

## ON TAIL BOUNDS FOR RANDOM RECURSIVE TREES

GÖTZ OLAF MUNSONIUS,\* *Goethe University Frankfurt*

### Abstract

We consider a multivariate distributional recursion of sum type, as arises in the probabilistic analysis of algorithms and random trees. We prove an upper tail bound for the solution using Chernoff's bounding technique by estimating the Laplace transform. The problem is traced back to the corresponding problem for binary search trees by stochastic domination. The result obtained is applied to the internal path length and Wiener index of random  $b$ -ary recursive trees with weighted edges and random linear recursive trees. Finally, lower tail bounds for the Wiener index of these trees are given.

*Keywords:* Random tree; probabilistic analysis of algorithms; tail bound; path length; Wiener index

2010 Mathematics Subject Classification: Primary 60C05; 05C05; 05C80; 60E15

### 1. Introduction

Many parameters of recursive algorithms, trees, or other recursive structures can often be described by a recursion of sum type:

$$X_n \stackrel{D}{=} \sum_{i=1}^b A_i(I_n) X_{I_{n,i}}^{(i)} + d(I_n, Z), \quad n \geq 2. \quad (1)$$

Here  $X_n^{(1)}, \dots, X_n^{(b)}$  have the same distribution as  $X_n$ ,  $d: \mathbb{R}^b \times \mathbb{R}^b \rightarrow \mathbb{R}^k$  and  $A_i: \mathbb{R}^b \rightarrow \mathbb{R}^{k \times k}$  are deterministic functions,  $I_n = (I_{n,1}, \dots, I_{n,b}) \in \{0, \dots, n-1\}^b$  and  $Z \in \mathbb{R}_{\geq 0}^b$  are random vectors with  $E[d(I_n, Z)] = 0$ , and  $X_n^{(1)}, \dots, X_n^{(b)}$ ,  $I_n$ , and  $Z$  are independent. By ' $\stackrel{D}{=}$ ', we denote equality in distribution.

From the algorithmic point of view, such a recurrence arises by considering so-called divide-and-conquer algorithms. Let  $Y_n$  denote the parameter of interest of the algorithm applied to a problem of size  $n$ . The algorithm splits the large problem into  $b$  subproblems of the smaller sizes  $I_{n,1}, \dots, I_{n,b}$ . If the considered parameter  $Y_n$  is essentially given by the (possible weighted) sum of the corresponding parameters of the smaller subproblems, for a matrix  $C_n$ , the vector  $X_n := C_n(Y_n - E[Y_n])$  satisfies recurrence (1), where the coefficients  $A_i(I_n)$  are the weights of the subproblems (scaled by  $C_n$  and  $C_{I_{n,i}}$ ) and the additional function  $d$  gives the cost for splitting the problem in this manner and merging the solutions of the subproblems to a solution of the size- $n$  problem. The vector  $Z$  provides more universality.

One famous example of a parameter satisfying recursion (1) is the distribution of the number of comparisons made by Quicksort, which is equal in distribution to the internal path length of the random binary search tree. McDiarmid and Hayward [11] used martingale difference methods

Received 5 July 2011; revision received 2 December 2011.

\* Postal address: Institute of Mathematics, Goethe University Frankfurt, 60054 Frankfurt am Main, Germany.

Email address: munsonius@math.uni-frankfurt.de

to show upper tail bounds for the number of comparisons made by Quicksort. Rösler [15] as well as Fill and Janson [6] obtained upper bounds for its Laplace transform by induction. Having upper bounds for the Laplace transform, they obtained upper bounds for the tails of the distribution by application of Chernoff’s bounding technique. Ali Khan and Neininger [2] generalized this procedure to the two-dimensional recursion for the Wiener index and the internal path length of the random binary search tree, extending their technique used in [1] to analyze tail bounds for the complexity of a randomized algorithm to evaluate game trees.

In this paper we apply the method of [2] to multivariate functionals satisfying recursion (1), where the operator norm of the coefficient matrices  $A_i$  can be stochastically bounded in a certain way.

We denote by ‘ $\leq_{st}$ ’ the stochastic order, by  $\|\cdot\|$  the Euclidean norm in  $\mathbb{R}^k$ , and by  $\|\cdot\|_{op}$  the operator norm for matrices. Let  $U$  be a random variable uniformly distributed on  $[0, 1]$ . The fundamental result of this paper is the following theorem.

**Theorem 1.** *Let  $X_n$  be a solution of the distributional recursion (1). Assume that  $X_1 = 0$ ,  $\|d(I_n, Z)\| \leq D$  almost surely for all  $n \in \mathbb{N}$  and a constant  $D \in \mathbb{R}$ , and that*

$$\sum_{i=1}^b \|A_i(I_n)\|_{op}^2 \leq_{st} 1 - U(1 - U)$$

as well as  $\|A_i\|_{op} \leq 1$  for all  $i \in \{1, \dots, b\}$ . Let  $\gamma \approx 2.0047$  be the positive solution of  $\frac{12}{7} = e^{2/\gamma} - 2/\gamma$ , and let  $L_0 \approx 5.0177$  be the largest root of  $e^L = 6L^2$ . Then we have, for all  $t > 0$ ,  $n \in \mathbb{N}$ , and any component  $X_{n,j}$  of  $X_n$  ( $j \in \{1, \dots, k\}$ ) with  $C := 48D/\gamma + D\sqrt{48(48/\gamma^2 - 5)}$ ,

$$P(X_{n,j} > t) \leq \begin{cases} \exp\left(-\frac{t^2}{10\gamma^2 D^2}\right) & \text{if } 0 \leq t \leq 5\gamma D, \\ \exp\left(\frac{5}{2} - \frac{t}{\gamma D}\right) & \text{if } 5\gamma D < t \leq C, \\ \exp\left(-\frac{t^2}{96D^2}\right) & \text{if } C < t \leq 48DL_0, \\ \exp\left(24L_0^2 - \frac{L_0}{D}t\right) & \text{if } 48DL_0 < t \leq 4De^{L_0}, \\ \exp\left(\frac{t}{D} - \frac{t}{D} \log\left(\frac{t}{4D}\right)\right) & \text{if } 4De^{L_0} < t. \end{cases} \tag{2}$$

The same bounds hold for the left tail  $P(X_{n,j} < -t)$ .

As an application of Theorem 1, we obtain upper tail bounds for the distribution of the internal path length and the Wiener index in random  $b$ -ary recursive trees with weighted edges by showing the stochastic domination condition. The distance between two nodes in a tree with weighted edges is defined as the sum of the edge weights along the unique path between the two nodes. Then, the internal path length of a rooted tree is the sum of all node depths of the tree where the depth of a node is its distance to the root, and the Wiener index is the sum of the distances between all unordered pairs of nodes.

The  $b$ -ary recursive tree with weighted edges is a rooted, ordered, labelled tree where the outdegree is bounded by  $b$ , the labels along each path beginning at the root increase, and the edges are endowed with random weights. It can be considered as a special case of the tree model of Broutin and Devroye [4] in discrete time, where the lifetimes of the edges are independent,

exponentially distributed random variables. The shape of the random tree is obtained as an increasing tree by Bergeron *et al.* [3] and is a special case of the general model of random trees in [5].

