Fifth Colloquium on Mathematics and Computer Science DMTCS proc. **AI**, 2008, 223–234

Polynomial tails of additive-type recursions

Eva-Maria Schopp[1]

[1] *University of Freiburg, Mathematical Stochastics, Eckerstr. 1, 79104 Freiburg, Germany*

Polynomial bounds and tail estimates are derived for additive random recursive sequences, which typically arise as functionals of recursive structures, of random trees, or in recursive algorithms. In particular they arise as parameters of divide and conquer type algorithms. We mainly focuss on polynomial tails that arise due to heavy tail bounds of the toll term and the starting distributions. Besides estimating the tail probability directly we use a modified version of a theorem from regular variation theory. This theorem states that upper bounds on the asymptotic tail probability can be derived from upper bounds of the Laplace–Stieltjes transforms near zero.

Keywords: polynomial tails, regular variation, random recursive sequences

1 Introduction

We are interested in sequences $(X_n)_{n\in\mathbb{N}_0}$ satisfying

$$X_n \stackrel{d}{=} \sum_{r=1}^{M} A_r(n) X^{(r)}_{I_r^{(n)}} + b_n, \quad n \geq n_0 \geq 1. \tag{1}$$

Here, X_n is the parameter of a problem of size n, which is split into $M \geq 1$ subproblems r of random sizes $I_r^{(n)} \in \{0, \dots, n-1\}$. $(X_n^{(r)})$ are distributional copies of (X_n) that correspond to the contribution of subgroup r. b_n is a random toll function term and $A_r(n)$ are random factors weighting the subproblems. Further, $(X_n^{(1)}), \dots, (X_n^{(M)}), (I^{(n)}, b_n, A(n))$ are independent where $I^{(n)} = (I_1^{(n)}, \dots, I_M^{(n)})$ and $A(n) = (A_1(n), \dots, A_M(n))$ denote the corresponding vectors. Finally, $\stackrel{d}{=}$ denotes equality in distribution.

The field of algorithms and data structure provides a substantial amount of sequences as in (1). In typical examples X_n ranges from depth, size, and path length of random trees, identification numbers of graph algorithms, the number of various substructures or components of combinatorial structures, space requirement, the number of comparisons, and other cost measures of algorithms, various parameters of communication and network models, and, in particular, to typical instances of the 'divide and conquer' paradigm. For numerous examples of this type we refer to the books of Mahmoud (1992), Sedgewick and Flajolet (1996), Szpankowski (2001), and Arratia et al. (2003). From those classical problems we can deduce variations where polynomial tails occur and our tail bounds apply, compare Section 2.

Many authors engaged themselves in the establishment of concentration results and tail bounds for (randomized) algorithms in general, which serve in particular for establishing approximation bounds and error estimates. Besides classical tail bounds like Chernoff's, Hoeffding's, Bennett's, Bernstein's bounds, and martingale bounds more recently techniques like induction methods, entropy methods, Talagrand's *convex-distance* inequality and others have been developed and applied. We refer to McDiarmid (1998), Motwani and Raghavan (1995), Boucheron et al. (2000), Boucheron et al. (2003)), and the references given therein. Concentration results supplement the asymptotic distributional analysis. They allow sharp error estimates, and in some cases they allow to establish laws of large numbers for related parameters like the height of trees (see e.g. Devroye (2002) and Broutin and Devroye (2006)). The following paper is the follower of the two papers Rüschendorf and Schopp (2007) and Rüschendorf and Schopp (2006) giving exponential tail bounds for additive-type and max-type recursive sequences.

In this paper we give some criteria to decide whether we can bound the tail of a recursive sequence as in (1) by a polynomial term. In Section 2 we use well-known results about tails of sums of regularly varying functions and give some examples. In Section 3 we prove upper bounds for the Laplace–Stieltjes transform. Then, using a modified version of a statement of Bingham and Doney (1974) we get upper bounds on the tail probabilities.

2 Bounds on Tail Probabilities

We are interested in sequences $(X_n)_{n\in\mathbb{N}_0}$ satisfying (1). Usually, the original recursion does not converge but we can obtain convergence of the corresponding modified recursion that is centered and normalized by a positive sequence $s(n)$. Typically, this sequence is of the order of the square root of the variance (if such exists) and the centering is about the mean if it exists. Consequently, we study the following modified sequence

$$Y_n := \frac{X_n - \mu_n}{s(n)}, \tag{2}$$

where, if (X_n) is s-integrable for some $s > 0$, we set for all $n \in \mathbb{N}_0$,

$$\mu_n = EX_n, \text{ if } s > 1, \quad s(n) = \max\{1, \sqrt{\operatorname{Var}(X_n)}\}, \text{ if } s > 2,$$

and $s(n)$ arbitrarily but positive else. Then, (Y_n) satisfies the following modified recurrence:

$$Y_n = \sum_{r=1}^{M} \underbrace{A_r(n)\frac{s(I_r^{(n)})}{s(n)}}_{=:A_r^{(n)}} Y^{(r)}_{I_r^{(n)}} + \underbrace{\frac{1}{s(n)}\left(b_n - \mu_n + \sum_{r=1}^{M} A_r(n)\mu_{I_r^{(n)}}\right)}_{=:b^{(n)}}. \tag{3}$$

