Discrete Mathematics and Theoretical Computer Science DMTCS vol. **9:1**, 2007, 333–352

Exponential bounds and tails for additive random recursive sequences

Ludger Rüschendorf and Eva-Maria Schopp

University of Freiburg, Mathematical Stochastics, Eckerstr. 1, 79104 Freiburg, Germany

received April 3, 2007, *revised September 26, 2007*, *accepted November 21, 2007*.

Exponential 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 derive tail bounds from estimates of the Laplace transforms and of the moment sequences. For the proof we use some classical exponential bounds and some variants of the induction method as well as various characterization results of distributions with exponential tails. The paper generalizes previous results on linear exponential tails and on subgaussian tails to more general classes of additive random recursive sequences. We establish conditions which imply exponential bounds and tail bounds of different strength and order. In particular we give sufficient conditions for tail bounds of the order $\exp(-at^p)$. The proof of these tail bounds is based on a classical characterization result of Kasahara (1978).

Keywords: exponential tails, exponential inequalitites, subgaussian tails, recursive algorithms

1 Introduction

A large number of functionals $(X_n)_{n\geq 0}$ of random recursive structures, random trees, and recursive algorithms satisfy recurrences of the form

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

Here X_n is the interesting parameter of a problem of size n, which is split into $M \geq 1$ subproblems r of random sizes $I_r^{(n)} \in \{0, \dots, n-1\}$. $(X_k^{(r)})$ are distributional copies of (X_k), 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.

In typical examples X_n ranges from depth, size, and path length of random trees, the number of various substructures or components of combinatorial structures, the number of comparisons, space requirement, and other cost measures of algorithms to various parameters of communication 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).

A general distributional limit theorem for this type of recurrences was given together with many examples in Neininger and Rüschendorf (2004a,b) and Rösler (2001) by means of the contraction method. The limit of X_n after normalization is characterized as unique solution of some fixed point equation of the form

$$X \stackrel{d}{=} \sum_{i=1}^{M} A_i X^{(i)} + b, \tag{2}$$

where the distributions are restricted to some subclass as e.g. to $M_{2,c}$, the class of all distributions with finite second moment and first moment equal to c. Here $X^{(1)}, \dots, X^{(M)}$ are independent copies of X and the random vector $(A_1, \dots, A_M, b)$ arises as the limit of the corresponding normalized vector $(A_1(n), \dots, A_M(n), b_n)$.

There has been quite a lot of work going on for establishing 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 from statistics 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. For references we refer to McDiarmid (1998), Motwani and Raghavan (1995), Boucheron et al. (2000, 2003), and many references given in these papers. The concentration results supplement the asymptotic distributional analysis. They allow sharp error estimates. 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 (2005)).

Exponential bounds for solutions of the fixed point equation in (2) were given by an induction method in Rösler (1992). For the Quicksort algorithm exponential bounds were established in Rösler (1991), see also Fill and Janson (2002) for some improvements in this case. For some classes of recurrences as in (1) bounds were given in Karp (1994) while Neininger (2005) established subgaussian bounds for (X_n) in cases, where the rescaled quantities converge to a normal limit. In this paper we aim to develop some exponential tail bounds for more general classes of additive recursive sequences.

In the first part of this paper we establish exponential bounds for Laplace transforms leading to exponential tail bounds for recursive algorithms of type (1) and for solutions of recursive equations of type (2). The proof is based on an induction argument. In the case that the convergence theorem for recursive algorithms is applicable we alternatively also indicate a proof which is based on a uniform approximation argument for Laplace transforms. Under stronger assumptions on the coefficients we establish a stronger exponential bound valid on all finite intervals.

In the second part we derive various exponential inequalities for the Laplace transform and for the tail probabilities on the whole real line. In particular we derive exponential bounds of subgaussian type for a general class of recursive sequences extending previous results in Neininger (2005). By alternative methods of proof based on Bernstein's inequality and on a moment characterization of subgaussianity we derive under corresponding conditions on the coefficients subgaussian behaviour resp. linear exponential bounds on the whole real line.

In the final part of the paper we establish general asymptotic bounds for the moments of X_n. A characterization of exponential tails by the asymptotics of moments due to Kasahara (1978) allows to establish some general conditions for exponential tails of any order.

The paper gives a series of results on exponential (tail) bounds for recursive algorithms. Depending on the strength of the conditions on the coefficients of the recursive sequence we obtain a broad spectrum of exponential bounds of different strength. For the proofs we use a variety of different techniques, making use in particular of induction arguments, approximation arguments for Laplace transforms, classical inequalities as Hoeffding's and Bernstein's inequalities, and various characterization results of exponential tails by moment conditions and by bounds on the Laplace transforms. All results in this paper make use in some sense of the 'contradiction' condition $E\sum_i A_i^2 < 1$ resp. $E\sum_r A_r^2(n) < 1$ for the recursive sequence. Thus the exponential bounds in this paper are usually not well suited for (asymptotically) normal sequences, which are 'typically' characterized by the limit equation (2) with $b = 0$ and $\sum_i A_i^2 = 1$. Our results apply only in some of these normal cases, when special properties of the expansion of the mean hold true (see for example Remark 3.2).

2 Bounds on the Laplace transform

The fixed point equation (2) can be written in operator form as

$$F \stackrel{d}{=} S(F), \tag{3}$$

where $S(F) \stackrel{d}{=} \mathcal{L}\big(\sum_{i=1}^M A_i X_i + b\big)$ is the distribution of the right-hand side of (2). Under suitable conditions on A_i, b it is known that a unique solution F^* of (3) with finite second moment and specified first moment exists and that the iterated operator $S^n F$ converges to F^* w.r.t. the minimal ℓ_2-metric (see Rösler (1992), Rachev and Rüschendorf (1995)). Rösler (1992) established the following exponential bound for a solution Z of the fixed point equation (2).