We define this tree model by the following recursive procedure. We consider the infinite complete  $b$ -ary rooted, ordered tree and start with the root as the first internal node and its  $b$  children as external nodes. Given *the random  $b$ -ary recursive tree* with  $n - 1$  internal nodes, the  $n$ th internal node is added in the following way. We choose a random node uniformly distributed on the set of all current external nodes, change it to an internal node, and add the  $b$  children of this new node to the set of external nodes. Finally, the nodes are labelled in the order of their appearance.

Let  $Z := (Z_1, \dots, Z_b) \in \mathbb{R}_{\geq 0}^b$  be a random vector with nonnegative entries and attach to every node  $u$  of the complete infinite  $b$ -ary tree an independent copy  $Z^{(u)}$  of  $Z$ . We consider the entries of  $Z^{(u)}$  to be the weights of the edges from  $u$  to its  $b$  children. If all the  $Z^{(u)}$  are independent of  $T_n$ , we refer to  $T_n$  associated with the family  $\{Z^{(u)}\}$  as a *random  $b$ -ary recursive tree with edge weights  $Z$* .

**Theorem 2.** *Let  $Y_n := (\Psi_n, \Phi_n)^\top$  denote the vector consisting of the Wiener index and the internal path length of a random  $b$ -ary recursive tree of size  $n$  with edge weights  $Z$ , where  $\|Z\|$  is bounded almost surely. Then there exists a constant  $D$  such that  $X_n$  in the recursive formula (1) is given by*

$$X_n := \begin{bmatrix} \frac{1}{n^2} & 0 \\ 0 & \frac{1}{n} \end{bmatrix} (Y_n - E[Y_n])$$

almost surely,  $\|d(I_n, Z)\| \leq D$ , and the bounds in (2) are valid for

$$P\left(\frac{\Psi_n - E[\Psi_n]}{n^2} > t\right) \quad \text{and} \quad P\left(\frac{\Phi_n - E[\Phi_n]}{n} > t\right)$$

as well as for the corresponding left tails  $P(X_{n,j} < -t)$  for  $j = 1, 2$ .

Using the asymptotic expansion of the expectation of the internal path length and the Wiener index, the following asymptotic tail bounds are obtained. We use the notation  $\log^{(k)} n := \log(\log^{(k-1)} n)$  for  $k \geq 2$ , where  $\log^{(1)} n := \log n$ .

**Corollary 1.** *Let  $\Phi_n$  denote the internal path length, and let  $\Psi_n$  be the Wiener index of a random  $b$ -ary recursive tree of size  $n$  with edge weights  $Z$ , where  $\|Z\|$  is bounded almost surely. Then, there exists a constant  $D > 0$  such that, for  $t > 0$  and  $n \rightarrow \infty$ ,*

$$P(|\Phi_n - E[\Phi_n]| > t E[\Phi_n]) \leq \exp\left(-\frac{b}{b-1} \frac{\mu}{D} t \log n (\log^{(2)} n + \log t + \alpha + o(1))\right)$$

and

$$P(|\Psi_n - E[\Psi_n]| > t E[\Psi_n]) \leq \exp\left(-\frac{b}{b-1} \frac{\mu}{D} t \log n (\log^{(2)} n + \log t + \alpha + o(1))\right),$$

where  $\mu = E[Z_1]$  and  $\alpha := \log(b\mu/(4D(b-1)e))$ .

Finally, by special choices of the edge weights and the use of transfer results in [13], the corresponding bounds for random linear recursive trees are obtained. The model of linear

recursive trees was introduced in [14]. Starting with the root, the linear recursive tree grows node by node. In each step the new node is attached to a randomly chosen node of the previous steps. The probability that node  $u$  is chosen is proportional to the weight  $w_u = 1 + \beta \deg(u)$ , where  $\deg(u)$  is the number of children of  $u$  and  $\beta \in \mathbb{R}_{\geq 0}$  is the parameter of the tree. This tree model encompasses as special cases the random recursive tree ( $\beta = 0$ ) and the plane-oriented recursive tree ( $\beta = 1$ ).

**Corollary 2.** *Let  $\Phi_n$  denote the internal path length of a random linear recursive tree of size  $n$  with weight function  $u \mapsto 1 + (b - 2) \deg(u)$  for  $b \in \mathbb{N}$  and  $b \geq 2$ . Then there exists  $D > 0$  such that, for  $t > 0$  and  $n \rightarrow \infty$ , we have the same tail bounds as in Theorem 2 for  $(\Phi_n - E[\Phi_n])/n$  and, in particular, we have, for  $t > 0$  and  $n \rightarrow \infty$ ,*

$$P(|\Phi_n - E[\Phi_n]| > t E[\Phi_n]) \leq \exp\left(-\frac{1}{b-1} \frac{1}{D} t \log n (\log^{(2)} n + \log t + \alpha + o(1))\right)$$

with  $\alpha := -\log(4D(b-1)e)$ .

**Corollary 3.** *Let  $\Psi_n$  denote the Wiener index of a random linear recursive tree of size  $n$  with weight function  $u \mapsto 1 + (b - 2) \deg(u)$  for  $b \in \mathbb{N}$  and  $b \geq 2$ . Then there exists  $D > 0$  such that we have, for  $t > 0$  and  $n \rightarrow \infty$ ,*

$$P(|\Psi_n - E[\Psi_n]| > t E[\Psi_n]) \leq \exp\left(-\frac{1}{b-1} \frac{1}{D} t \log n (\log^{(2)} n + \log t + \alpha + o(1))\right)$$

with  $\alpha := -\log(4D(b-1)e)$ .

Using the Wentzel–Kramers–Brillouin method, Knessl and Szpankowski [9] obtained very sharp bounds for the tail of the limit distribution of the internal path length of random binary search trees. Rüschemdorf and Schopp [16] obtained general upper bounds for tails of distributions given by a recursion of sum type in the one-dimensional case. For simply generated trees, asymptotics for the right tail of the limit distribution of the total path length and the Wiener index are shown in [7] and [8].

This paper is organized as follows. In Section 2 we consider the general recursion formula (1) and give a proof of the upper tail bound given in Theorem 1. The  $b$ -ary recursive tree with weighted edges is defined in Section 3. We then show the stochastic domination condition in this case by a coupling argument, and provide proofs of Theorem 2 and Corollary 1 in Section 3.1. Finally, using the transfer results from [13], we obtain the upper tail bounds in the case of random linear recursive trees (Corollary 2 and Corollary 3) in Section 3.2. In Section 4 we give a summary of corresponding results concerning lower tail bounds for the Wiener index.

For functions  $f$  and  $g$ , we respectively write  $f = o(g)$ ,  $f = O(g)$ , and  $f = \Theta(g)$  if  $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$ ,  $|f(n)/g(n)| \leq C$ , and  $c \leq |f(n)/g(n)| \leq C$  for all  $n$  and some constants  $0 < c \leq C < \infty$ .

### 2. Upper tail bound for a general recursion

We consider a random  $k$ -dimensional vector  $X_n = (X_{n,1}, \dots, X_{n,k})$  which solves the distributional recursion formula

$$X_n \stackrel{D}{=} \sum_{i=1}^b A_i(I_n) X_{I_n,i}^{(i)} + d(I_n, Z),$$

where  $X_n^{(1)}, \dots, X_n^{(b)}$  have the same distribution as  $X_n$ ,  $d: \mathbb{R}^b \times \mathbb{R}^b \rightarrow \mathbb{R}^k$  and  $A_i: \mathbb{R}^b \rightarrow \mathbb{R}^{k \times k}$

are deterministic functions,  $Z \in \mathbb{R}_{\geq 0}^b$  and  $I_n = (I_{n,1}, \dots, I_{n,b}) \in \{0, \dots, n-1\}^b$  are random vectors with  $E[d(I_n, Z)] = 0$ , and  $(X_k^{(1)})_{k \geq 0}, \dots, (X_k^{(b)})_{k \geq 0}, I_n$ , and  $Z$  are independent.

