^{2\alpha_t - 2}, & \text{if } t \geqslant 59, \end{cases} \\
\operatorname{Cov}(\mathcal{S}_n, \mathcal{X}_n) &\sim \begin{cases} D_R n, & \text{if } 1 \leqslant t \leqslant 28; \\ G_2\left(\beta_t \log n\right) n^{\alpha_t}, & \text{if } t \geqslant 29, \end{cases} \\
\mathbb{V}(\mathcal{X}_n) &\sim D_X n^2,
\end{aligned}$$

where D_S, D_R are suitable constants, $\beta_t = \Im(\varrho_2)$, and all other constants and functions are given below.

The periodic functions in the above theorem are given by

$$\begin{aligned}
G_1(z) = 2\frac{|C_2|^2}{|\Gamma(\varrho_2)|^2} &\left(-1 + \frac{2(2t+1)!|\Gamma(\varrho_2 + t)|^2}{t!^2\Gamma(2\alpha_t + 2t) - 2t!(2t+1)!\Gamma(2\alpha_t + t - 1)}\right) \\
&+ 2\Re\left(\frac{C_2^2 e^{2iz}}{\Gamma(\varrho_2)^2}\left(-1 + 2\frac{2(2t+1)!\Gamma(\varrho_2 + t)^2}{t!^2\Gamma(2\varrho_2 + 2t) - 2t!(2t+1)!\Gamma(2\varrho_2 + t - 1)}\right)\right)
\end{aligned}$$

and

$$\begin{aligned}
G_2(z) = \Re\Bigg(\frac{C_2 e^{iz}}{\Gamma(\varrho_2)}\Bigg(&\frac{\varrho_2 + 2t + 1}{t+1} \\
&- \frac{(\varrho_2 + 2t + 1)\psi(\varrho_2 + 2t + 2) - (\varrho_2 + t)\psi(\varrho_2 + t + 1) - (t+1)(H_{t+1} - \gamma)}{(t+1)(H_{2t+2} - H_{t+1})}\Bigg)\Bigg),
\end{aligned}$$

respectively. Moreover, we have

$$D_X = \frac{1}{(H_{2t+2} - H_{t+1})^2}\left(\frac{2t+3}{t+1} H_{2t+2}^{(2)} - \frac{t+2}{t+1} H_{t+1}^{(2)} - \frac{\pi^2}{6}\right).$$

The limit law for the normalized TPL of random FBBSTs was first shown in the dissertation of Bruhn, [5]; see also [4, 8, 34, 40]. The phase change of the limit law of the normalized $\mathcal{S}_n$ was first discovered in [8].

To describe the joint limiting behavior of $\mathcal{S}_n$ and $\mathcal{X}_n$, we denote by $\mathcal{V}$ a random variable that is the median of $(2t+1)$ independent, identically distributed uniform $[0,1]$ random variables, i.e., a Beta$(t+1, t+1)$ distribution. We define the map T_{med} by

$$\begin{aligned} &T_{\text{med}} : \mathcal{M}^{\mathbb{R}\times\mathbb{C}} \to \mathcal{M}^{\mathbb{R}\times\mathbb{C}}, \\ &\quad \mathcal{L}(Z,W) \mapsto \mathcal{L}\left(\begin{bmatrix} \mathcal{V} & 0 \\ 0 & \mathcal{V}^{\varrho_2} \end{bmatrix}\begin{pmatrix} Z^{(1)} \\ W^{(1)} \end{pmatrix} + \begin{bmatrix} 1-\mathcal{V} & 0 \\ 0 & (1-\mathcal{V})^{\varrho_2} \end{bmatrix}\begin{pmatrix} Z^{(2)} \\ W^{(2)} \end{pmatrix} + \begin{pmatrix} b_M \\ 0 \end{pmatrix}\right), \end{aligned}$$

with conditions on independence and distributions as in (30) and

$$b_M := 1 + \frac{1}{H_{2t+2} - H_{t+1}} \left(\mathcal{V}\log\mathcal{V} + (1-\mathcal{V})\log(1-\mathcal{V})\right).$$

Then Lemma 4.1 and its proof also apply to the map T_{med} as long as $t \geqslant 59$. The normalization used is given by

$$\mathcal{Y}_n := \left(\frac{\mathcal{X}_n - \mathbb{E}(\mathcal{X}_n)}{n}, \frac{\mathcal{S}_n - C_1 n}{n^{\alpha_t - 1}}\right), \qquad (n \geqslant 1). \tag{44}$$

We have the following asymptotic behavior for $t \geqslant 59$. Rewrite (43) as

$$\mathbb{E}(\mathcal{S}_n) = C_1(n+1) - 1 + \Re(\vartheta n^{\varrho_2}) + o(n^{\alpha_t - 1}), \tag{45}$$

where $\vartheta := 2\Re(C_2/\Gamma(\varrho_2))$.