Theorem 2.1 (Exponential bound for fixed point equation (Rösler (1992)))
Consider the fixed point equation (2) and assume the following conditions:

a) $E\sum_{i=1}^M A_i^2 < 1, \quad \sup_i |A_i| \le 1,$

b) $Ee^{\lambda b} < \infty, \quad \forall \lambda \in [-L, L].$

Further, in case

c1) $E\sum_{i=1}^M A_i \neq 1$, *define c by:* $Eb = c(1 - E\sum_i A_i)$

c2) $E\sum_{i=1}^M A_i = 1$, *let* $Eb = 0$ *and let* $c \in \mathbb{R}^1$ *be any fixed number.*

Then for any random variable $X \stackrel{d}{=} F$ *with* $EX = c$, $Ee^{\lambda X} \le e^{c\lambda + K\lambda^2}$ *for some* $K > 0$ *and all* $\lambda \in [-L, L]$ *holds:*

1) $S^n F$ *converges to the unique fixed point* F^* *in* $M_{2,c}$ *w.r.t.* ℓ_2 *and*

2) for $Z \stackrel{d}{=} F^*$ *or* $Z \stackrel{d}{=} S^n F$ *holds uniformly in* $n \in \mathbb{N}$ *for some constant* $\widetilde{K} \ge 0$

$$Ee^{\lambda Z} \le e^{c\lambda + \widetilde{K}\lambda^2}, \quad \forall \lambda \in [-L, L]. \tag{4}$$

Remark 2.2 *As consequence of Theorem 2.1 one obtains for $Z \stackrel{d}{=} F^*$ or $Z \stackrel{d}{=} S^n F$ uniformly in n*

$$Ee^{\lambda(Z-EZ)} \le e^{\widetilde{K}\lambda^2}, \quad \lambda \in [-L, L]. \tag{5}$$

This implies by Chernoff's bounding technique the tail estimates for all $\lambda \in [-L, L]$

$$P(Z - EZ \ge t) \le P(\exp(\lambda(Z - EZ)) \ge \exp(\lambda t)) \le \exp(\widetilde{K}\lambda^2 - \lambda t).$$

This becomes minimal for $\lambda = t/(2\widetilde{K})$ if this is less or equal to L and in the other case we set $\lambda = L$. Since the same bound holds for $P(EZ - Z \ge t)$ we obtain

$$P(|Z - EZ| \ge t) \le \begin{cases} 2\exp\left(-\dfrac{t^2}{4\widetilde{K}}\right), & 0 < t \le 2L\widetilde{K}, \\ 2\exp(\widetilde{K}L^2 - Lt), \ t \ge 2L\widetilde{K}. \end{cases} \tag{6}$$

For recursive sequences general limit theorems have been established by means of the contraction method. The following is a typical result for this approach (see Rösler (2001) or Neininger and Rüschendorf (2004a)). We consider convergence of X_n without further normalization.

Theorem 2.3 (Convergence theorem) *Let (X_n) be a recursive sequence as in (1), square-integrable, and with expectation 0. We assume that*

stabilization: $(A_1(n), \dots, A_M(n), b_n) \to (A_1, \dots, A_M, b)$ *in* L^2

contraction: $E\sum_r A_r^2 < 1$

nondegeneracy: $E1_{\{I_r^{(n)} \le \ell\}}(A_r(n))^2 \xrightarrow[n\to\infty]{} 0, \quad \forall \ell \in \mathbb{N}, \quad 1 \le r \le M.$

Then $\ell_2(X_n, X^*) \to 0$ as $n \to \infty$, where X^* is the unique solution of the fixed point equation (2) in the class $M_{2,0}$ of all distributions with finite second moment and expectation 0 *(as in Theorem 2.1).*

The convergence theorem does in general not imply that for large n X_n has a similar tail behaviour as the limit X. Some boundedness condition however allows to derive uniform convergence of the Laplace transforms, which can be used to derive exponential tails of (X_n) (see Remark 2.6).

Proposition 2.4 (Convergence of Laplace transforms) *Let (X_n) be a sequence of r.v.s with $X_n \xrightarrow{d} X$ and let $E\exp(\pm L(1+\varepsilon)X_n) \le K$, $\forall n \in \mathbb{N}_0$, and $E\exp(\pm L(1+\varepsilon)X) \le K$ for some $L > 0, \varepsilon > 0$, then*

$$E\exp(\lambda X_n) \to E\exp(\lambda X) \quad \text{as } n \to \infty \tag{7}$$

uniformly for $|\lambda| \le L$.

Proof: Without loss of generality let X_n, X be a.s. convergent versions $X_n \to X$ a.s. for $n \to \infty$. Then we prove $E|\exp(\lambda X_n) - \exp(\lambda X)| \to 0,\ n \to \infty$ uniformly for $|\lambda| \le L$.
To prove uniformity of the convergence for $|\lambda| \le L$ observe that

$$e^{\lambda X} = e^{\lambda X^+} + e^{-\lambda X^-} - 1 \tag{8}$$

and, therefore, by the triangle inequality

$$E|e^{\lambda X_n} - e^{\lambda X}| \leq E|e^{\lambda X_n^+} - e^{\lambda X^+}| + E|e^{-\lambda X_n^-} - e^{-\lambda X^-}|.$$

Thus we can restrict to the case $X_n \geq 0$, $X \geq 0$ since $Ee^{\lambda X_n^+} \leq Ee^{\lambda X_n} + 1$ and $Ee^{\lambda X_n^+} \leq 1$ for $\lambda < 0$ and similarly for X_n^-, X^-. For $0 \leq \lambda \leq L$ and $X_n \geq 0$, $X \geq 0$ holds

$$E|e^{\lambda X_n} - e^{\lambda X}| \leq E|e^{LX_n} - e^{LX}|$$

by monotonicity in λ of the integrand. Further, for $\lambda > 0$

$$E|e^{-\lambda X_n} - e^{-\lambda X}| = E\frac{|e^{\lambda X} - e^{\lambda X_n}|}{e^{\lambda(X_n + X)}} \leq E|e^{\lambda X_n} - e^{\lambda X}| \leq E|e^{LX_n} - e^{LX}|.$$

Now following the proof of Theorem 4.5 in Rösler (1991) for w.l.g. $X_n, X \geq 0$ and $N > 0$ we have

$$\begin{aligned} E\Big|e^{\lambda X_n} - e^{\lambda X}\Big| \;&\leq\; E\Big|e^{LX_n} 1_{(X_n \leq N)} - e^{LX} 1_{(X \leq N)}\Big| \\ &\quad + Ee^{LX_n} 1_{(X_n > N)} + Ee^{LX} 1_{(X > N)}. \end{aligned}$$

By the bounded convergence theorem the first term on the right-hand side converges towards zero. For the second term we use Hölder's inequality to obtain

$$Ee^{LX_n} 1_{(X_n > N)} \leq \Big(Ee^{L(1+\varepsilon)X_n}\Big)^{\frac{1}{1+\varepsilon}} \left(P(X_n > N)\right)^{\frac{1+\varepsilon}{\varepsilon}} \to 0$$

for $N \to \infty$, since the first factor on the right-hand side is uniformly bounded and the second converges towards zero. Similarly we obtain $Ee^{LX} 1_{(X>N)} \to 0$ as $N \to \infty$. □

In the following theorem we establish some general conditions on the coefficients which imply exponential bounds for recurrences as in (1).