**Lemma 1.** *Let  $X_n$  be a solution of the distributional recursion (1). Assume that  $X_1 = 0$ ,  $\|d(I_n, Z)\| \leq D$  almost surely for all  $n \in \mathbb{N}$  and a constant  $D \in \mathbb{R}$ , and that*

$$\sum_{i=1}^b \|A_i(I_n)\|_{\text{op}}^2 \leq_{\text{st}} 1 - U(1 - U) \tag{3}$$

as well as  $\|A_i\|_{\text{op}} \leq 1$ . Let  $\gamma \approx 2.0047$  be the positive solution of

$$\frac{12}{7} = e^{2/\gamma} - \frac{2}{\gamma} \quad \text{and} \quad K = \frac{5}{2} D^2 \gamma^2.$$

Then we have, for all  $s \in \mathbb{R}^k$  with  $\|s\| \leq 1/(\gamma D)$  and all  $n \in \mathbb{N}$ ,

$$E[\exp(\langle s, X_n \rangle)] \leq \exp(K \|s\|^2).$$

*Proof.* We show the claim by induction on  $n$ . For  $n = 1$ , we have  $X_1 = 0$  and there is nothing to show.

Condition (3) implies that there is a coupling of  $I_n$  with a uniformly distributed random variable  $U \in [0, 1]$  such that

$$\sum_{i=1}^b \|A_i(I_n)\|_{\text{op}}^2 \leq 1 - U(1 - U) \tag{4}$$

holds almost surely.

Using the recursion formula and the given independence, we obtain, for  $n \geq 2$ ,

$$\begin{aligned} & E[\exp(\langle s, X_n \rangle)] \\ &= E \left[ \exp \left( \left\langle s, \sum_{i=1}^b A_i(I_n) X_{I_n,i}^{(i)} + d(I_n, Z) \right\rangle \right) \right] \\ &= \sum_{x \in \{0, \dots, n-1\}^b} E[e^{\langle s, d(x, Z) \rangle}] \prod_{i=1}^b E[\exp(\langle (A_i(x))^{\top} s, X_{x_i}^{(i)} \rangle) \mid I_n = x] P(I_n = x). \end{aligned}$$

The assumption that  $\|A_i(x)\|_{\text{op}} \leq 1$  implies that  $\|A_i(x)^{\top} s\| \leq \|s\| \|A_i(x)\|_{\text{op}} \leq \|s\|$ . Since, for every  $i \in \{1, \dots, b\}$ , we have  $x_i \leq n-1$ , we can apply the induction hypothesis to obtain

$$\begin{aligned} E[\exp(\langle s, X_n \rangle)] &\leq \sum_{x \in \{0, \dots, n-1\}^b} E[\exp(\langle s, d(x, Z) \rangle)] \exp \left( \sum_{i=1}^b K \|s\|^2 \|A_i(x)\|_{\text{op}}^2 \right) P(I_n = x) \\ &= E \left[ \exp(\langle s, d(I_n, Z) \rangle) \exp \left( K \|s\|^2 \sum_{i=1}^b \|A_i(I_n)\|_{\text{op}}^2 \right) \right]. \end{aligned}$$

By condition (4) and the monotonicity of  $x \mapsto e^{\lambda x}$  for  $\lambda > 0$ , we conclude that

$$\begin{aligned} E[\exp(\langle s, X_n \rangle)] &\leq E[\exp(\langle s, d(I_n, Z) \rangle) \exp(K \|s\|^2 (1 - U(1 - U)))] \\ &= e^{K \|s\|^2} E[\exp(\langle s, d(I_n, Z) \rangle) \exp(-K \|s\|^2 U(1 - U))]. \end{aligned} \tag{5}$$

Hence, using the Cauchy–Schwarz inequality, it suffices to show that

$$\begin{aligned} & (\mathbb{E}[\exp(\langle s, d(I_n, Z) \rangle) \exp(-K \|s\|^2 U(1 - U))]^2 \\ & \leq \mathbb{E}[\exp(2\langle s, d(I_n, Z) \rangle)] \mathbb{E}[\exp(-2K \|s\|^2 U(1 - U))] \\ & \leq 1. \end{aligned} \tag{6}$$

By assumption,  $\|d(I_n, Z)\| \leq D$  holds almost surely and  $\mathbb{E}[d(I_n, Z)] = 0$ . Thus, we obtain, for  $\|s\| \leq 1/(\gamma D)$ ,

$$\begin{aligned} \mathbb{E}[\exp(2\langle s, d(I_n, Z) \rangle)] &= 1 + \mathbb{E}\left[\langle s, d(I_n, Z) \rangle^2 \sum_{k=2}^{\infty} \frac{2^k \langle s, d(I_n, Z) \rangle^{k-2}}{k!}\right] \\ &\leq 1 + \|s\|^2 D^2 \sum_{k=2}^{\infty} \frac{2^k}{k! \gamma^{k-2}} \\ &= 1 + \|s\|^2 D^2 \gamma^2 \left( e^{2/\gamma} - 1 - \frac{2}{\gamma} \right). \end{aligned} \tag{7}$$

For all  $x > 0$ , we have  $e^{-x} \leq 1 - x + \frac{1}{2}x^2$ . This yields

$$\begin{aligned} \mathbb{E}[\exp(-2K \|s\|^2 U(1 - U))] &\leq \mathbb{E}\left[1 - 2K \|s\|^2 U(1 - U) + 2K^2 \|s\|^4 U^2(1 - U)^2\right] \\ &= 1 - \frac{1}{3}K \|s\|^2 + \frac{1}{15}K^2 \|s\|^4. \end{aligned} \tag{8}$$

The estimates (7) and (8) imply that (6) follows from

$$\left(1 + \|s\|^2 D^2 \gamma^2 \left( e^{2/\gamma} - 1 - \frac{2}{\gamma} \right)\right) \left(1 - \frac{1}{3}K \|s\|^2 + \frac{1}{15}K^2 \|s\|^4\right) \leq 1.$$

This in turn is equivalent to  $f(\|s\|) \leq 0$  for

$$\begin{aligned} f(\|s\|) &:= D^2 \gamma^2 \left( e^{2/\gamma} - 1 - \frac{2}{\gamma} \right) - \frac{1}{3}K - \left( \frac{1}{3}K D^2 \gamma^2 \left( e^{2/\gamma} - 1 - \frac{2}{\gamma} \right) - \frac{1}{15}K^2 \right) \|s\|^2 \\ &\quad + \frac{1}{15}K^2 D^2 \gamma^2 \left( e^{2/\gamma} - 1 - \frac{2}{\gamma} \right) \|s\|^4. \end{aligned}$$

Substituting  $K = \frac{5}{2}D^2\gamma^2$  and  $e^{2/\gamma} - 1 - 2/\gamma = \frac{5}{7}$  we obtain

$$f(\|s\|) = D^2 \gamma^2 \left( \frac{5}{7} - \frac{5}{6} + \left( \frac{5}{12} - \frac{25}{42} \right) (D\gamma \|s\|)^2 + \frac{25}{84} (D\gamma \|s\|)^4 \right).$$

We see that  $f(0) \leq 0$  and  $f(1/(\gamma D)) = 0$ . Since  $f$  is a biquadratic function in  $\|s\|$  with a positive coefficient corresponding to  $\|s\|^4$  and  $f(0) \leq 0$ , it has at most two real roots. On the interval between these two roots the function is negative and outside this interval the function takes only positive values. Since  $f(1/(\gamma D)) = 0$ , we therefore obtain  $f(\|s\|) \leq 0$  for all  $s$  with  $0 \leq \|s\| \leq 1/(\gamma D)$ .