Theorem 5.2. *Assume $t \geqslant 59$. Let $\mathcal{Y}_n$ be the normalization of TPL and the number of non-leaf nodes in a random FBBST defined in (44). Denote by $\mathcal{L}(X_{\text{med}}, \Lambda_{\text{med}})$ the unique fixed point of the restriction of T_{med} to $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}((0,\vartheta))$ with ϑ defined in (45). Then, denoting by $\beta_t := \Im(\varrho_2)$, we have*

$$\ell_2\left(\mathcal{Y}_n, (X_{\text{med}}, \Re(n^{i\beta_t}\Lambda_{\text{med}}))\right) \to 0, \qquad (n \to \infty).$$

For the range of $1 \leqslant t \leqslant 58$, we define $b^*_{\text{med}} := (D_X^{-1/2} b_M, 0)^{\mathrm{t}}$ and the map T'_{med} on $\mathcal{M}^2$:

$$\begin{aligned} &T'_{\text{med}} : \mathcal{M}^2 \to \mathcal{M}^2, \\ &\quad \mathcal{L}(Z,W) \mapsto \mathcal{L}\left(\begin{bmatrix} \mathcal{V} & 0 \\ 0 & \mathcal{V}^{1/2} \end{bmatrix}\begin{pmatrix} Z^{(1)} \\ W^{(1)} \end{pmatrix} + \begin{bmatrix} 1-\mathcal{V} & 0 \\ 0 & (1-\mathcal{V})^{1/2} \end{bmatrix}\begin{pmatrix} Z^{(2)} \\ W^{(2)} \end{pmatrix} + b^*_{\text{med}}\right), \end{aligned}$$

with conditions on independence and distributions as in (38). Again Lemma 4.2 and its proof apply to T'_{med} and imply that the restriction of T'_{med} to $\mathcal{M}_3^2(0, \mathrm{Id}_2)$ has a unique fixed point $\mathcal{L}(X'_{\text{med}}, \Lambda'_{\text{med}})$.

Similar to the small m case of m-ary search trees, the remaining range $1 \leqslant t \leqslant 58$ also leads to a convergence in distribution.

Theorem 5.3. *Assume $1 \leqslant t \leqslant 58$. Let $\mathcal{Q}_n = (\mathcal{X}_n, \mathcal{S}_n)$ be the vector of TPL and the number of non-leaf nodes in a random FBBST. With $\mathcal{L}(X'_{\text{med}}, \Lambda'_{\text{med}})$ as above, we have*

$$\mathrm{Cov}(\mathcal{Q}_n)^{-1/2}\left(\mathcal{Q}_n - \mathbb{E}[\mathcal{Q}_n]\right) \xrightarrow{d} \mathcal{L}\left(X'_{\text{med}}, \Lambda'_{\text{med}}\right),$$

where Λ'_{med} is a standard normal distribution. Moreover, X'_{med} and Λ'_{med} are independent.

5.2 Random quadtrees

Point quadtrees, first proposed by Finkel and Bentley [15], are one of the most natural extensions of binary search trees to multivariate data in which each point splits the d-dimensional space into 2^d subspaces, corresponding to 2^d subtrees in the corresponding tree structure. For a precise definition of random d-dimensional quadtrees; see [7, 30]. Since the space requirement is a constant, we discuss the number of leaves L_n and the internal path length Ξ_n in this section. Note that for the pair $\mathcal{W}_n := (\Xi_n, L_n)$, we have, for all $n \geqslant 2$,

$$(\mathcal{W}_n)^{\mathrm{t}} \stackrel{d}{=} \sum_{1\leqslant r\leqslant 2^d} \left(\mathcal{W}_{J_r}^{(r)}\right)^{\mathrm{t}} + \begin{pmatrix} n-1 \\ 0 \end{pmatrix},$$

with conditions on independence and identical distributions as in (12), where the initial conditions are $L_0 = 0, L_1 = 1, \Xi_0 = \Xi_1 = 0$. Moreover, the underlying splitting probabilities are given by