Theorem 2.5 (Laplace transform, exponential bounds)
Let (X_n) be a recursive sequence as in (1) with $EX_n = 0$, $n \in \mathbb{N}_0$. Assume that for some constant $\widetilde{K}$ and for λ in some open interval $U = U(0)$ around zero $Ee^{\lambda X_i} \leq e^{\widetilde{K}\lambda^2}$, $0 \leq i \leq n_0 - 1$. Further assume that for some sequence $(t_n)_{n \geq n_0}$ with $0 < t_n < 1$

a) $Ee^{\lambda b_n} \leq \widetilde{C}$, $Eb_n^2 \leq \widetilde{C}t_n$, $\forall n \geq n_0$, *for some finite $\widetilde{C}$, and $\lambda \in U$,*

b) $E\sum_r (A_r(n))^2 < 1 - t_n$, $\forall n \geq n_0$

c) $\sup_r |A_r(n)| \leq 1$, $n \geq n_0$.

Then

$$Ee^{\lambda X_m} \leq e^{\overline{K}\lambda^2} \tag{9}$$

for all $m \in \mathbb{N}_0$ and for all λ in some open interval $V = V(0)$ around zero, where $\overline{K}$ is a finite constant.

Proof: The proof is given by induction. The assertion is true for $m \le n_0 - 1$ by assumption with $\widetilde{K} = \bar{K}$. For the induction step we consider the case $n \ge n_0$ and assume that (9) is true for $m \le n-1$. Define the function $f_K^{(n)}(\lambda) := Ee^{\lambda b_n + K\lambda^2(\sum_{r=1}^M (A_r(n))^2 - 1)}$ for some positive constant K and $n \ge n_0$. We will write $(A(n), I^{(n)}, b_n)$ for short for the vector $(A_1(n), \dots, A_M(n), I_1^{(n)}, \dots, I_M^{(n)}, b_n)$. Since $X_n \stackrel{d}{=} \sum_{r=1}^M A_r(n) X_{I_r^{(n)}} + b_n$, we obtain by the induction hypothesis and boundedness of $A_r(n)$

$$\begin{aligned} Ee^{\lambda X_n} &= E\left(e^{\lambda b_n} \prod_{r=1}^M E\left(e^{\lambda A_r(n) X^{(r)}_{I_r^{(n)}}} \mid \left(A(n), I^{(n)}, b^{(n)}\right)\right)\right) \\ &\le e^{K\lambda^2} f_K^{(n)}(\lambda), \quad \lambda \in U. \end{aligned}$$

Therefore we have proven the induction step if we can show that for some $\bar{K} \ge \widetilde{K}$, independent of n, holds:

$$f_{\bar{K}}^{(n)}(\lambda) \le 1 \quad \text{for } |\lambda| < L \text{ for some } L > 0. \tag{10}$$

By differentiation we obtain for $\bar{K} \ge \widetilde{K}$ large enough

$$\begin{aligned} f_{\bar{K}}^{(n)}(0) &= 1, \\ \frac{d}{d\lambda} f_{\bar{K}}^{(n)}(\lambda) &= E\left(e^{\lambda b_n + \bar{K}\lambda^2\left(\sum_{r=1}^M A_r^2(n) - 1\right)} \left(b_n + 2\lambda\bar{K}\Big(\sum_{r=1}^M A_r^2(n) - 1\Big)\right)\right), \\ \frac{d}{d\lambda} f_{\bar{K}}^{(n)}(0) &= 0, \\ \frac{d^2}{d\lambda^2} f_{\bar{K}}^{(n)}(\lambda) &= E\Bigg(e^{\lambda b_n + \bar{K}\lambda^2\left(\sum_{r=1}^M A_r^2(n) - 1\right)} \\ &\qquad \cdot \left(\left(b_n + 2\lambda\bar{K}\Big(\sum_{r=1}^M A_r^2(n) - 1\Big)\right)^2 + 2\bar{K}\Big(\sum_{r=1}^M A_r^2(n) - 1\Big)\right)\Bigg), \\ \frac{d^2}{d\lambda^2} f_{\bar{K}}^{(n)}(0) &= Eb_n^2 + E\left(2\bar{K}\Big(\sum_{r=1}^M A_r^2(n) - 1\Big)\right) \\ &\le \widetilde{C} t_n - 2\bar{K} t_n < 0 \end{aligned}$$

using condition b) and noting that $EX_n = 0, \forall n \in \mathbb{N}_0$, implies $Eb_n = 0, \quad \forall n \ge n_0$. We also note that $\bar{K} \ge \max\{\widetilde{K}, \widetilde{C}\}$ can be chosen independently of n.

It remains to show $f_{\bar{K}}^{(n)}(\lambda) \le 1$ for all λ in some open interval $V = V(0)$. For the proof we establish that $\frac{d^3}{d\lambda^3} f_{\bar{K}}^{(n)}(\lambda)$ is bounded in some neighbourhood of zero independently of n. Therefore consider

$$\begin{aligned} \frac{d^3}{d\lambda^3} f_{\bar{K}}^{(n)}(\lambda) &= E\Bigg(e^{\lambda b_n + \bar{K}\lambda^2\left(\sum_{r=1}^M A_r^2(n) - 1\right)} \left(b_n + 2\lambda\bar{K}\Big(\sum_{r=1}^M A_r^2(n) - 1\Big)\right) \\ &\qquad \cdot \left(\left(b_n + 2\lambda\bar{K}\Big(\sum_{r=1}^M A_r^2(n) - 1\Big)\right)^2 + 6\bar{K}\Big(\sum_{r=1}^M A_r^2(n) - 1\Big)\right)\Bigg). \end{aligned}$$

By the uniform boundedness of the Laplace transform of b_n and assumptions b) and c), we obtain that this expression is bounded for $\bar{K}$ fixed and $|\lambda| \leq K_0$ independently of n. It follows that the second derivative is less than 0 in an open interval around zero uniformly in n and therefore $f_{\bar{K}}^{(n)}(\lambda) \leq 1$ for all λ in some open interval $V = V(0)$.