For recursive sequences general limit theorems have been established by means of the contraction method. The sequence (Y_n) converges under suitable conditions in distribution to a random variable Y which is a solution of a suitable fixed point equation,

$$Y \stackrel{d}{=} \sum_{i=1}^{M} A_i Y_i + b. \tag{4}$$

We refer to Neininger and Rüschendorf (2004, Theorem 4.1) for one possible version of the contraction method and for more details and references.

Since we deal with sums of random variables in (1) it is very natural to use a variation of a well-known statement due to Feller (1971, p. 278). Lemma 2.1 reformulates that assertion that was originally formulated in terms of asymptotic equivalence instead of upper resp. lower bounds.
Note that with R_0 we denote the set of all slowly varying functions (at infinity) in Karamata's sense, compare Bingham et al. (1987). Additionally, define for functions $f(x), g(x) > 0$ on $\mathbb{R}_+$,

$$f(x) \leq_{as} g(x) \quad \text{if} \quad \overline{\lim_{x\to\infty}} \frac{f(x)}{g(x)} \leq 1 \quad \text{and} \quad f(x) \sim_{as} g(x) \quad \text{if} \quad \lim_{x\to\infty} \frac{f(x)}{g(x)} = 1. \tag{5}$$

Lemma 2.1 *Let F_1 and F_2 be two distribution functions such that as $x \to \infty$*

$$1 - F_i(x) \leq_{as} (\geq_{as})\, x^{-\rho} l_i(x), \quad i = 1, 2$$

where $l_i \in R_0,\ i = 1, 2$ and $\rho > 0$. Then, the convolution of F_1 with F_2, $G := F_1 * F_2$, satisfies

$$1 - G(x) \leq_{as} (\geq_{as})\, x^{-\rho}(l_1(x) + l_2(x)), \quad x \to \infty.$$

To prove the following Theorem 2.8 we will use a lemma due to Grey (1994), which we reformulate for the ease of reference.

Lemma 2.2 *Let $L \in R_0$ and $\alpha > 0$. Further assume*

1. $L(\lambda t) \leq \lambda^\alpha L(t)$, *for all* $\lambda \geq 1$,
2. $1 \geq t^{-\alpha} L(t) \geq t_0^{-\alpha} L(t_0)$, *for all* $0 < t \leq t_0$.

Then, given $\delta > 0$, there exists $K > 1$ such that

$$\frac{L(\lambda t)}{L(t)} \leq \max\{\lambda^\alpha, K\lambda^{-\delta}\}, \quad \textit{for all } t > 0, \lambda > 0. \tag{6}$$

Proof: Follows from Lemma 1 in Grey (1994) and its proof. □

Corollary 2.3 *Let $\alpha > 0$, $L \in R_0$. If X is a random variable with $P(X > t) = t^{-\alpha} L(t)$ for $t > 0$, then (6) holds for $L(\cdot)$ and α.*

We will formulate two more useful lemmas for the ease of reference. The first one is a restatement of Lemma 2.1 of Davis and Resnick (1996) for non-negative random variables. The second, Lemma 2.5, was proven in Goovaerts et al. (2005, proof of Theorem 2.1, p. 14 ff).

Lemma 2.4 *For a sequence of non-negative random variables $(X_i)_{1\leq i\leq n}$ and a distribution function $F \in R_{-\alpha}$, $\alpha > 0$, let*

$$\lim_{x\to\infty} \frac{P(X_i > x)}{1 - F(x)} = c_i, \quad \textit{for } i = 1, \ldots, n$$

and

$$\lim_{x\to\infty} \frac{P(X_i > x, X_j > x)}{1 - F(x)} = 0, \quad \textit{for } 1 \leq i \neq j \leq n.$$

Then

$$\lim_{x\to\infty} \frac{P(\sum_{i=1}^n X_i > x)}{1 - F(x)} = \sum_{i=1}^n c_i.$$

Lemma 2.5 *Let $X_1, \dots, X_n$ be independent random variables and $(\theta_1, \dots, \theta_n)$ be non-negative random variables independent of $(X_i)_{1 \le i \le n}$. Let $F \in R_{-\alpha}$ for some $\alpha > 0$ and assume*

$$E \sum_{k=1}^{n} \theta_k^{\beta} < \infty, \quad \exists \beta > \alpha > 0.$$

1. *If $X_1, \dots, X_n$ are identically distributed with distribution function F, we have*

$$\lim_{x \to \infty} \frac{P\left(\sum_{k=1}^{n} \theta_k X_k > x\right)}{1 - F(x)} = \sum_{k=1}^{n} E\theta_k^{\alpha}.$$

2. *Let*

$$\lim_{x \to \infty} \frac{P(X_i > x)}{1 - F(x)} = c_i, \quad \text{for } i = 1, \dots, n,$$

then

$$\lim_{x \to \infty} \frac{P\left(\sum_{k=1}^{n} \theta_k X_k > x\right)}{1 - F(x)} \le \sum_{k=1}^{n} c_k E\theta_k^{\alpha}.$$

Now we formulate the announced tail result.