$$\mathbb{P}(J_1 = j_1, \dots, J_{2^d} = j_{2^d}) = \binom{n-1}{j_1, \dots, j_{2^d}} \int_{[0,1]^d} q_1(\mathbf{x})^{j_1} \cdots q_{2^d}(\mathbf{x})^{j_{2^d}} \mathrm{d}\mathbf{x},$$

where $j_1, \dots, j_{2^d} \geqslant 0, j_1 + \cdots + j_{2^d} = n-1, \mathbf{x} = (x_1, \dots, x_d)$ and

$$q_h(\mathbf{x}) = \prod_{1\leqslant l\leqslant d} \left((1-b_l)x_l + b_l(1-x_l)\right),$$

with $(b_1, \dots, b_d)_2$ being the binary representation of $h-1$.

First, it was proved in [7] that the mean of L_n satisfies, for $d \geqslant 2$,

$$\mathbb{E}(L_n) = \chi_d n + c_+ n^{\hat{\alpha}+i\hat{\beta}} + c_- n^{\hat{\alpha}-i\hat{\beta}} + \frac{\chi_d}{2^d-1} + o(n^{\hat{\alpha}}), \tag{46}$$

where χ_d, c_+, c_- (which is the conjugate of c_+) are given in [7], and $2e^{2\pi i/d} = \hat{\alpha} + 1 + i\hat{\beta}$. Moreover, the asymptotic transfer results in [7] also lead to the asymptotic approximation (see also [16])

$$\mathbb{E}(\Xi_n) = \frac{2}{d} n \log n + \hat{c} n + o(n),$$

for some explicitly computable constant $\hat{c}$. In a similar manner, we can characterize the asymptotics of the variances and the covariance.

Theorem 5.4. *For the number of leaves L_n and the internal path length Ξ_n in random d-dimensional quadtrees, we have*

$$\begin{aligned}
\mathbb{V}(L_n) &\sim \begin{cases} E_L n, & \text{if } 1 \leqslant d \leqslant 8; \\ P_1\big(\hat{\beta}\log n\big) n^{2\hat{\alpha}}, & \text{if } d \geqslant 9, \end{cases} \\
\mathrm{Cov}(\Xi_n, L_n) &\sim \begin{cases} E_R n, & \text{if } 1 \leqslant d \leqslant 5; \\ P_2\big(\hat{\beta}\log n\big) n^{\hat{\alpha}+1}, & \text{if } d \geqslant 6, \end{cases} \\
\mathbb{V}(\Xi_n) &\sim E_X n^2,
\end{aligned}$$

where E_L, E_R are suitable constants, $\hat{\beta} := 2\sin\frac{2\pi}{d}$, and all other constants and functions are given below.

The periodic functions above are given by

$$P_1(z) = 2\frac{(2\hat{\alpha}+1)^d}{(2\hat{\alpha}+1)^d - 2^d}|c_+|^2 c_L(\hat{\alpha}+i\hat{\beta}, \hat{\alpha}-i\hat{\beta}) \\ + 2\Re\left(\frac{(2\hat{\alpha}+2i\hat{\beta}+1)^d}{(2\hat{\alpha}+2i\hat{\beta}+1)^d - 2^d} c_+^2 c_L(\hat{\alpha}+i\hat{\beta}, \hat{\alpha}+i\hat{\beta}) e^{2iz}\right),$$

where $c_L(u,v) = 1 - \eta(0,u) - \eta(0,v) + 2^d\eta(u,v)$ with

$$\eta(u,v) := \left(\frac{1}{u+v+1} + \frac{\Gamma(u+1)\Gamma(v+1)}{\Gamma(u+v+2)}\right)^d$$

and

$$P_2(z) = 2\Re\left(\frac{(\hat{\alpha}+i\hat{\beta}+2)^d}{(\hat{\alpha}+i\hat{\beta}+2)^d - 2^d} c_+ c_K(\hat{\alpha}+i\hat{\beta}) e^{iz}\right),$$

where

$$c_K(u,v) = \eta(0,u) + \frac{2^{d+1}}{d}\frac{\partial}{\partial v}\eta(u,v)\Big|_{v=1}.$$

Finally,

$$E_X = \frac{3^d}{3^d - 2^d} \cdot \frac{21 - 2\pi^2}{9d}.$$

The limit law for the normalized internal path length of random d-dimensional quadtrees was first obtained in [38]; see also [4, 7, 34]. The asymptotic behavior of the normalized number of leaves together with its phase change was first discovered in [7]; see also [9, 23, 24, 25] for closely related types of phase changes.