**Lemma 2.** Let  $X_n$  be a solution of the distributional recursion (1). Assume that  $X_1 = 0$ ,  $\|d(I_n, Z)\| \leq D$  almost surely for all  $n \in \mathbb{N}$  and a constant  $D \in \mathbb{R}$ , and that

$$\sum_{i=1}^b \|A_i(I_n)\|_{\text{op}}^2 \leq_{\text{st}} 1 - U(1 - U)$$

as well as  $\|A_i\|_{\text{op}} \leq 1$ . Let  $\gamma \approx 2.0047$  be the positive solution of  $\frac{12}{\gamma} = e^{2/\gamma} - 2/\gamma$ , and let  $L_0 \approx 5.0177$  be the largest root of  $e^L = 6L^2$ . Then we have, for  $1/(\gamma D) \leq \|s\| \leq L$ ,

$$E[\exp(\langle s, X_n \rangle)] \leq \exp(K_L \|s\|^2),$$

where

$$K_L := \begin{cases} 24D^2 & \text{for } 1/(\gamma D) < L \leq L_0/D, \\ 4\frac{1}{L^2}e^{LD} & \text{for } L_0/D < L. \end{cases}$$

*Proof.* We again use induction on  $n$ . For  $n = 1$ , there is nothing to show. We use the same arguments as in the beginning of the proof of Lemma 1 and obtain (5), i.e.

$$E[\exp(\langle s, X_n \rangle)] \leq e^{K_L \|s\|^2} E[\exp(\langle s, d(I_n, Z) \rangle) \exp(-K_L \|s\|^2 U(1 - U))]$$

for a random variable  $U$  which is uniformly distributed on  $[0, 1]$  and coupled with  $I_n$  such that (4) holds. Hence, it suffices to prove (6) under the new assumptions. Since  $\|d(I_n, Z)\| \leq D$  almost surely, the proof is completed by showing that

$$e^{D\|s\|} E[\exp(-K_L \|s\|^2 U(1 - U))] \leq 1.$$

Fill and Janson [6, Section 4] proved that, for any  $K > 0$ ,

$$E[\exp(-2K \|s\|^2 U(1 - U))] \leq \frac{1 - \exp(-K \|s\|^2/2)}{K \|s\|^2/2}$$

and, for  $0.42 \leq |\lambda| \leq M$ ,

$$e^{|\lambda|} \frac{1 - \exp(-K_M \lambda^2/2)}{K_M \lambda^2/2} \leq 1 \quad \text{when} \quad K_M = \begin{cases} 12 & \text{for } M \leq L_0, \\ \frac{2e^M}{M^2} & \text{for } L_0 < M. \end{cases}$$

In the present situation, it follows that

$$e^{D\|s\|} E[\exp(-K_L \|s\|^2 U(1 - U))] \leq e^{D\|s\|} \left( 1 - \exp\left(-\frac{K_L D^2 \|s\|^2}{2D^2} \frac{D^2 \|s\|^2}{2}\right) \right) / \left( \frac{K_L D^2 \|s\|^2}{2D^2} \frac{D^2 \|s\|^2}{2} \right) \leq 1$$

when  $1/\gamma \leq D\|s\| \leq LD$  and

$$K_L = \begin{cases} 24D^2 & \text{for } L \leq L_0/D, \\ 4\frac{1}{L^2}e^{LD} & \text{for } L_0/D < L. \end{cases}$$

Thus, we obtain the claim because  $1/\gamma \geq 0.42$ .

We summarize the results of the two preceding lemmas.

**Corollary 4.** *Let  $X_n$  be a solution of the distributional recursion (1). Assume that  $X_1 = 0$ ,  $\|d(I_n, Z)\| \leq D$  almost surely for all  $n \in \mathbb{N}$  and a constant  $D \in \mathbb{R}$ , and that*

$$\sum_{i=1}^b \|A_i(I_n)\|_{\text{op}}^2 \leq_{\text{st}} 1 - U(1 - U)$$

as well as  $\|A_i\|_{\text{op}} \leq 1$ . Let  $\gamma \approx 2.0047$  be the positive solution of  $\frac{12}{7} = e^{2/\gamma} - 2/\gamma$ , and let  $L_0 \approx 5.0177$  be the largest root of  $e^L = 6L^2$ . Then we have, for every  $s$  and  $n \geq 1$ ,

$$E[\exp(\langle s, X_n \rangle)] \leq \begin{cases} \exp(\frac{5}{2}\gamma^2 D^2 \|s\|^2) & \text{for } 0 \leq \|s\| \leq 1/(\gamma D), \\ \exp(24D^2 \|s\|^2) & \text{for } 1/(\gamma D) < \|s\| \leq L_0/D, \\ \exp(4e^D \|s\|) & \text{for } L_0/D < \|s\|. \end{cases}$$

*Proof.* The bounds for  $\|s\| \leq L_0/D$  follow immediately from Lemma 1 and Lemma 2. Since the function  $x \mapsto e^{Dx}/x^2$  is monotonically increasing on the interval  $[L_0/D, \infty)$ , Lemma 2 also yields the bound in the  $\|s\| > L_0/D$  case.

We now obtain the tail bound for any entry of the vector  $X_n$ .

*Proof of Theorem 1.* We denote by  $e_j$  the vector with a 1 in the  $j$ th position and 0s elsewhere. We use Chernoff’s bounding technique and obtain, for  $u > 0$  and  $j \in \{1, \dots, k\}$  with Corollary 4,

$$\begin{aligned} P(X_{n,j} > t) &= P(\exp(uX_{n,j}) > \exp(ut)) \\ &\leq E[\exp(uX_{n,j} - ut)] \\ &= E[\exp(u\langle e_j, X_n \rangle - ut)] \\ &\leq \exp(K_u u^2 - ut), \end{aligned}$$

where

$$K_u = \begin{cases} \frac{5}{2}\gamma^2 D^2 & \text{for } 0 \leq u \leq 1/(\gamma D), \\ 24D^2 & \text{for } 1/(\gamma D) < u \leq L_0/D, \\ 4\frac{e^{Du}}{u^2} & \text{for } L_0/D < u. \end{cases}$$

For the left tail, we analogously obtain

$$\begin{aligned} P(X_{n,j} < -t) &= P(\exp(uX_{n,j}) < \exp(-ut)) \\ &\leq E[\exp(-uX_{n,j} - ut)] \\ &= E[\exp(-u\langle e_j, X_n \rangle - ut)] \\ &\leq \exp(K_u u^2 - ut). \end{aligned}$$

In order to minimize this bound, we look for the minimum of the function  $f(u) := K_u u^2 - ut$ . This function takes its minimum at  $\check{u}_i(t)$  and has the value  $f(\check{u}_i(t))$ :

$$\begin{aligned} \check{u}_1(t) &= \frac{t}{5D^2\gamma^2}, & f(\check{u}_1(t)) &= -\frac{t^2}{10\gamma^2 D^2}, & \text{for } K_u &= \frac{5}{2}D^2\gamma^2, \\ \check{u}_2(t) &= \frac{t}{48D^2}, & f(\check{u}_2(t)) &= -\frac{t^2}{96D^2}, & \text{for } K_u &= 24D^2, \\ \check{u}_3(t) &= \frac{1}{D} \log \frac{t}{4D}, & f(\check{u}_3(t)) &= \frac{t}{D} - \frac{t}{D} \log \frac{t}{4D}, & \text{for } K_u &= 4\frac{e^{Du}}{u^2}. \end{aligned}$$

Here  $\check{u}_i(t) \in U_i$  with  $U_1 := [0, 1/(\gamma D)]$ ,  $U_2 := (1/(\gamma D), L_0/D]$ , and  $U_3 := (L_0/D, \infty)$ .