The neighbourhood of zero thus can be chosen independently of n. Thus we obtain some constant L such that (10) holds. $\square$

Remark 2.6 *In the case that additionally to the assumptions in Theorem 2.5 the conditions of the convergence theorem hold a direct proof of the exponential bounds in Theorem 2.5 is possible without referring to the induction argument in the proof of Theorem 2.5 but using instead the convergence of the Laplace transforms in Proposition 2.4 applied to* $f_K^{(n)}(\lambda) = E\exp(\lambda b_n + K\lambda^2(\sum A_r^2(n) - 1))$ *as well as to its first and second derivative. This approximation argument allows to derive from the limiting case the crucial bound* $f_K^{(n)}(\lambda) \leq 1$ *for* λ *in a neighbourhood of zero and uniformly for all* $n \geq N$*. Together with the corresponding inequality for the finitely many indices* $n = 1, \ldots, N$ *in a possibly smaller neighbourhood the exponential inequality follows.*

In the following theorem we derive exponential bounds for $|\lambda| < L$ for any given $L > 0$. In order to obtain this stronger conclusion we strengthen the condition on the coefficients $A_r(n)$. For the purpose of ease of application we formulate this result for the normalized version (Y_n) of the recursive sequence (X_n). Let (X_n) be a recursive sequence as in (1) and denote by

$$Y_n = \frac{X_n - EX_n}{s_n} \tag{11}$$

a normalized version of X_n with nondecreasing scalings $s_n = s(n) > 0$. We define $\mu_n := EX_n, n \in \mathbb{N}_0$. Then (Y_n) satisfies the recursive equation

$$Y_n \stackrel{d}{=} \sum_{r=1}^{M} A_r(n) \frac{s(I_r^{(n)})}{s_n} Y_{I_r^{(n)}}^{(r)} + b^{(n)}, \quad n \geq n_0 \tag{12}$$

where

$$b^{(n)} = \frac{1}{s_n}\Big(b_n - \mu_n + \sum_{r=1}^{M} A_r(n)\mu_{I_r^{(n)}}\Big).$$

Thus (Y_n) again is a recursive sequence of the form (1) with modified coefficients $b^{(n)}$ and $A_r^{(n)} := A_r(n)\frac{s(I_r^{(n)})}{s_n}$.

The following theorem gives under stronger conditions on the coefficients compared to Theorem 2.5 a stronger result on exponential bounds for the Laplace transform of the scaled version (Y_n) of (X_n) valid on any prescribed finite interval $[-L, L]$.

Theorem 2.7 (Laplace transforms of the normalized sequence)
Let (X_n) *satisfy the recurrence (1) with normalized version* (Y_n) *as in (11). Assume that* $Ee^{\lambda Y_i} \leq e^{\widetilde{K}\lambda^2}$, $0 \leq i \leq n_0 - 1$, $\lambda \in \mathbb{R}^1$*. Further assume that for some sequence* $(t_n)_{n \geq n_0}$ *with* $0 < t_n < 1$,

a) $Ee^{\lambda b^{(n)}} \leq \widetilde{C}$, $\forall \lambda \in \mathbb{R}^1$, $E(b^{(n)})^2 \leq \widetilde{C} t_n$, $\forall n \geq n_0$, *for some finite constant* $\widetilde{C}$.

b) $U_n := \sum_{r=1}^{M} \big(A_r^{(n)}\big)^2 - 1 \le 0$ *and* $EU_n \le -t_n < 0$*,* $n \ge n_0$*,*

c) $\lim_{K\to\infty} \sup_{n\ge n_0} E\exp(KU_n) = 0$.

Then for all $L > 0$ *there exists a finite constant* K_L *such that for all* $n \in \mathbb{N}_0$ *and* $\lambda \in [-L, L]$

$$Ee^{\lambda Y_n} \le e^{\lambda^2 K_L}. \tag{13}$$

Proof: Let $f(K, \lambda, n) := E\exp(\lambda b^{(n)} + \lambda^2 KU_n)$, $K > 0$, $\lambda \in \mathbb{R}$, $n \ge n_0$. We establish in the following by induction that for any $L > 0$ there exists a K_L such that

$$f(K_L, \lambda, n) \le 1, \quad \forall\lambda \in [-L, L], \forall n \ge n_0. \tag{14}$$

Step 1: There exist constants $K_1 > 0$ and $L_1 > 0$, such that $f(K_1, \lambda, n) \le 1$, $\forall\lambda \in [-L_1, L_1]$, $\forall n \ge n_0$.

For the proof of step 1 we note that $f(K, \lambda, n)$ is well defined and twice differentiable (as shown in the proof of Theorem 2.1 of Rösler (1992)) and it holds

$$\frac{d}{d\lambda} f(K, \lambda, n)\big|_{\lambda=0} = Eb^{(n)} = 0 \quad \text{as } EY_n = 0.$$

Further

$$\frac{d^2}{d\lambda^2} f(K, \lambda, n)\big|_{\lambda=0} = E(b^{(n)})^2 + 2KEU_n \le \sup_{n\ge n_0} E(b^{(n)})^2 - 2Kt_n \le t_n(\widetilde{C} - 2K).$$

Choosing $K_1 > 0$ with $\frac{d^2}{d\lambda^2} f(K_1, \lambda, n)|_{\lambda=0} < 0$, then $f(K_1, \lambda, n)$ has a strict local maximum in $\lambda = 0$ for all $n \ge n_0$. Thus $f(K_1, \cdot, n) \le 1$ in a neighbourhood $U = U(0)$ of zero. Since $\frac{d^2}{d\lambda^2} f(K_1, \lambda, n) < 0$ in some neighbourhood $V = V(0)$ independent of n, the neighbourhood U can be chosen also independently of n (similar as is shown in the proof of Theorem 2.5).

Step 2: For $L > 0$ there exists $K_2 > 0$ such that $f(K_2, L, n) \le 1$, $f(K_2, -L, n) \le 1$ for all n.