We now describe the joint behavior of Ξ_n and L_n. A random variable U uniformly distributed over the unit hypercube $[0,1]^d$ decomposes this cube into 2^d quadrants by drawing the d hyperplanes through U perpendicular to the edges of the cube. Choose an ordering of these quadrants and denote their volumes by $\langle U\rangle_1, \ldots, \langle U\rangle_{2^d}$; see [38, Section 2]. Now define the map T_{quad} by (with $\delta_2 := 2e^{2\pi i/d}$)

$$T_{\text{quad}} : \mathcal{M}^{\mathbb{R}\times\mathbb{C}} \to \mathcal{M}^{\mathbb{R}\times\mathbb{C}}, \\ \mathcal{L}(Z,W) \mapsto \mathcal{L}\left(\sum_{1\leqslant r\leqslant 2^d} \begin{bmatrix} \langle U\rangle_r & 0 \\ 0 & \langle U\rangle_r^{\delta_2} \end{bmatrix} \begin{pmatrix} Z^{(r)} \\ W^{(r)} \end{pmatrix} + \begin{pmatrix} b_Q \\ 0 \end{pmatrix}\right),$$

with conditions on independence and distributions as in (30), and

$$b_Q := 1 + \frac{2}{d}\sum_{1\leqslant r\leqslant 2^d} \langle U\rangle_r \log\langle U\rangle_r.$$

Then Lemma 4.1 and its proof also apply to map T_{quad} as long as $d \geqslant 9$. The normalization used is given by

$$\mathcal{V}_n := \left(\frac{\Xi_n - \mathbb{E}(\Xi_n)}{n}, \frac{L_n - \chi_d n}{n^{\hat{\alpha}}}\right) \qquad (n \geqslant 1). \tag{47}$$

Rewrite (46) as

$$\mathbb{E}(L_n) = \chi_d n + \Re(\hat{\vartheta} n^{\hat{\alpha}+i\hat{\beta}}) + \frac{\chi_d}{2^d - 1} + o(n^{\hat{\alpha}}), \tag{48}$$

where $\hat{\vartheta} = 2c_+$.

Theorem 5.5. *Assume $d \geqslant 9$. Let $\mathcal{V}_n$ denote the normalization of the internal path length and the number of leaves in a random d-dimensional quadtree defined in (47). Denote by $\mathcal{L}(X_{\mathrm{quad}}, \Lambda_{\mathrm{quad}})$ the unique fixed point of the restriction of T_{quad} to $\mathcal{M}_2^{\mathbb{R}\times\mathbb{C}}((0, \hat{\vartheta}))$ with $\hat{\vartheta}$ defined in (48). Then we have*

$$\ell_2\left(\mathcal{V}_n, \left(X_{\mathrm{quad}}, \Re\left(n^{i\hat{\beta}}\Lambda_{\mathrm{quad}}\right)\right)\right) \to 0.$$

For the remaining range of $1 \leqslant d \leqslant 8$, we define $b^*_{\mathrm{quad}} := (E_X^{-1/2} b_M, 0)^{\mathrm{t}}$ and the map T'_{quad} on $\mathcal{M}^2$

$$\begin{aligned} &T'_{\mathrm{quad}} : \mathcal{M}^2 \to \mathcal{M}^2, \\ &\mathcal{L}(Z, W) \mapsto \mathcal{L}\left(\sum_{1\leqslant r\leqslant 2^d} \begin{bmatrix} \langle U\rangle_r & 0 \\ 0 & \langle U\rangle_r^{1/2} \end{bmatrix} \begin{pmatrix} Z^{(r)} \\ W^{(r)} \end{pmatrix} + b^*_{\mathrm{quad}}\right), \end{aligned}$$

with conditions on independence and distributions as in (38). Similarly, Lemma 4.2 and its proof again apply to T'_{quad} and imply that the restriction of T'_{quad} to $\mathcal{M}_3^2(0, \mathrm{Id}_2)$ has a unique fixed point $\mathcal{L}(X'_{\mathrm{quad}}, \Lambda'_{\mathrm{quad}})$.