Theorem 2.6 *Let Y_k satisfy recursion (3) and let $\alpha \in \mathbb{R}^+$. The coefficients $A_r(n)$ (respectively $A_r^{(n)}$) are assumed to be non-negative. Assume for $\beta > \alpha > 0$ and some $c > 0$*

1.) $P(Y_k \ge x) \le_{as} c x^{-\alpha} l(x), \quad k = 0, \dots, n_0 - 1,\ x \to \infty,$

2.) $\forall k \ge n_0 :\ \sum_{r=1}^{M} E\left(A_r^{(k)}\right)^{\beta} < \infty,$

3.) $\forall k \ge n_0 :\ P(b^{(k)} > x) \sim_{as} x^{-\alpha} l(x), \quad x \to \infty,$

with $l \in R_0$ satisfying (6). Define

$$c_n := \begin{cases} c & n < n_0, \\ 1 + E\sum_{r=1}^{M} \left(A_r^{(n)}\right)^{\alpha} c_{I_r^{(n)}} & n \ge n_0. \end{cases}$$

Then we have polynomial bounds for $(Y_k)_{k \in \mathbb{N}_0}$:

$$P(Y_k \ge x) \le_{as} c_k \frac{1}{x^{\alpha}} l(x), \quad x \to \infty,\ k \in \mathbb{N}_0. \tag{7}$$

If condition 3.) is replaced by

3.') $\forall k \ge n_0 :\ P(b^{(k)} > x) = o(x^{-\alpha} l(x)),$

then (7) is valid with $\widetilde{c}_k$ instead of c_k where

$$\widetilde{c}_n := \begin{cases} c & n < n_0, \\ E\sum_{r=1}^{M} \left(A_r^{(n)}\right)^{\alpha} \widetilde{c}_{I_r^{(n)}} & n \ge n_0. \end{cases}$$

Proof: We assume first conditions 1.)-3.). We prove the assertion for $Y_k,\ k \in \mathbb{N}_0$, by induction on k. The statement is valid for $|Y_0|, \dots, |Y_{n_0-1}|$ by assumption 1). The calculations for (7) are then straightforward: Assume the induction hypothesis (7) holds for all $|Y_l|,\ l \le k-1$. Then, by independence, induction hypothesis, and Lemma 2.1 we obtain for all $\varepsilon > 0$

$$
\begin{aligned}
P(Y_k > t) =&\ P\left(\sum_{r=1}^{M} A_r(k) \frac{s(I_r^{(k)})}{s(k)} Y^{(r)}_{I_r^{(k)}} + b^{(k)} > t\right) \\
=&\ P\left(b^{(k)} > (1+\varepsilon)t\right) - P\left(b^{(k)} > (1+\varepsilon)t,\ b^{(k)} + \sum_{i=1}^{M} A_i^{(k)} Y^{(i)}_{I_i^{(k)}} \le t\right) \\
&+ P\left((1-\varepsilon)t < b^{(k)} \le (1+\varepsilon)t,\ b^{(k)} + \sum_{i=1}^{M} A_i^{(k)} Y^{(i)}_{I_i^{(k)}} > t\right) \\
&+ P\left(b^{(k)} \le (1-\varepsilon)t,\ b^{(k)} + \sum_{i=1}^{M} A_i^{(k)} Y^{(i)}_{I_i^{(k)}} > t\right) \\
=:&\ I_1^{(k)}(t) - I_2^{(k)}(t) + I_3^{(k)}(t) + I_4^{(k)}(t).
\end{aligned}
$$

First note that

$$
\begin{aligned}
I_1^{(k)}(t) &\sim_{as} ((1+\varepsilon)t)^{-\alpha}\, l((1+\varepsilon)t) \\
&\sim_{as} t^{-\alpha} l(t),\ t \to \infty.
\end{aligned}
$$

Next,

$$
\begin{aligned}
0 \le I_3^{(k)}(t) \le&\ \ P((1-\varepsilon)t < |b^{(k)}| \le (1+\varepsilon)t) \\
\sim_{as}&\ \left((1-\varepsilon)^{-\alpha} - (1+\varepsilon)^{-\alpha}\right) t^{-\alpha} l(t).
\end{aligned}
$$