If  $\check{u}_i(t) \notin U_i$  for a given  $t$ , we can take  $u$  at the proper boundary of  $U_i$ .



Comparing the different values of the minimum for  $i = 1, 2, 3$  we obtain the total minimum. For  $t \in [0, 5\gamma D]$ , we have the following possibilities:

$$\begin{aligned} \check{u}_1(t) &= \frac{t}{5D^2\gamma^2}, & f(\check{u}_1(t)) &= -\frac{t^2}{10\gamma^2D^2}, \\ u_2 &= \frac{1}{\gamma D}, & f(u_2) &= \frac{24}{\gamma^2} - \frac{t}{\gamma D}, \\ u_3 &= \frac{L_0}{D}, & f(u_3) &= 4e^{L_0} - \frac{L_0}{D}t. \end{aligned}$$

The minimum is given for  $\check{u}_1(t)$ .

Similarly, we obtain the minimum in the other cases by making the following choices. For  $t \in [5\gamma D, 48D/\gamma + D\sqrt{48\sqrt{48/\gamma^2 - 5}}]$ ,

$$u_1 = \frac{1}{\gamma D}, \quad K_u = \frac{5}{2}\gamma^2D^2, \quad f(u_1) = \frac{5}{2} - \frac{t}{\gamma D};$$

for  $t \in [48D/\gamma + D\sqrt{48\sqrt{48/\gamma^2 - 5}}, 48DL_0]$ ,

$$\check{u}_2(t) = \frac{t}{48D^2}, \quad K_u = 24D^2, \quad f(\check{u}_2(t)) = -\frac{t^2}{96D^2};$$

for  $t \in [48DL_0, 4De^{L_0}]$ ,

$$u_2 = \frac{L_0}{D}, \quad K_u = 24D^2, \quad f(u_2) = 24L_0^2 - \frac{L_0}{D}t;$$

and, for  $t \in [4De^{L_0}, \infty)$ ,

$$\check{u}_3(t) = \frac{1}{D} \log \frac{t}{4D}, \quad K_{\check{u}_3(t)} = \frac{4e^{D\check{u}_3(t)}}{\check{u}_3(t)^2}, \quad f(\check{u}_3(t)) = \frac{t}{D} - \frac{t}{D} \log \frac{t}{4D}.$$

### 3. Applications to random trees

An example of a vector which satisfies the recursion formula (1) is the vector consisting of the internal path length and the Wiener index of a random tree in which all subtrees are (conditioned upon their sizes) an independent copy of the whole tree.

The internal path length of a rooted tree is the sum of all node depths of the tree. The depth of a node is given by the number of edges on the path from the node to the root. Analogously, the Wiener index is the sum of the distances between all unordered pairs of nodes where the distance is given by the number of edges on the unique path between the two nodes. If the edges of the tree are endowed with weights, the distance between two nodes and the depth of a node are given by the sum of the weights along the corresponding path.

#### 3.1. The random $b$ -ary recursive tree with weighted edges

In this subsection we consider the special case of a random  $b$ -ary recursive tree with weighted edges. We recall the recursive definition of this tree model given in the introduction. Considering the infinite complete  $b$ -ary rooted, ordered tree, we start with the root as the first internal node and its  $b$  children as external nodes. Given the random  $b$ -ary recursive tree with  $n - 1$  internal nodes, the  $n$ th internal node is added in the following way. We choose a random

node uniformly distributed on the set of all current external nodes, change it to an internal node, and add the  $b$  children of this new node to the set of external nodes. Finally, the nodes are labelled in the order of their appearance.

The edge weights are given by the entries of independent copies of a random vector  $Z := (Z_1, \dots, Z_b) \in \mathbb{R}_{\geq 0}^b$  attached to every node  $u$  of the complete infinite  $b$ -ary tree.

While the entries of the vector  $Z$  may depend on each other, we assume that they are identically distributed, i.e. for all  $i, j \in \{1, \dots, b\}$ , we have  $Z_i \stackrel{D}{=} Z_j$ , denoting its expectation by  $\mu := E[Z_1]$ . This assumption is not restrictive for the intended limit theorems as can be seen by a permutation argument (see [12, pp. 14–15]).

The shape of the random binary search tree has the same distribution as the shape of the random  $b$ -ary recursive tree with edge weights  $(Z_1, Z_2) = (1, 1)$  for  $b = 2$ .

Let  $Y_n = (\Psi_n, \Phi_n)$  denote the vector consisting of the Wiener index and the internal path length of the random  $b$ -ary recursive tree of size  $n$  with edge weights  $Z$ . In [13] it was shown that the vector

$$X_n := \begin{bmatrix} \frac{1}{n^2} & 0 \\ 0 & \frac{1}{n} \end{bmatrix} (Y_n - E[Y_n]) \tag{9}$$

satisfies the recursion formula (1), where the matrices  $A_i(I_n)$  are given by

$$A_i(I_n) = \begin{bmatrix} \frac{I_{n,i}^2}{n^2} & \frac{I_{n,i}(n - I_{n,i})}{n^2} \\ 0 & \frac{I_{n,i}}{n} \end{bmatrix},$$

and the vector  $d(I_n, Z)$  is given by

$$d_1^{(n)} = \frac{b}{b-1} \mu \sum_{i=1}^b \frac{I_{n,i}}{n} \log \frac{I_{n,i}}{n} + \sum_{i \neq j} \left( \frac{1}{2}(Z_i + Z_j) + \frac{b}{b-1} \mu \right) \frac{I_{n,i}}{n} \frac{I_{n,j}}{n} + o(1) \tag{10}$$

and

$$d_2^{(n)} = \frac{b}{b-1} \mu \sum_{i=1}^b \frac{I_{n,i}}{n} \log \frac{I_{n,i}}{n} + \sum_{i=1}^b Z_i \frac{I_{n,i}}{n} + o(1). \tag{11}$$

To apply the result of the previous section, we have to prove the stochastic domination condition for the  $b$ -ary recursive tree.

3.1.1. *Coupling.* For  $(x_1, \dots, x_b) \in \mathbb{R}^b$ , we denote by  $(x_{(1)}, \dots, x_{(b)})$  the order statistic, i.e.

$$x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(b)};$$

the entries of  $(x_1, \dots, x_b)$  and  $(x_{(1)}, \dots, x_{(b)})$  are the same. We consider the space  $\mathbb{R}^b$  with the partial order given by

$$(x_1, \dots, x_b) \leq (y_1, \dots, y_b) \iff x_i \leq y_i \text{ for all } i \in \{1, \dots, b\},$$

and define  $\mathcal{E}_b := \{(x_1, \dots, x_b) \in \mathbb{R}_{\geq 0}^b \mid x_1 \geq x_2 \geq \dots \geq x_b\}$ . Moreover, we denote by  $\text{PU}(b)$  a Pólya urn with balls of  $b$  different colours  $\{1, \dots, b\}$ . The urn initially contains one ball of

each colour. After a ball of colour  $j$  is drawn, it is returned to the urn together with another  $b - 1$  balls of the same colour.