For the proof of step 2 note that by assumption b)

$$\frac{d}{dK} f(K, \lambda, n) = E\lambda^2 U_n \exp(\lambda b^{(n)} + \lambda^2 KU_n) \le 0$$

and thus $f(K, \lambda, n)$ is monotonically nonincreasing in K for λ, n fixed.
By assumption c) $\lim_{K\to\infty} \sup_n f(K, \lambda, n) = 0$ for all $\lambda \ne 0$; thus step 2 follows.

To prove (14) we define $K_L := \max\{K_1, \frac{L^2}{L_1^2} K_2\}$. By step 1 and monotonicity in K holds $f(K_L, \lambda, n) \le 1$ for $|\lambda| \le L_1$. For $L_1 \le |\lambda| \le L$ define $\alpha := \frac{L}{|\lambda|} \ge 1$, then

$$\begin{aligned} f(K_L, \lambda, n) &\le \Big(E\exp\big(\alpha(\lambda b^{(n)} + \lambda^2 K_L U_n)\big)\Big)^{1/\alpha} \\ &\le \Big(E\exp\big(Lb^{(n)} + \frac{L_1^2 L^2}{L_1^2} K_2 U_n\big)\Big)^{1/\alpha} \\ &\le 1 \qquad \text{by step 2.} \end{aligned}$$

Now we prove (13) by induction.

By assumption $E\exp(\lambda Y_k) \leq \exp(\lambda^2 \widetilde{K})$, $k = 0, \dots, n_0 - 1$. For the induction step define $\widetilde{K}_L = \max\{\widetilde{K}, K_L\}$. Then using the induction hypothesis

$$\begin{aligned} E\exp(\lambda Y_n) &= E\exp\left(\lambda\Big(\sum_{r=1}^{M} A_r(n)\frac{s(I_r^{(n)})}{s_n} Y^{(r)}_{I_r^{(n)}} + b^{(n)}\Big)\right) \\ &\leq \exp(\lambda^2 \widetilde{K}_L) E\exp(\lambda^2 \widetilde{K}_L U_n + \lambda b^{(n)}) \\ &\leq \exp(\lambda^2 \widetilde{K}_L) f(\widetilde{K}_L, \lambda, n) \\ &\leq \exp(\lambda^2 \widetilde{K}_L), \qquad \forall \lambda \in [-L, L]. \end{aligned}$$

□

Remark 2.8 *The conditions of Theorems 2.5 and 2.7 are in many examples not difficult to check. The Quicksort algorithm satisfies the recurrence (1) in the form*

$$X_n \stackrel{d}{=} X^{(1)}_{\lfloor nU \rfloor} + X^{(2)}_{\lfloor n(1-U) \rfloor} + C_n, \tag{15}$$

where $C_n = n - 1$. The normalized version $Y_n = \frac{X_n - EX_n}{n}$ converges to the Quicksort distribution and the exponential bounds of Theorem 2.5 are established by Rösler (1991) in this case. For Quicksort type recursions with a monomial toll term $C_n = n^\alpha$ exponential bounds have been established in Neininger (2002). For several further examples of a similar type as (15) which fit our conditions, see Hwang and Neininger (2002).

3 Subgaussian behaviour of recurrences

A random variable X has subgaussian right (left) tails if for some constants $L > 0, C > 0$

$$E\exp(\lambda X) \leq C\exp(L\lambda^2), \quad \forall \lambda > 0 \ (\forall \lambda < 0). \tag{16}$$

X has subgaussian tails if it has subgaussian right and left tails. In this section we aim to sharpen the bounds in Section 2 to imply that a recurrence (X_n) or its normalized form $Y_n = \frac{X_n - EX_n}{s_n}$ has subgaussian tails, i.e. it holds independently of $n \in \mathbb{N}_0$:

$$E\exp(\lambda Z_n) \leq \exp(L\lambda^2), \quad \forall \lambda \in \mathbb{R}, \forall n \in \mathbb{N}_0, \tag{17}$$

where $Z_n = X_n$ resp. $Z_n = Y_n$.

As consequence this implies improved exponential bounds compared to (6), (9), and (13)

$$P(|X_n - EX_n| \geq t|EX_n|) \leq 2\exp\left(-\frac{t^2}{4L}\Big(\frac{EX_n}{s_n}\Big)^2\right) \text{ for all } t > 0. \tag{18}$$

We assume that the normalizing sequence s_n is nondecreasing, $s_n \geq \varepsilon > 0$, $\forall n$ and that s_n^2 is the order of magnitude of the expectation:

E) $\qquad \mu_n := EX_n = \mu s_n^2 + r_n$ for some $\mu > 0$

where $r_n = o(s_n^2)$ is the remainder term.

The normalized sequence (Y_n) satisfies the recursive equation

$$Y_n \stackrel{d}{=} \sum_{r=1}^{M} A_r(n) \frac{s(I_r^{(n)})}{s_n} Y^{(r)}_{I_r^{(n)}} + b^{(n)} \tag{19}$$

with

$$\begin{aligned} b^{(n)} &= \frac{1}{s_n}\Big(b_n - \mu_n + \sum_{r=1}^{M} A_r(n)\mu_{I_r^{(n)}}\Big) \\ &= \frac{1}{s_n}\Big(b_n - \mu\Big(s_n^2 - \sum_{r=1}^{M} A_r(n) s(I_r^{(n)})^2\Big) + R_n\Big) \end{aligned}$$

where $R_n = -r_n + \sum_{r=1}^{M} A_r(n) r(I_r^{(n)}),\ r(k) := r_k$. Thus the bounds in Section 2 apply under corresponding conditions on the modified coefficients $b^{(n)}$, $A_r^{(n)} := A_r(n)\frac{s(I_r^{(n)})}{s_n}$.

The following result on subgaussian tails extends Theorem 1 in Neininger (2005), who considers the special case where $s_n^2 = n$, $A_r(n) = 1$ and $r_n = O(1)$. In comparison to the results in Section 2 we postulate some stronger assumptions. In particular we assume boundedness of b_n where the bounds are of order $O(r_n)$, further we strengthen condition b) of Theorem 2.5 resp. conditions b) and c) of Theorem 2.7.