And last with Lemma 2.1 we have

$$
\begin{aligned}
I_4^{(k)}(t) =&\ \ t^{-\alpha} l(t) E\left(\frac{P\left(\sum_{i=1}^{M} A_i^{(k)} Y^{(i)}_{I_i^{(k)}} > t - b^{(k)} \,\big|\, \left(A^{(k)}, I^{(k)}, b^{(k)}\right)\right)}{t^{-\alpha} l(t)} 1_{\left\{b^{(k)} \le (1-\varepsilon)t\right\}}\right) \\
=&\ \ t^{-\alpha} l(t) E\Bigg(\sum_{\widetilde{J} \in \{1,\dots,M\}} \Bigg[\frac{P\left(\sum_{i \in \widetilde{J}} A_i^{(k)} Y^{(i)}_{I_i^{(k)}} > t - b^{(k)} \,\big|\, \left(A^{(k)}, I^{(k)}, b^{(k)}\right)\right)}{t^{-\alpha} l(t)} \\
&\qquad \times 1_{\left\{\bigcap_{i \in \widetilde{J}} \{A_i^{(k)} > 0\} \,\cap\, \bigcap_{i \in \widetilde{J}^c} \{A_i^{(k)} = 0\} \,\cap\, \{b^{(k)} \le (1-\varepsilon)t\}\right\}}\Bigg]\Bigg).
\end{aligned}
$$

The limes superior of the random variable in brackets converges pointwise to a finite limit less or equal than

$$
\sum_{i \in \widetilde{J}} c_{I_i^{(k)}} \left(A_i^{(k)}\right)^{\alpha} 1_{\left\{\bigcap_{i \in \widetilde{J}} \{A_i^{(k)} > 0\} \,\cap\, \bigcap_{i \in \widetilde{J}^c} \{A_i^{(k)} = 0\}\right\}}
$$

as $t \to \infty$ and is dominated by

$$\sum_{i \in \widetilde{J}} \frac{P\left(Y^{(r)}_{I^{(k)}_r} > \left(A^{(k)}_r\right)^{-1} \varepsilon t\right)}{t^{-\alpha} l(t)},$$

where the summands are by Lemma 2.2 for $\delta = \beta - \alpha$ dominated by $\max\{1, K\varepsilon^{-\beta}|A^{(k)}_r|^{\beta}\}$ for some constant $K > 0$. This random variable is integrable and so, by dominated convergence, we conclude that

$$I^{(k)}_4(t) \leq_{as} E \sum_{r=1}^{M} c_{I^{(k)}_r} \left(A^{(k)}_r\right)^{\alpha} t^{-\alpha} l(t), \quad t \to \infty.$$

Now collecting our results and letting $\varepsilon \searrow 0$ we obtain the statement.

Now using condition 3.′) instead of condition 3.), we prove the statement again by induction. We set $\bar{F}(x) := x^{-\alpha} l(x)$ and $\varepsilon := \beta - \alpha$. The statement is true for $Y_0, \ldots, Y_{n_0 - 1}$, so for the induction step we assume that the assertion is valid for $Y_k,\ k \leq n-1$. Then

$$P(Y_n \geq x) = P\left(\sum_{r=1}^{M} A^{(n)}_r Y^{(r)}_{I^{(n)}_r} + b^{(n)} > x\right).$$

Further note that

$$P\left(A^{(n)}_r Y^{(r)}_{I^{(n)}_r} > x\right) = P\left(A^{(n)}_r \sum_{k=0}^{n-1} 1_{\{I^{(n)}_r = k\}} Y^{(r)}_k > x\right).$$

Since for $l \neq k$

$$P\left(1_{\{I^{(n)}_r = k\}} Y^{(r)}_k > x, 1_{\{I^{(n)}_r = l\}} Y^{(r)}_l > x\right) \leq P\left(\{I^{(n)}_r = k\} \cap \{I^{(n)}_r = l\}\right) = P(\emptyset) = 0$$

we obtain with Lemma 2.5

$$\begin{aligned} \lim_{x\to\infty} \frac{P\left(A^{(n)}_r Y^{(r)}_{I^{(n)}_r} > x\right)}{\bar{F}(x)} &= \lim_{x\to\infty} \frac{P\left(\sum_{k=0}^{n-1} (A^{(n)}_r 1_{\{I^{(n)}_r = k\}}) Y^{(r)}_k > x\right)}{\bar{F}(x)} \\ &\leq \sum_{k=0}^{n-1} E\left(\left(A^{(n)}_r\right)^{\alpha} 1_{\{I^{(n)}_r = k\}} \widetilde{c}_k\right) \\ &= E\left(\left(A^{(n)}_r\right)^{\alpha} \widetilde{c}_{I^{(n)}_r}\right) \end{aligned}$$