Considering the evolution process which yields the random  $b$ -ary recursive tree, it is not difficult to see that the vector of the sizes of the subtrees has the same distribution as the vector of the numbers of drawings of a ball of each colour in the urn described above (for more details, see [12, Section 2.2]). Using this, the next two lemmas provide the estimate we need.

**Lemma 3.** *For  $j \in \{1, \dots, b\}$ , let  $J_{n,j}$  denote the number of times that the drawn ball is of colour  $j$  during the first  $n$  drawings from the Pólya urn  $\text{PU}(b)$ , and let  $I_{n,j}$  be the corresponding size of the Pólya urn  $\text{PU}(b + 1)$ . Then we have, for the vectors  $J_n := (J_{n,1}, \dots, J_{n,b})$  and  $I_n := (I_{n,1}, \dots, I_{n,b+1})$ ,*

$$(I_{n,(1)}, \dots, I_{n,(b)}) \leq_{\text{st}} (J_{n,(1)}, \dots, J_{n,(b)}).$$

We prove this lemma by using a result about stochastic domination between Markov chains (see [10, Section IV.5, Theorem 5.8]). Recall that  $e_j \in \mathbb{R}^b$  denotes the vector with a 1 in the  $j$ th position and 0s elsewhere.

*Proof of Lemma 3.* It suffices to show that there is a coupling of  $I'_n := (I_{n,(1)}, \dots, I_{n,(b)})$  and  $J'_n := (J_{n,(1)}, \dots, J_{n,(b)})$  such that  $I'_n \leq J'_n$  almost surely.

The sequences  $J'_n$  and  $I'_n$  are Markov chains. To write down the transition probabilities, we define  $\alpha_j : \mathcal{E}_b \rightarrow \mathbb{N}_0$  by

$$\alpha_j(x_1, \dots, x_b) := \begin{cases} |\{i \mid x_j = x_i\}| & \text{if } x_{j-1} > x_j, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, the transition probability for  $J'_n$  is given by the kernel  $K_n : \mathcal{E}_b \times \mathcal{E}_b \rightarrow [0, 1]$  with

$$K_n(x, x + e_j) := P(J'_{n+1} = x + e_j \mid J'_n = x) = \frac{1 + x_j(b - 1)}{b + n(b - 1)} \alpha_j(x)$$

for  $x = (x_1, \dots, x_b) \in \mathcal{E}_b$  and  $j = 1, \dots, b$ . For the transition probability of  $I'_n$ , we have the kernel  $K'_n : \mathcal{E}_b \times \mathcal{E}_b \rightarrow [0, 1]$  with

$$K'_n(x, x + e_j) := P(I'_{n+1} = x + e_j \mid I'_n = x) = \frac{1 + x_j b}{b + 1 + nb} \alpha_j(x)$$

for  $j = 1, \dots, b$  and

$$K'_n(x, x) := P(I'_{n+1} = x \mid I'_n = x) = \frac{1 + (n - \sum_{i=1}^b x_i)b}{b + 1 + nb}.$$

Let  $x, y \in \mathcal{E}_b$  with  $y \leq x$ . We claim that  $K'_n(y, \cdot)$  is stochastically dominated by  $K_n(x, \cdot)$ . If

$$P(I'_{n+1} = y + e_j \mid I'_n = y) > 0,$$

we have  $\alpha_j(y) \neq 0$ . For  $y_j < x_j$ , we obtain  $y + e_j \leq x$ . Thus, we only have to consider the case where  $\alpha_j(y) \neq 0$  and  $y_j = x_j$ . Let  $j_1, \dots, j_m$  be the components for which  $\alpha_{j_l}(y) \neq 0$  and  $x_{j_l} = y_{j_l}$  for  $1 \leq l \leq m$ . Then we have  $\alpha_{j_l}(x) \geq \alpha_{j_l}(y)$  because  $x_{j_l-1} \geq y_{j_l-1} > y_{j_l} = x_{j_l}$ . Since  $x_{j_l} \leq n$ , we obtain

$$\frac{1 + x_{j_l}(b - 1)}{b + n(b - 1)} \geq \frac{1 + x_{j_l} b}{b + 1 + nb}.$$

This yields, for all  $l \in \{1, \dots, m\}$ ,

$$K_n(x, x + e_{j_l}) = \frac{1 + x_{j_l}(b - 1)}{b + n(b - 1)} \alpha_{j_l}(x) \geq \frac{1 + x_{j_l}b}{b + 1 + nb} \alpha_{j_l}(y) = K'_n(y, y + e_{j_l}).$$

For  $i \in \{1, \dots, b\} \setminus \{j_1, \dots, j_m\}$ , we obviously have

$$K_n(x, x + e_i) \geq 0 = K'_n(y, y + e_i).$$

Hence, we can find a coupling  $(\tilde{J}_{n+1}, \tilde{I}_{n+1})$  of  $J'_{n+1}$  and  $I'_{n+1}$  with

$$P(\tilde{I}_{n+1} \leq \tilde{J}_{n+1} \mid (\tilde{J}_n, \tilde{I}_n) = (x, y)) = 1$$

for all  $x, y \in \mathcal{E}_b$  with  $y \leq x$ . This implies that  $K_n(x, \cdot)$  stochastically dominates  $K'_n(y, \cdot)$ . Because of the Markov property we conclude by Theorem 5.8 of [10, Section IV.5] that there exists a coupling  $(\tilde{J}, \tilde{I})$  of  $J'$  and  $I'$  such that  $\tilde{I}_n \leq \tilde{J}_n$  almost surely for all  $n \in \mathbb{N}$ .

**Lemma 4.** *Let  $f : [0, 1] \rightarrow \mathbb{R}$  be the function given by*

$$f(x) = x^4 + (x^2 - x^3)(1 + \sqrt{x^2 + 1}).$$

*Then, for  $(x_1, \dots, x_b) \in \mathcal{E}_b$  and  $(y_1, \dots, y_{b+1}) \in \mathcal{E}_{b+1}$  with*

$$\sum_{i=1}^b x_i = \sum_{j=1}^{b+1} y_j = 1$$

*and  $(y_1, \dots, y_b) \leq (x_1, \dots, x_b)$ , we have*

$$\sum_{i=1}^{b+1} f(y_i) \leq \sum_{i=1}^b f(x_i).$$

*Proof.* We first show that the function  $f$  is convex. To do this, we derive the second derivative, which is given by

$$\begin{aligned} f''(x) &= 12x^2 + (2 - 6x)(1 + \sqrt{x^2 + 1}) + \frac{2(2x^2 - 3x^3)}{\sqrt{x^2 + 1}} + \frac{x^2 - x^3}{(x^2 + 1)^{3/2}} \\ &\geq 10x^2 + (2 - 6x)(1 + \sqrt{x^2 + 1}) + \frac{2(3x^2 - 3x^3)}{\sqrt{x^2 + 1}} + \frac{x^2 - x^3}{(x^2 + 1)^{3/2}}. \end{aligned}$$

To show convexity, it suffices to show that  $f'' \geq 0$ . Since, for  $x \in [0, 1]$ ,  $x^2 \geq x^3$ , it remains to show that

$$g(x) := 10x^2 + (2 - 6x)(1 + \sqrt{x^2 + 1}) \geq 0.$$

By consideration of the first and second derivatives we see that the minimum of  $g$  is obtained for  $x = \frac{3}{4}$ :  $g(\frac{3}{4}) = 0$ . Taking everything into account, we obtain  $f''(x) \geq 0$  for all  $x \in [0, 1]$ , which implies that the first derivative  $f'$  is monotone increasing.