Theorem 3.1 *Let (X_n) be a recursive sequence as in (1) satisfying condition E) and assume that for some $Q > 0$, $Ee^{\lambda Y_i} \le \exp(Q\lambda^2)$, $0 \le i \le n_0 - 1$, $\lambda \in \mathbb{R}^1$. Further assume*

a) $\quad \|b_n\|_\infty \le M_1 \bar{r}_n, \quad \bar{r}_n := \max\{r_n, 1\}$, *for $n \ge n_0$ and some $M_1 \ge 0$*

b)
$$|\mu_n - \sum_{r=1}^{M} A_r(n)\mu_{I_r^{(n)}}| \le C\bar{r}_n, \quad \text{for } n \ge n_0 \text{ and for some } C \ge 0 \tag{20}$$

c)
$$E\exp\left(\frac{1}{\bar{r}_n^2}\Big(\sum_{r=1}^{M} A_r^2(n) s(I_r^{(n)})^2 - s_n^2\Big)\right)^x \le \eta^x, \ \text{for some } \eta < 1,\ x \ge 0,\ n \ge n_0. \tag{21}$$

Then (Y_n) is subgaussian, i.e. (17) holds.

Proof: By assumption we have

$$E\exp(\lambda Y_j) \le \exp(Q\lambda^2), \quad j = 0, \ldots, n_0{-}1 \tag{22}$$

for some $Q > 0$. We next establish by induction that for some finite constant $L < \infty$

$$E\exp(\lambda Y_j) \le \exp(L\lambda^2), \qquad \forall j \ge 0, \forall \lambda \in \mathbb{R}. \tag{23}$$

Let (23) be true for $j \le n-1$ and let Υ_n denote the distribution of $(I^{(n)}, b^{(n)}, A(n))$. Then by conditioning and by the induction hypothesis we obtain

$$\begin{aligned}
E\exp(\lambda Y_n) &= \int E\exp\Big(\lambda \sum_{r=1}^M a_r \frac{s(j_r)}{s(n)} Y_{j_r}^{(r)} + \lambda\beta\Big) d\Upsilon_n(j,\beta,a) \\
&\le \int E\exp\Big(L\lambda^2 \sum_{r=1}^M a_r^2 \Big(\frac{s(j_r)}{s(n)}\Big)^2 + \lambda\beta\Big) d\Upsilon_n(j,\beta,a) \\
&= \exp(L\lambda^2) E\exp\Big(L\lambda^2\Big(\sum_{r=1}^M A_r^2(n)\Big(\frac{s(I_r^{(n)})}{s(n)}\Big)^2 - 1\Big) + \lambda b^{(n)}\Big) \\
&\le E\exp\Big(2L\lambda^2 \frac{\bar r_n^2}{s_n^2}\frac{1}{\bar r_n^2}\Big(\sum_{r=1}^M A_r^2(n) s^2(I_r^{(n)}) - s_n^2\Big)\Big)^{\frac12} E\exp(2\lambda b^{(n)})^{\frac12}\exp(L\lambda^2) \\
&= E\exp\Big(\frac{1}{\bar r_n^2}\Big(\sum_{r=1}^M A_r^2(n) s^2(I_r^{(n)}) - s_n^2\Big)\Big)^{L\lambda^2\frac{\bar r_n^2}{s_n^2}} E\exp(\lambda b^{(n)})\exp(L\lambda^2).
\end{aligned}$$

Using (20) we obtain

$$\|b^{(n)}\|_\infty \le \frac{1}{s_n}(\|b_n\|_\infty + C\bar r_n) \le \frac{M_2 \bar r_n}{s_n}, \quad M_2 := M_1 + C.$$

Next we apply Hoeffding's inequality to $b^{(n)}$: If X is a random variable with $a \le X \le b$ and $EX = 0$, then for all $\lambda \in \mathbb{R}$ it holds (see e.g. Lugosi (2005), Hoeffding (1963))

$$E\exp(\lambda X) \le \exp\Big(\frac{(b-a)^2\lambda^2}{8}\Big).$$

As consequence we obtain

$$\begin{aligned}
E\exp(\lambda Y_n) &\le \eta^{L\lambda^2\frac{\bar r_n^2}{s_n^2}} \exp\Big(\frac{(M_2\bar r_n\lambda)^2}{2s_n^2}\Big)\exp(L\lambda^2) \\
&= \exp\Big(L(\log\eta)\lambda^2\frac{\bar r_n^2}{s_n^2} + \frac{M_2^2\bar r_n^2}{2s_n^2}\lambda^2\Big)\exp(L\lambda^2) \\
&\le \exp(L\lambda^2)
\end{aligned}$$

if we choose $2L \ge \frac{M_2^2}{\log\frac{1}{\eta}}$ and large enough such that it holds true also for $n = 0,\dots,n_0-1$. □

Remark 3.2 *a) Condition (20) can be replaced by the following conditions: Let for some generic constants* C_1, C_2

A1)
$$
\begin{cases}
\left|s_n^2 - \sum_{r=1}^{M} A_r(n) s^2(I_r^{(n)})\right| \le C_1 \bar{r}_n & (24)\\
\sup_{k\le n} |r_k| \le C_2 \bar{r}_n, \quad \textit{and} & \\
\sup_r |A_r(n)| \le C_2. & (25)
\end{cases}
$$

In the particular case $r_n = O(1)$ we thus get the following sufficient conditions for the conclusion of Theorem 3.1:

A2)
$$
\begin{cases}
\sup_{n\ge n_0} \|b_n\|_\infty < \infty, \quad \sup_{\substack{n\ge n_0\\ r\ge 1}} \|A_r(n)\|_\infty < \infty & \\
1 \le s_n^2 - \sum_{r=1}^{M} A_r(n) s^2(I_r^{(n)}) & (26)\\
\left|s_n^2 - \sum_{r=1}^{M} A_r(n) s^2(I_r^{(n)})\right| \le C_1. & (27)
\end{cases}
$$

If e.g. X_n denotes the number of leaves in a random binary tree, then (X_n) satisfies the recursive equation

$$X_n \stackrel{d}{=} X_{I_1^{(n)}} + X_{n-1-I_1^{(n)}}, \quad X_0 = 0, X_1 = 1 \tag{28}$$

and $I_1^{(n)} \stackrel{d}{=} \mathrm{unif}\{0,\dots,n-1\}$ the uniform distribution on $\{0,\dots,n-1\}$.