by independence of $(A^{(n)}, I^{(n)})$ and Y_k. Furthermore,

$$\lim_{x\to\infty} \frac{P\left(\sum_{r=1}^{M} A^{(n)}_r Y^{(r)}_{I^{(n)}_r} > x,\ b^{(n)}_r > x\right)}{\bar{F}(x)} \leq \lim_{x\to\infty} \frac{P(b^{(n)}_r > x)}{\bar{F}(x)} = 0$$

by assumption on the tail of $b_r^{(n)}$, and since for $1 \le l \ne k \le M$ we have for $\delta > 0$ small enough

$$\begin{aligned}
&P\left(A_l^{(n)} Y_{I_l^{(n)}}^{(l)} > x,\ A_k^{(n)} Y_{I_k^{(n)}}^{(k)} > x\right) \\
&\le P\left(A_l^{(n)} \left(Y_{I_l^{(n)}}^{(l)}\right)^+ > x,\ A_k^{(n)} \left(Y_{I_k^{(n)}}^{(k)}\right)^+ > x\right) \\
&\le P\left(A_l^{(n)} > x^{1-\delta}\right) + P\left(A_l^{(n)} \left(Y_{I_l^{(n)}}^{(l)}\right)^+ > x,\ A_k^{(n)} \left(Y_{I_k^{(n)}}^{(k)}\right)^+ > x,\ A_l^{(n)} \le x^{1-\delta}\right) \\
&\le x^{-(1-\delta)(\alpha+\varepsilon)} E(A_l^{(n)})^{\alpha+\varepsilon} + E\left[P\left(\left(Y_{I_l^{(n)}}^{(l)}\right)^+ > x^\delta \Big| \left(A_r^{(n)}, I^{(n)}, b^{(n)}\right)\right)\right. \\
&\qquad\qquad\qquad \left.\times P\left(A_k^{(n)} \left(Y_{I_k^{(n)}}^{(l)}\right)^+ > x \Big| \left(A_r^{(n)}, I^{(n)}, b^{(n)}\right)\right)\right] \\
&= o(\bar{F}(x)).
\end{aligned}$$

Now combining our results we can again use Lemma 2.4 and Lemma 2.5 to obtain

$$\begin{aligned}
\lim_{x\to\infty} \frac{P\left(\sum_{r=1}^M A_r^{(n)} Y_{I_r^{(n)}}^{(r)} + b_r^{(n)} > x\right)}{\bar{F}(x)} &= \lim_{x\to\infty} \frac{P\left(\sum_{r=1}^M A_r^{(n)} Y_{I_r^{(n)}}^{(r)} > x\right)}{\bar{F}(x)} \\
&= \sum_{r=1}^M \lim_{x\to\infty} \frac{P\left(A_r^{(n)} Y_{I_r^{(n)}}^{(r)} > x\right)}{\bar{F}(x)} \\
&\le \sum_{r=1}^M E\left(\left(A_r^{(n)}\right)^\alpha \widetilde{c}_{I_r^{(n)}}\right) \\
&= \widetilde{c}_n. \qquad \square
\end{aligned}$$

Remark 2.7 *If $\sum_{r=1}^M \left(A_r^{(n)}\right)^\alpha \le a < 1$ for all $n \ge n_0$ and if $(c_n)_{n\ge n_0}$ is a non-decreasing sequence, then $c_n \le \frac{1}{1-a}$ for all $n \ge n_0$.*

As an immediate consequence we get the following statement.

Lemma 2.8 *Let Y_k satisfy recursion (3) and assume*

1) $P(|Y_k| > x) \le_{as} x^{-\alpha} l(x), \quad 0 \le k < n_0$, *for some* $\alpha > 0,\ l \in R_0$,

2) $\forall k \ge n_0 :\ \sum_{r=1}^M E|A_r^{(k)}|^\alpha \le 1$,

3) $\forall k \ge n_0 :\ \sum_{r=1}^M E|A_r^{(k)}|^{\widetilde{\alpha}} < \infty, \quad$ *for some* $\widetilde{\alpha} > \alpha$,

4) $\forall k \ge n_0 :\ P(|b^{(k)}| > x) = o\left(x^{-\alpha} l(x)\right)$,

whereas l satisfies (6). Then we have polynomial bounds for the sequence $(Y_k)_{k\in\mathbb{N}_0}$

$$P(|\, Y_k \,| \ge x) \le_{as} \frac{1}{x^\alpha} l(x), \quad x \to \infty,\ k \in \mathbb{N}_0. \tag{8}$$

Proof: Using the upper bound

$$\begin{aligned} P(|Y_k| \geq x) = & P\left(\left|\sum_{r=1}^{M} A_r(k)\frac{s(I_r^{(k)})}{s(k)} Y^{(r)}_{I_r^{(k)}} + b^{(k)}\right| \geq x\right) \\ \leq & P\left(\sum_{r=1}^{M} \left|A_r^{(k)}\right| \left|Y^{(r)}_{I_r^{(k)}}\right| + \left|b^{(k)}\right| \geq x\right) \end{aligned}$$

and the proof of Theorem 2.6 applied on $\sum_{r=1}^{M} \left|A_r^{(k)}\right| \left|Y^{(r)}_{I_r^{(k)}}\right| + \left|b^{(k)}\right|$ we obtain the statement. □