By assumption, there exist numbers  $\alpha_1, \dots, \alpha_b \geq 0$  with  $x_i = y_i + \alpha_i y_{b+1}$  for  $i = 1, \dots, b$  and  $\sum_{i=1}^b \alpha_i = 1$ . The monotonicity of  $f'$  and  $f(0) = 0$  imply that, using the mean value

theorem,

$$\frac{f(x_i) - f(y_i)}{\alpha_i y_{b+1}} = f'(\xi) \geq f'(\eta) = \frac{f(y_{b+1})}{y_{b+1}}$$

for some  $\xi \in [y_i, x_i]$  and  $\eta \in [0, y_{b+1}] \subset [0, y_i]$ . This finally yields

$$\sum_{i=1}^b f(x_i) \geq \sum_{i=1}^b (f(y_i) + \alpha_i f(y_{b+1})) = \sum_{i=1}^{b+1} f(y_i).$$

*Proof of Theorem 2.* As seen in (9), we have for the vector  $X_n$  the recursion formula (1), where

$$A_i(I_n) = \begin{bmatrix} \frac{I_{n,i}^2}{n^2} & \frac{I_{n,i}(n - I_{n,i})}{n^2} \\ 0 & \frac{I_{n,i}}{n} \end{bmatrix}.$$

For the operator norm, we obtain

$$\|A_i(I_n)\|_{\text{op}}^2 = \|A_i(I_n)^\top A_i(I_n)\|_{\text{op}}.$$

The matrix  $A_i(I_n)^\top A_i(I_n)$  is symmetric. Thus, its operator norm is given by the largest absolute eigenvalue. Solving the characteristic equation for the matrix we find that its eigenvalue being larger in absolute value is given by

$$\frac{I_{n,i}^2}{n^2} \left( 1 - \frac{I_{n,i}}{n} + \frac{I_{n,i}^2}{n^2} + \left( 1 - \frac{I_{n,i}}{n} \right) \sqrt{\frac{I_{n,i}^2}{n^2} + 1} \right) = f\left(\frac{I_{n,i}}{n}\right),$$

with the function  $f$  as given in Lemma 4. The relation between the sizes of the subtrees and the number of drawings in a Pólya urn gives, by repeated application of Lemma 3 and Lemma 4,

$$\sum_{i=1}^b \|A_i(I_n)\|_{\text{op}}^2 \leq_{\text{st}} \|A_1(J_n)\|_{\text{op}}^2 + \|A_2(J_n)\|_{\text{op}}^2,$$

where  $J_n = (J_{n,1}, J_{n,2})$  is the vector of the sizes of the subtrees of a random binary search tree, i.e.  $J_{n,1}$  is uniformly distributed on  $\{0, \dots, n - 1\}$  and  $J_{n,2} = n - 1 - J_{n,1}$ . By Lemma 2.2 of [2] we obtain the stochastic domination condition

$$\sum_{i=1}^b \|A_i(I_n)\|_{\text{op}}^2 \leq_{\text{st}} 1 - U(1 - U).$$

Considering the toll vector  $d(I_n, Z)$  in (10) and (11), the boundedness of  $\|Z\|$  implies that its norm is bounded almost surely by some constant  $D$ . Furthermore, we trivially have  $X_1 = 0$  since the tree with one node is only the root. The claim follows by Theorem 1.

In [13] the asymptotic expansion of the expectation of  $\Phi_n$  and  $\Psi_n$  is determined. Using these results, we obtain asymptotic tail bounds.

*Proof of Corollary 1.* With

$$y_n = \frac{t E[\Phi_n]}{n} = \frac{b}{b - 1} \mu t \log n + O(1)$$

we obtain, by Theorem 2 and because  $\lim_{n \rightarrow \infty} y_n = \infty$ ,

$$\begin{aligned} &P(|\Phi_n - E[\Phi_n]| \geq t E[\Phi_n]) \\ &= P\left(\frac{|\Phi_n - E[\Phi_n]|}{n} \geq y_n\right) \\ &\leq \exp\left(\left(\frac{b}{b-1} \frac{\mu}{D} t \log n + O(1)\right) \left(1 - \log \frac{tb\mu \log n}{4D(b-1)}\right)\right) \\ &= \exp\left(-\frac{b}{b-1} \frac{\mu}{D} t \log n (\log^{(2)} n + \log t + \alpha + o(1))\right). \end{aligned}$$

With  $z_n = t E[\Psi_n]/n^2 = y_n + O(1)$ , the claim for  $\Psi_n$  follows as well.

**3.2. Random linear recursive trees**

In this subsection we transfer the results for the random  $b$ -ary recursive tree with weighted edges to linear recursive trees. In this tree, every node  $u$  has a weight  $w_u$ . Starting with the root, the tree grows node by node. In each step the new node is attached to a randomly chosen node of the previous steps. The probability that node  $u$  is chosen is proportional to the weight  $w_u$  of the node. In the case of linear recursive trees the weight is given by  $w_u = 1 + \beta \deg(u)$ , where  $\deg(u)$  is the number of children of  $u$  and  $\beta \in \mathbb{R}_{\geq 0}$  is the parameter of the tree.

Given a random linear recursive tree  $T_n$  of size  $n$  with weight function  $u \mapsto 1 + (b-2) \deg(u)$ , we consider a  $b$ -ary recursive tree  $\tilde{T}_{n-1}$  of size  $n - 1$  where the edges are weighted by the random vector  $Z$  which is obtained by a uniformly distributed permutation of the entries of  $(1, 0, \dots, 0) \in \mathbb{R}^b$ . In particular, we have  $\mu = 1/b$ . Denote by  $\tilde{\Phi}_{n-1}$ ,  $\tilde{\Psi}_{n-1}$ , and  $\Phi_n, \Psi_n$  the internal path lengths and the Wiener indices of  $\tilde{T}_{n-1}$  and  $T_n$ , respectively.

*Proof of Corollary 2.* With the notation above, it was shown in [13] that

$$\Phi_n \stackrel{D}{=} \tilde{\Phi}_{n-1} + n - 1$$

holds. Therefore, we have

$$\frac{\Phi_n - E[\Phi_n]}{n} \stackrel{D}{=} \frac{\tilde{\Phi}_{n-1} - E[\tilde{\Phi}_{n-1}]}{n}.$$

The claim follows immediately by Theorem 2 and Corollary 1.