The expectation has the expansion

$$EX_n = \frac{n+1}{3} = \frac{n}{3} + O(1). \tag{29}$$

The conditions A2) are satisfied with $s_n = \sqrt{n}$, $\mu = \frac{1}{3}$, $r_n = O(1)$, $A_r(n) = 1$, and thus we get subgaussian behaviour. This example fits also the conditions of Neininger's Theorem 1 2005.

Further similar examples fit in this scheme like the size of m-ary search trees, $m \le 13$, several parameters of binary search trees with bounded toll functions and some others (see Hwang and Neininger (2002), Neininger (2002), and Neininger and Rüschendorf (2004a)).

For the size of m-ary search trees subgaussian behaviour is known to hold for $m \le 26$ as follows from the moment expansions in Chern and Hwang (2001) and Lemma 3.3. But as mentioned in the introduction our methods do not suit well to this normal case and thus we get subgaussian behaviour due to the special properties of the expansion of the mean in this case for $m \le 13$ only.

b) If the condition $s_n > 0$ does not hold true for all $n \ge 0$ as in the case $s_n = n$ or $s_n = \log n$, then we may choose $s_\varepsilon(n) := \max(\varepsilon, s_n)$ and obtain with this modification analogous bounds.

An example for this situation are recursive equations of the form

$$X_n \stackrel{d}{=} X_{I_n} + b_n, \quad n \ge n_0, \tag{30}$$

where $s_n^2 = \log n$, $n \geq 2$, $r_n = O(1)$. It suffices in that case to assume

$$\sup_{n \geq n_0} E \log^2 \Big(\frac{I_n \vee 1}{n}\Big) < \infty, \tag{31}$$

$$E\Big(\frac{I_n \vee 2}{n}\Big)^k \leq \eta^k, \quad k \geq 1, n \geq n_0. \tag{32}$$

Recursions of the type (30) *are considered in Neininger (2005, Theorem 2). Conditions (31) and (32) are not fulfilled for $I_n \sim \mathrm{unif}\{0, \dots, n-1\}$ but fit to more concentrated subgroup sizes I_n where e.g. $I_n \sim \mathfrak{B}(n-1, \vartheta)$.*

The following characterization of subgaussian right tails allows to dismiss with the boundedness of the toll terms b_n in Theorem 3.1. The conditions on the weight coefficients $A_r(n)$ in Theorem 3.1 and in the following Theorem 3.4 are however not comparable. For the proof we shall make use of the following lemma which can be found in Lugosi (2005) (up to a slight modification).

Lemma 3.3 *Let X be a real random variable.*

1. *If for some constant $c > 0$ holds $(EX_+^k)^{1/k} \leq \sqrt{ck}$ for all $k \in \mathbb{N}$, then*

$$Ee^{\lambda X} \leq \sqrt{2} e^{1/6} e^{ce\lambda^2/2}, \quad \text{for all } \lambda > 0. \tag{33}$$

2. *If for some $c, C > 0$ we have $Ee^{\lambda X} \leq ce^{C\lambda^2}$, $\forall \lambda > 0$, then there exists an universal constant K, so that for all $k \in \mathbb{N}$*

$$(EX_+^k)^{1/k} \leq c^{1/k} K \sqrt{Ck}.$$

Theorem 3.4 *Let (X_n) be a recursive sequence as in (1) with E) and with normalized sequence $Y_n = \frac{X_n - EX_n}{s_n}$. We assume the following conditions:*

a) $(E|Y_i|^k)^{1/k} \leq \sqrt{ck}$, *for some $c > 0$, $i = 0, \dots, n_0-1$, $k \in \mathbb{N}$,*

b) $E\Big|\dfrac{b_n}{s_n}\Big|^k \leq (\sqrt{c_b k}\,)^k$ *for some $c_b > 0$, $k \in \mathbb{N}$, $n \geq n_0$,*

c)
$$\sum_{r=1}^M \Big|\frac{s(I_r^{(n)})}{s_n} A_r(n)\Big| \leq 1 - \frac{1}{\eta}, \quad n \geq n_0 \quad \text{for some } \eta > 1, \tag{34}$$

d) $\Big|\mu_n - \sum_{r=1}^M A_r(n) \mu_{I_r^{(n)}}\Big| \leq C\widetilde{r}_n$, *where $\widetilde{r}_n = O(s_n)$.*

Then for all $t > 0$

$$P(|Y_n| \geq t) \leq 2\sqrt{2} e^{1/6} \exp\Big(-\frac{t^2}{2e\widetilde{c}}\Big),$$

with $\widetilde{c} := \max\{c, \eta^2(\sqrt{c_b} + \bar{C})^2\}$ where $\bar{C} := \sup_n \{C \frac{\widetilde{r}_n}{s_n}\}$.

Proof: We prove, that for some constant $\widetilde{c}$

$$(E|Y_n|^k)^{1/k} \le \sqrt{\widetilde{c}k}, \qquad \forall k \in \mathbb{N} \text{ and } n \ge n_0. \tag{35}$$

Using the definition of $b^{(n)}$ in (19) and condition d) we estimate

$$\begin{aligned}
E|b^{(n)}|^k &\le E\Big|\frac{b_n}{s_n} + C\frac{\widetilde{r}_n}{s_n}\Big|^k \\
&\le \sum_{\ell=0}^{k} \binom{k}{\ell} E\Big|\frac{b_n}{s_n}\Big|^\ell \Big(C\frac{\widetilde{r}_n}{s_n}\Big)^{k-\ell} \\
&\le \sum_{\ell=0}^{k} \binom{k}{\ell} \Big(\sqrt{c_b \ell}\Big)^\ell \Big(C\frac{\widetilde{r}_n}{s_n}\Big)^{k-\ell} \\
&\le \Big(\sqrt{c_b k} + C\frac{\widetilde{r}_n}{s_n}\Big)^k \le \Big(\sqrt{\widetilde{c}_b k}\Big)^k, \text{where } \widetilde{c}_b = (\sqrt{c}_b + \bar{C})^2.
\end{aligned}$$