Theorem 5.6. *Assume $1 \leqslant d \leqslant 8$. Let $\mathcal{V}_n = (\Xi_n, L_n)$ denote the vector of internal path length and the number of leaves in a random d-dimensional quadtree. With $\mathcal{L}(X'_{\mathrm{quad}}, \Lambda'_{\mathrm{quad}})$ as above, we have*

$$\mathrm{Cov}(\mathcal{V}_n)^{-1/2}(\mathcal{V}_n - \mathbb{E}[\mathcal{V}_n]) \xrightarrow{d} \mathcal{L}\left(X'_{\mathrm{quad}}, \Lambda'_{\mathrm{quad}}\right),$$

where Λ'_{quad} is a standard normal distribution, and X'_{quad}, and Λ'_{quad} are independent.

The case when $d = 1$ corresponds to binary search trees, or equivalently, to Hoare's quicksort, and the above theorem can be re-worded as follows. *The number of comparisons and the number of partitioning stages used by Hoare's quicksort are asymptotically uncorrelated and independent.* Note that our results in the previous section for random FBBSTs give indeed a stronger statement for the asymptotic independence or asymptotical periodicity for quicksort using median-of-$(2t + 1)$.

5.3 More general shape parameters

Our study can be extended to other shape parameters. For random m-ary search trees, the generality of Proposition 3.1 provides an effective means of widening our study to a broader class of "toll functions" in the definitions of S_n, K_n and N_n. For example, the following extensions are straightforward.

$$S_n \stackrel{d}{=} S^{(1)}_{I_1} + \cdots + S^{(m)}_{I_m} + \begin{cases} c + o(n^{-\varepsilon}), & \text{if } 2 \leqslant m \leqslant 13; \\ o(n^{\alpha-1}), & \text{if } m \geqslant 14 \end{cases} \tag{49}$$

for some constant c, and

– $K_n \stackrel{d}{=} K_{I_1}^{(1)} + \cdots + K_{I_m}^{(m)} + n + t_n$ with

$$t_n = o(n) \quad \text{and} \quad \left| \sum_n t_n n^{-2} \right| < \infty, \tag{50}$$

and

– $N_n \stackrel{d}{=} N_{I_1}^{(1)} + \cdots + N_{I_m}^{(m)} + S_{I_1}^{(1)} + \cdots + S_{I_m}^{(m)} + t_n$, where the S_n's satisfy (49) and t_n satisfies (50).

Because the same iff-condition (50) also appears in the recurrence relations arising from the two other classes of random trees (see [7, 8]), exactly the same conditions can be used to extend the consideration for FBBSTs and quadtrees. Details are omitted here.

References

[1] R. A. Baeza-Yates. Some average measures in m-ary search trees. *Inform. Process. Lett.*, 25(6):375–381, 1987.

[2] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[3] P. Bindjeme and J. A. Fill. Exact l^2-distance from the limit for quicksort key comparisons (extended abstract). *Discrete Math. Theor. Comput. Sci., Nancy*, pages 339–348, 2012.

[4] N. Broutin and C. Holmgren. The total path length of split trees. *Ann. Appl. Probab*, 22:1745–1777, 2012.

[5] V. Bruhn. *Eine Methode zur asymptotischen Behandlung einer Klasse von Rekursionsgleichungen mit einer Anwendung in der stochastischen Analyse des Quicksort-Algorithmus*. PhD thesis, Christian-Albrechts-Universität zu Kiel Dissertation, 1996.

[6] B. Chauvin and N. Pouyanne. m-ary search trees when $m \geqslant 27$: a strong asymptotics for the space requirements. *Random Structures Algorithms*, 24(2):133–154, 2004.

[7] H.-H. Chern, M. Fuchs, and H.-K. Hwang. Phase changes in random point quadtrees. *ACM Trans. Algorithms*, 3(2):Art. 12, 51, 2007.

[8] H.-H. Chern and H.-K. Hwang. Phase changes in random m-ary search trees and generalized quicksort. *Random Structures Algorithms*, 19(3-4):316–358, 2001.

[9] D. S. Dean and S. N. Majumdar. Phase transition in a random fragmentation problem with applications to computer science. *J. Phys. A*, 35(32):L501–L507, 2002.

[10] L. Devroye. On the expected height of fringe-balanced trees. *Acta Inform.*, 30(5):459–466, 1993.

[11] M. Drmota, S. Janson, and R. Neininger. A functional limit theorem for the profile of search trees. *Ann. Appl. Probab.*, 18(1):288–333, 2008.

[12] J. A. Fill. Distributional convergence for the number of symbol comparisons used by quicksort. *Ann. Appl. Probab*, 23:1129–1147, 2013.

[13] J. A. Fill and N. Kapur. The space requirement of m-ary search trees: distributional asymptotics for $m \geqslant 27$. *Invited paper, Proceedings of the 7th Iranian Statistical Conference. Available via* http://www.ams.jhu.edu/~fill/papers/periodic.pdf, 7, 2004.