Examples 2.9

1. *Let us consider a variation of Kolmogorov's rock model also called Bisection model. An object is initially of mass X where X is a real valued random variable. At time one it splits into two objects with uniform size. At time n each of the 2^n objects are broken independently of the other ones into two objects with uniform size. The size of an object is therefore a product of independent uniform random variables times X. Taking logarithms gives a discrete time branching random walk. Various parameters of branching random walks have recursive structure and their tails depend now highly on the tail of X. If X has stable like tails then this example fits into our framework.*

2. *Consider a set of $\lceil N \rceil := \inf\{i \in \mathbb{N} : i \geq N\}$ different real valued numbers where $N > 0$ is a non-negative random variable with heavy tails. Then, applying for examples the well-known sorting algorithm Quicksort in oder to sort this set of random size we have for the number of comparisions $C_{\lceil N \rceil}$ needed by Quicksort*

$$C_{\lceil N \rceil} \stackrel{d}{=} C_{I_{\lceil N \rceil}} + C_{n-1-I_{\lceil N \rceil}} + (\lceil N \rceil - 1).$$

   Now this key performance number fits into our tail framework.

3. *Let $b \in \mathbb{N}$ and $c_1 \leq c_2 \leq \ldots \leq c_b$ be positive integers. Lopsided trees are b-ary rooted trees where for each node, the edge to its i-th child has length c_i (compare for example Broutin and Devroye (2006) and references given therein). Consider a lopsided tree in which the lengths of the edges are random and have polynomial tails. Then various recursive tree parameters can be handled with our tail estimates as, for example, the internal path length.*
   In addition, other random trees such as random binary search trees or random recursive trees can also be weighted by different heavy-tailed edge lengths and then be studied with our tail results.

3 Bounds on Laplace–Stieltjes transforms

One major disadvantage of the previous tail bound results is the fact that they do not hold uniformly and can, therefore, not be transferred to the limit described as a fixed point. However, there are cases when we can prove uniform upper tail estimates, for example in the case $M = 2$ and if α, the tail index, lies in $(0, 1]$. Note that estimates become highly involved and need more preliminary assumptions if we want to extend the method to $\alpha \in (1, \infty)$. Before we proceed with Theorem 3.3 that contains our results in the special case, we present a modified version of an assertion due to Bingham and Doney (1974). The original statement is formulated in terms of asymptotic equivalence instead of asymptotic upper bounds.

Note that Lemma 3.1 is neither a direct consequence of the original assertion nor a direct generalization of the corresponding proof. For the definition of O-regularly varying functions (short: OR) and functions of bounded increase (short BI) we refer to the book of Bingham et al. (1987).

Lemma 3.1 *Let X be non-negative and unbounded. Let $l \in R_0$. For every $n \in \mathbb{Z}^+$ such that $0 < \mu_n < \infty$, for every $\beta, 0 \leq \beta \leq 1$, let $\alpha = n + \beta$. If*

(i) $\exists c > 0 : \; g_n(s) \leq_{as} cs^{\beta} l\left(\frac{1}{s}\right), \quad s \searrow 0,$

then

(ii) $\exists C > 0 : \; 1 - F(x) \leq_{as} Cx^{-\alpha} l(x), \quad x \to \infty.$

Proof: First note that $sg_n(1/s)$ is an O-regularly varying function and observe that $s^{-1}g_n(s)$ is the Laplace–Stieltjes transform of the function $\int_0^x dt \int_t^\infty y^n dF(y)$. Further, $\int_0^x dt \int_t^\infty y^n dF(y)$ is a non-negative, non-decreasing function, so by the de-Haan–Stadtmüller-Theorem (compare Bingham et al. (1987)) (i) is equivalent to

(iii) $\exists c_1 > 0 : \; U(x) := \int_0^x dt \int_t^\infty y^n dF(y) \leq_{as} c_1 x^{1-\beta} l(x), \quad x \to \infty$ and $U \in$ OR.