*Proof of Corollary 3.* With the notation above, it was shown in [13] that

$$\Psi_n \stackrel{D}{=} \tilde{\Psi}_{n-1} - \tilde{\Phi}_{n-1} + (n - 1)^2$$

holds. This yields

$$P(|\Psi_n - E[\Psi_n]| > t E[\Psi_n]) = P(|\tilde{\Psi}_{n-1} - E[\tilde{\Psi}_{n-1}] - \tilde{\Phi}_{n-1} + E[\tilde{\Phi}_{n-1}]| > t E[\Psi_{n-1}]).$$

Moreover, in [13] it was shown that  $\text{var}(\tilde{\Phi}_{n-1}) = \Theta(n^2)$  for  $n \rightarrow \infty$ . Applying Chebycheff's inequality, we obtain, for any  $\varepsilon > 0$ ,

$$P\left(\frac{|\tilde{\Phi}_{k-1,n} - E[\tilde{\Phi}_{n-1}]|}{n^2} > \varepsilon\right) \leq O\left(\frac{1}{n^2}\right).$$

Since  $E[\tilde{\Psi}_{n-1}] = E[\Psi_n] + O(n \log n) = \Theta(n^2 \log n)$ , this yields, using Corollary 1,

$$\begin{aligned} P(|\Psi_n - E[\Psi_n]| > t E[\Psi_n]) &\leq P\left(\frac{|\tilde{\Psi}_{n-1} - E[\tilde{\Psi}_{n-1}]|}{n^2} + \varepsilon > \frac{t E[\Psi_n]}{n^2}\right) + O\left(\frac{1}{n^2}\right) \\ &\leq \exp\left(-\frac{1}{b-1} \frac{1}{D} t \log n (\log^{(2)} n + \log t + \alpha + o(1))\right). \end{aligned}$$

**Remark 1.** The random plane-oriented recursive tree without a node order is equal in distribution to the random linear recursive tree with parameter  $\beta = 1$ . Since the internal path length as well as the Wiener index are invariant under a change in the order of the tree, the tail bounds in Corollary 2 and Corollary 3 with  $b = 3$  provide in particular the corresponding tail bounds for the plane-oriented recursive tree.

#### 4. Lower tail bounds for the Wiener index

McDiarmid and Hayward [11] proved a lower bound for the tail of the distribution of the number of comparisons made by Quicksort. They constructed a set of binary trees that had high probability and implied a large number of comparisons. They succeeded in finding lower and upper bounds which have the same asymptotical behaviour. This idea was employed by Ali Khan and Neiningger [2] to prove a lower tail bound for the Wiener index of binary search trees.

In Section 7.2 of [12] the construction from [2] is extended to random  $b$ -ary recursive trees with weighted edges where at least one entry of  $Z$  is 1. This yields the following lower bound on the tail of the distribution of the Wiener index.

**Theorem 3.** *Let  $\Psi_n$  denote the Wiener index of a random  $b$ -ary recursive tree of size  $n$  with edge weights  $Z$ , where  $\{Z_1, \dots, Z_b\} \cap \{1\} \neq \emptyset$  and  $Z_i \geq 0$  hold almost surely. Then we have, for fixed  $t > 0$  and  $n \rightarrow \infty$ ,*

$$P(|\Psi_n - E[\Psi_n]| > t E[\Psi_n]) \geq \exp\left(-4 \frac{b}{b-1} \mu t \log n (\log^{(2)} n + O(\log^{(3)} n))\right).$$

With the transfer results already used in Section 3.2 we obtain a lower bound for the distribution of the Wiener index of random linear recursive trees.

**Theorem 4.** *Let  $\Psi_n$  denote the Wiener index of a random linear recursive tree of size  $n$  with weight function  $u \mapsto 1 + (b - 2) \deg(u)$  for  $b \in \mathbb{N}$  and  $b \geq 2$ . Then we have, for fixed  $t > 0$  and  $n \rightarrow \infty$ ,*

$$P(|\Psi_{n+1} - E[\Psi_{n+1}]| > t E[\Psi_{n+1}]) \geq \exp\left(-4 \frac{1}{b-1} t \log n (\log^{(2)} n + O(\log^{(3)} n))\right).$$

Remark 1 also holds for the lower tail bound.

**Remark 2.** The constants  $D$  arising in the results depend on the specific toll function which in turn depends on the functional and the tree model considered. Since the toll function in (10) and (11) is only known up to an  $o(1)$  term, it is in general not possible to determine this constant. Nevertheless, it is an analytical problem and should be solvable for special functionals and tree models. For instance, in the case of the vector  $(\Psi_n, \Phi_n)$  of the binary search tree, Ali Khan and Neiningger [2] showed that  $D \leq 1$ .

### Acknowledgement

The author thanks Ralph Neininger for posing the problem of tail bounds and highlighting some related literature.

### References

- [1] ALI KHAN, T. AND NEININGER, R. (2004). Probabilistic analysis for randomized game tree evaluation. In *Mathematics and Computer Science. III*, Birkhäuser, Basel, pp. 163–174.
- [2] ALI KHAN, T. AND NEININGER, R. (2007). Tail bounds for the Wiener index of random trees. In *2007 Conf. Analysis of Algorithms (AofA '07 Discrete Math. Theoret. Comput. Sci. Proc. AH)*, Assoc. Discrete Math. Theoret. Comput. Sci., Nancy, pp. 279–289.
- [3] BERGERON, F., FLAJOLET, P. AND SALVY, B. (1992). Varieties of increasing trees. In *CAAP '92 (Rennes, 1992; Lecture Notes Comput. Sci. 581)*, ed. J.-C. Raoult, Springer, Berlin, pp. 24–48.
- [4] BROUTIN, N. AND DEVROYE, L. (2006). Large deviations for the weighted height of an extended class of trees. *Algorithmica* **46**, 271–297.
- [5] BROUTIN, N., DEVROYE, L., MCLEISH, E. AND DE LA SALLE, M. (2008). The height of increasing trees. *Random Structures Algorithms* **32**, 494–518.
- [6] FILL, J. A. AND JANSON, S. (2001). Approximating the limiting Quicksort distribution. *Random Structures Algorithms* **19**, 376–406.
- [7] FILL, J. A. AND JANSON, S. (2009). Precise logarithmic asymptotics for the right tails of some limit random variables for random trees. *Ann. Combinatorics* **12**, 403–416.
- [8] JANSON, S. AND CHASSAING, P. (2004). The center of mass of the ISE and the Wiener index of trees. *Electron. Commun. Probab.* **9**, 178–187.
- [9] KNESSL, C. AND SZPANKOWSKI, W. (1999). Quicksort algorithm again revisited. *Discrete Math. Theoret. Comput. Sci.* **3**, 43–64.
- [10] LINDVALL, T. (1992). *Lectures on the Coupling Method*. John Wiley, New York.
- [11] MCDIARMID, C. J. H. AND HAYWARD, R. B. (1996). Large deviations for Quicksort. *J. Algorithms* **21**, 476–507.
- [12] MUNSONIUS, G. O. (2010). Limit theorems for functionals of recursive trees. Doctoral Thesis, University of Freiburg, Germany. Available at <http://www.freidok.uni-freiburg.de/volltexte/7472/>.
- [13] MUNSONIUS, G. O. (2010). The total Steiner  $k$ -distance for  $b$ -ary recursive trees and linear recursive trees. In *21st Internat. Conf. Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA '10 Discrete Math. Theoret. Comput. Sci. Proc. AH)*, Assoc. Discrete Math. Theoret. Comput. Sci., Nancy, pp. 527–548.
- [14] PITTEL, B. (1994). Note on the heights of random recursive trees and random  $m$ -ary search trees. *Random Structures Algorithms* **5**, 337–347.
- [15] RÖSLER, U. (1991). A limit theorem for “Quicksort”. *RAIRO Inf. Théor. Appl.* **25**, 85–100.
- [16] RÜSCHENDORF, L. AND SCHOPP, E.-M. (2007). Exponential bounds and tails for additive random recursive sequences. *Discrete Math. Theoret. Comput. Sci.* **9**, 333–352.