[14] J. A. Fill and N. Kapur. Transfer theorems and asymptotic distributional results for m-ary search trees. *Random Structures Algorithms*, 26(4):359–391, 2005.

[15] R. A. Finkel and J. L. Bentley. Quad trees: A data structure for retrieval on composite keys. *Acta Inf.*, 4:1–9, 1974.

[16] P. Flajolet, G. Labelle, L. Laforest, and B. Salvy. Hypergeometrics and the cost structure of quadtrees. *Random Structures Algorithms*, 7(2):117–144, 1995.

[17] M. Fuchs. A note on the quicksort asymptotics. *Random Structures Algorithms*, 46(4):677–687, 2015.

[18] M. Fuchs and H.-K. Hwang. Dependence between size and external path-length in random tries. preprint, 2016.

[19] R. Grübel. Search trees: metric aspects and strong limit theorems. *Ann. Appl. Probab.*, 24(3):1269–1297, 2014.

[20] R. Grübel and Z. Kabluchko. A functional central limit theorem for branching random walks, almost sure weak convergence, and applications to random trees. 2014.

[21] C. Holmgren. A weakly 1-stable distribution for the number of random records and cuttings in split trees. *Adv. in Appl. Probab.*, 43(1):151–177, 2011.

[22] H.-K. Hwang. Second phase changes in random m-ary search trees and generalized quicksort: convergence rates. *Ann. Probab.*, 31(2):609–629, 2003.

[23] S. Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.*, 110(2):177–245, 2004.

[24] S. Janson. Congruence properties of depths in some random trees. *ALEA Lat. Am. J. Probab. Math. Stat.*, 1:347–366, 2006.

[25] S. Janson and R. Neininger. The size of random fragmentation trees. *Probab. Theory Related Fields*, 142(3-4):399–442, 2008.

[26] M. Knape and R. Neininger. Pólya urns via the contraction method. *Combin. Probab. Comput.*, 23(6):1148–1186, 2014.

[27] D. E. Knuth. *The Art of Computer Programming. Vol. 3. Sorting and Searching.* Addison-Wesley, Reading, MA, 1998. Second edition.

[28] W. Lew and H. M. Mahmoud. The joint distribution of elastic buckets in multiway search trees. *SIAM J. Comput.*, 23(5):1050–1074, 1994.

[29] H. M. Mahmoud. On the average internal path length of m-ary search trees. *Acta Inform.*, 23(1):111–117, 1986.

[30] H. M. Mahmoud. *Evolution of Random Search Trees*. Wiley-Interscience. John Wiley & Sons, Inc., New York, 1992. A Wiley-Interscience Publication.

[31] H. M. Mahmoud and B. Pittel. Analysis of the space of search trees under the random insertion algorithm. *J. Algorithms*, 10(1):52–75, 1989.

[32] C. Mailler. Describing the asymptotic behaviour of multicolour Pólya urns via smoothing systems analysis. *arXiv:1407.2879*, 2014.

[33] S. N. Majumdar, D. S. Dean, and P. L. Krapivsky. Understanding search trees via statistical physics. *Pramana*, 64:1175–1189, 2005.

[34] G. O. Munsonius. On the asymptotic internal path length and the asymptotic Wiener index of random split trees. *Electron. J. Probab.*, 16:no. 35, 1020–1047, 2011.

[35] R. Muntz and R. Uzgalis. Dynamic storage allocation for binary search trees in a two-level memory. *Proceedings of the Princeton Conference on Information Sciences and Systems*, 4:345–349, 1971.

[36] R. Neininger. On a multivariate contraction method for random recursive structures with applications to Quicksort. *Random Structures Algorithms*, 19(3-4):498–524, 2001.

[37] R. Neininger. Refined quicksort asymptotics. *Random Structures Algorithms*, 46(2):346–361, 2015.

[38] R. Neininger and L. Rüschendorf. On the internal path length of d-dimensional quad trees. *Random Structures Algorithms*, 15(1):25–41, 1999.

[39] R. Neininger and L. Rüschendorf. A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.*, 14(1):378–418, 2004.

[40] U. Rösler. On the analysis of stochastic divide and conquer algorithms. *Algorithmica*, 29(1-2):238–261, 2001.

[41] H. Sulzbach. On martingale tail sums for the path length in random trees. *http://arXiv:1412.3508*, 2014.