Let $T_n(t) := u(t) := \int_t^\infty y^n dF(y)$. Then u is non-negative and monotone, but non-increasing with $U(x) = \int_0^x u(t)dt$. Therefore $U(x) \geq xu(x)$. So from (iii) we get

(iv) $\exists c_2 > 0 : \; T_n(x) = \int_x^\infty y^n dF(y) \leq_{as} c_2 x^{-\beta} l(x), \quad x \to \infty.$

As $\mu_n < \infty$, $\int_0^\infty t^{n-1}(1 - F(t))dt$ is convergent by Bingham et al. (1987, Lemma 8.1.5). Integration by parts yields $T_n(x) = x^n(1 - F(x)) + n \int_x^\infty y^{n-1}(1 - F(y))dy$. Hence

$$1 - F(x) = \frac{1}{x^n} T_n(x) - n \int_x^\infty y^{-(n+1)} T_n(y)dy \leq \frac{1}{x^n} T_n(x),$$

whence (iv) implies (ii). □

Remark 3.2 *Assume 1.(i) is valid in Lemma 3.1 for some $c > 0$. Then 1.(ii) is valid and we can choose $C = ec$ as a suitable constant since for a non-decreasing, non-negative function U in OR, $\frac{\hat{U}(1/t)}{U(t)} \geq e^{-1}$ (this can be obtained from the proof of the de Haan–Stadtmüller-Theorem).*

Theorem 3.3 *Let $(Y_n)_{n \in \mathbb{N}_0}$ be a sequence satisfying (3) with $M = 2$. We assume further*

$$\hat{F}_n(s) = Ee^{-s|Y_n|} \geq 1 - c_1 s^{\alpha}(1 + o_n(1)), \quad s \searrow 0, \; n = 0, \dots, n_0 - 1$$

with $c_1 \in \mathbb{R}_+$, $0 \leq \alpha < 1$, and $o_n(1) \to 0, \; s \searrow 0$.
Let for all $n \geq n_0$ and some $c_2 > 0$

1) $\hat{F}_{|b^{(n)}|}(s) \geq 1 - c_2 \left(\frac{s}{s(n)}\right)^{\alpha} (1 + o_b(1)), \quad o_b(1) \to 0, \; s \searrow 0,$

2) $\sum_{r=1}^2 \left(|A_r(n)| \frac{s(I_r^{(n)})}{s(n)}\right)^{\alpha} \leq 1 - \left(\frac{1}{s(n)}\right)^{\alpha}.$

Then we have uniform polynomial bounds for (Y_n) as well as for the solution of (4):

$$P(\mid Y_k \mid \geq x) \leq_{as} e \max\{c_1, c_2\} x^{-\alpha}, \quad x \to \infty,\ k \in \mathbb{N}_0. \tag{9}$$

Proof: We establish by induction uniformly for $n \in \mathbb{N}$ Laplace–Stieltjes estimates for (Y_n) which correspond by Lemma 3.1 and Remark 3.2 to the thesis above. Write $\hat{F}_j$ for the Laplace–Stieltjes transform of $|Y_j|$. For the induction step we have that $1 - \hat{F}_j(s) \leq_{as} c^* s^\alpha(1 + o(1)),\ j = 1, \ldots, n-1$, where we choose $o(1) := \max\{|o_1(1)|, \ldots, |o_{n_0-1}(1)|, |o_b(1)|\}$ and $c^* := \max\{c_1, c_2\}$. For the induction step consider for $s \searrow 0$

$$\begin{aligned}
\hat{F}_n(s) &= E e^{-s|Y_n|} \\
&\geq E\left(E\left(e^{-s\sum_{r=1}^2 |A_r^{(n)}|\left|Y_{I_r^{(n)}}^{(r)}\right| + |b^{(n)}|} \mid \left(A^{(n)}, I^{(n)}, b^{(n)}\right)\right)\right) \\
&= E\left[e^{-s|b^{(n)}|} \prod_{r=1}^{2} \hat{F}_{I_r^{(n)}}\left(s|A_r(n)|\frac{s(I_r^{(n)})}{s(n)}\right)\right] \\
&\geq E\left[e^{-s|b^{(n)}|} \prod_{r=1}^{2}\left(1 - c^*\left(s|A_r(n)|\frac{s(I_r^{(n)})}{s(n)}\right)^\alpha (1 + o(1))\right)\right] \\
&= E\left[e^{-s|b^{(n)}|}\left(1 - c^* s^\alpha (1 + o(1)) \sum_{r=1}^{2} |A_r^{(n)}|^\alpha + c^{*2} s^{2\alpha}(1 + o(1))^2 |A_1^{(n)} A_2^{(n)}|^\alpha\right)\right] \\
&\geq E\left[e^{-s|b^{(n)}|}\left(1 - c^* s^\alpha (1 + o(1)) \sum_{r=1}^{2} |A_r^{(n)}|^\alpha\right)\right] \\
&\geq \left(1 - c^*\left(\frac{1}{s(n)}\right)^\alpha s^\alpha (1 + o(1))\right)\left(1 - c^*\left(1 - \left(\frac{1}{s(n)}\right)^\alpha\right) s^\alpha (1 + o(1))\right) \\
&= 1 - c^* s^\alpha (1 + o(1)) + (c^*)^2 (1 + o(1))^2 \left(\frac{1}{s(n)}\right)^\alpha \left(1 - \left(\frac{1}{s(n)}\right)^\alpha\right) s^{2\alpha} \\
&\geq 1 - c^* s^\alpha (1 + o(1)).
\end{aligned}$$

□

Remark 3.4 *From Goldie and Grübel (1996, Theorem 4.1) we deduce that if one of the coefficients in (4) takes values greater than one with positive probability, we have at least a power-law tail. More results on the tail behaviour of fixed points as in (4) in the case $M = 1$ are given in Kesten (1973) (working in many dimensions), Grincevičius (1975), Goldie (1991) (working with a more general functional equation), Grey (1994), Embrechts and Goldie (1994), and Goldie and Maller (2000). An interesting summary about the state of the art is given in Embrechts and Goldie (1994). For fixed points as in (4) in the case $b = 0$, sometimes called smoothing transformations, we refer to Liu (1998) and Iksanov (2004) for tail results.*

References

R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic Combinatorial Structures: A Probabilistic Approach*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003. ISBN 3-03719-000-0.

N. Bingham and R. Doney. Asymptotic properties of super-critical branching processes. I: The Galton–Watson process. *Advances in Applied Probability*, 6:711–731, 1974.

N. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation*. Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 1987.

S. Boucheron, G. Lugosi, and P. Massart. A sharp concentration inequality with applications. *Random Structures Algorithms*, 16(3):277–292, 2000. ISSN 1042-9832.

S. Boucheron, G. Lugosi, and P. Massart. Concentration inequalities using the entropy method. *Annals of Probability*, 31(3):1583–1614, 2003. ISSN 0091-1798.

N. Broutin and L. Devroye. Large deviation for the weighted height of an extended class of trees. *Algorithmica*, 46:271–297, 2006.

R. Davis and S. Resnick. Limit theory for bilinear processes with heavy-tailed noise. *Annals of Applied Probability*, 6(4):1191–1210, 1996.

L. Devroye. Laws of large numbers and tail inequalities for random tries and PATRICIA trees. *J. Comput. Appl. Math.*, 142(1):27–37, 2002. ISSN 0377-0427.

P. Embrechts and C. M. Goldie. Perpetuities and random equations. In *Asymptotic Statistics: Proceedings of the Fifth Prague Symposium, Physica*, pages 75–86, Heidelberg, 1994.

W. Feller. *An Introduction to Probability Theory and Applications II*. Second Edition, Wiley Series in Probability and Mathematical Statistic. John Wiley & Sons, Inc., 1971. ISBN 0-471-257095.

C. M. Goldie. Implicit renewal theory and tails of solutions of random equations. *Annals of Applied Probability*, 1(1):126–166, 1991.

C. M. Goldie and R. Grübel. Perpetuities with thin tails. *Advances in Applied Probability*, 28:463–480, 1996.

C. M. Goldie and R. A. Maller. Stability of perpetuities. *Annals of Probability*, 28(3):1195–1218, 2000.

M. J. Goovaerts, R. Kaas, R. J. Laeven, Q. Tang, and R. Vernic. The tail probability of discounted sums of Pareto-like losses in insurance. *Scandinavian Actuarial Journal*, 6:446–461, 2005.

D. Grey. Regular variation in the tail behaviour of solutions of random difference equations. *The Annals of Applied Probability*, 4(1):169–183, 1994.

A. Grincevičius. One limit distribution for a random walk on the line. *Lithuanian Math.J.*, 15:580–589, 1975.

A. Iksanov. Elementary fixed points of the BRW smoothing transforms with finite number of summands. *Stochastic Processes and Their Applications*, 114:27–50, 2004.

H. Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131:207–248, 1973.

Q. Liu. Fixed points of a generalized smoothing transformation and applications to the branching random walk. *Advances in Applied Probability*, 30:85–112, 1998.

H. M. Mahmoud. *Evolution of Random Search Trees*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., New York, 1992. ISBN 0-471-53228-2. A Wiley-Interscience Publication.

C. McDiarmid. Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics*, volume 16 of *Algorithms Combin.*, pages 195–248. Springer, Berlin, 1998.

R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, Cambridge, 1995. ISBN 0-521-47465-5.

R. Neininger and L. Rüschendorf. A general limit theorem for recursive algorithms and combinatorial structures. *Annals of Applied Probability*, 14(1):378–418, 2004. ISSN 1050-5164.

L. Rüschendorf and E.-M. Schopp. Exponential bounds and tails for additive random recursive sequences. *DMTCS*, 9(1):333–352, 2007.

L. Rüschendorf and E.-M. Schopp. Exponential tail bounds for max-recursive sequences. *Electronic Communication in Probability*, 11:266–277, 2006.

R. Sedgewick and P. Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley, Amsterdam, 1996.

W. Szpankowski. *Average Case Analysis of Algorithms on Sequences*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2001. ISBN 0-471-24063-X.