Because of condition a) we have $(E|Y_i|^k)^{1/k} \le \sqrt{ck}$, $k \in \mathbb{N}$, $i = 0, \dots, n_0 - 1$. For the induction step now assume that (35) holds for $Y_{n_0}, \dots, Y_{n-1}$ with a suitable constant $\widetilde{c} \ge \max\{c, \eta^2 \widetilde{c}_b\}$ (therefore $(E|Y_i|^k)^{1/k} \le \sqrt{\widetilde{c}k}$, $i = 0, \dots, n_0 - 1$). Using the conditioning argument under $(A(n), I^{(n)}, b^{(n)})$ and with $\Upsilon_n = P^{(A(n), I^{(n)}, b^{(n)})}$ it holds

$$\begin{aligned}
E|Y_n|^q &= E\Big|\sum_{r=1}^{M} A_r(n)\frac{s(I_r^{(n)})}{s(n)} Y^{(r)}_{I_r^{(n)}} + b^{(n)}\Big|^q \\
&\le \int E\Big(\sum_{r=1}^{M}\Big|\frac{s(i_r)}{s_n}a_r\Big|\big|Y^{(r)}_{i_r}\big| + |\beta|\Big)^q d\Upsilon_n(a,i,\beta) \\
&= \int \sum_{\ell=0}^{q}\binom{q}{\ell}\Big(\sum_{r=1}^{M}\Big|\frac{s(i_r)}{s_n}a_r\Big|E\big|Y^{(r)}_{i_r}\big|\Big)^\ell |\beta|^{q-\ell} d\Upsilon_n(a,i,\beta) \\
(*) \qquad &\le \int \sum_{\ell=0}^{q}\binom{q}{\ell}\sqrt{\widetilde{c}^\ell} \sum_{j_1+\ldots+j_M=\ell}\binom{\ell}{j_1 \dots j_M} \\
&\quad \times \prod_{r=1}^{M}\Big|\sqrt{j_r}^{j_r}\Big(\frac{s(i_r)}{s_n}a_r\Big)^{j_r}\Big| |\beta|^{q-\ell} d\Upsilon_n(a,i,\beta) \\
&\le \int \sum_{\ell=0}^{q}\binom{q}{\ell}\sqrt{\widetilde{c}^\ell}\sqrt{\ell^\ell}\Big(\sum_{r=1}^{M}\frac{s(i_r)}{s_n}|a_r|\Big)^\ell |\beta|^{q-\ell} d\Upsilon_n(a,i,\beta) \\
&\le \sum_{\ell=0}^{q}\binom{q}{\ell}\sqrt{\widetilde{c}^\ell}\sqrt{\ell^\ell}\Big(1-\frac{1}{\eta}\Big)^\ell E|b^{(n)}|^{q-\ell} \\
&\le \sum_{\ell=0}^{q}\binom{q}{\ell}\sqrt{\widetilde{c}^\ell}\sqrt{\ell^\ell}\Big(1-\frac{1}{\eta}\Big)^\ell \Big(\frac{1}{\eta}\Big)^{q-\ell} \sqrt{\eta^2\widetilde{c}_b(q-\ell)}^{\,q-\ell}
\end{aligned}$$

$$\leq \sum_{\ell=0}^{q} \binom{q}{\ell} \sqrt{\widetilde{c}^{\ell}} \sqrt{q^{\ell}} \left(1-\frac{1}{\eta}\right)^{\ell} \left(\frac{1}{\eta}\right)^{q-\ell} \sqrt{\widetilde{c}q}^{\,q-\ell}$$
$$\leq \sqrt{\widetilde{c}^{q}} \sqrt{q^{q}}.$$

where $(*)$ follows from

$$E\left(\sum_{r=1}^{M} \frac{s(i_r)}{s_n} |a_r| |Y_{i_r}^{(r)}|\right)^{\ell} = \sum_{j_1+\ldots+j_M=\ell} \binom{\ell}{j_1 \ldots j_M} \prod_{r=1}^{M} \left(\frac{s(i_r)}{s_n}|a_r|\right)^{j_r} E|Y_{i_r}^{(r)}|^{j_r}$$
$$\leq \sqrt{\widetilde{c}^{\ell}} \sum_{j_1+\ldots+j_M=\ell} \binom{\ell}{j_1, \ldots j_M} \prod_{r=1}^{M} \sqrt{j_r}^{\,j_r} \left(\frac{s(i_r)}{s_n}|a_r|\right)^{j_r}$$

as $i_r < n$. This implies

$$(E|Y_n|^q)^{1/q} \leq \sqrt{\widetilde{c}q} \ \text{ with } \widetilde{c} := \max\{c, \eta^2 \widetilde{c}_b\}.$$

Thus by Lemma 3.3 Y_n has subgaussian tails and therefore similar to (18) with Chernoff's bounding technique we obtain the assertion. ☐

Under somewhat weaker growth conditions on the moments of $\frac{b_n}{s_n}$ than those in Theorem 3.4 we still obtain exponential tails of linear order instead of subgaussian tails. The proof is based on Bernstein's inequality which we state for the ease of reference.

Lemma 3.5 (Bernstein's inequality) *Let $X_1, \ldots, X_n$ be independent r.v.s, $S_n = \sum_{i=1}^n X_i$ with $EX_i = 0$, $\sigma_i^2 := EX_i^2 \leq L_i^2 < \infty$ and $\sigma^2 := ES_n^2 = \sum_{i=1}^n \sigma_i^2 \leq L^2 := \sum_{i=1}^n L_i^2$. If for some constant $W > 0$ and all $i \leq n$ and all $r \in \mathbb{N}$ holds*

$$\nu_{r,i} := E|X_i|^r \leq \frac{1}{2} L_i^2 W^{r-2} r! \tag{36}$$

then

$$P(S_n \geq t\sigma) \leq \exp\left(-\frac{t^2}{2 + 2\frac{W}{L}t}\right), \qquad \forall t > 0. \tag{37}$$

Remark 3.6 *Actually the Bernstein inequality in Bennet (1962) is formulated with $\sigma_i, 1 \leq i \leq n$, respectively σ instead of upper bounds $L_i, 1 \leq i \leq n$ respectively L. But the proof reveals immediately that the more general version formulated here is also valid.*

Theorem 3.7 *Let (X_n) be a recursive sequence as in (1) with condition E) and let $Y_n = \frac{X_n - EX_n}{s_n}$ be the normalized sequence. Assume that $\sigma_i^2 = \mathrm{Var}(X_i) < \infty$, $i = 0, \ldots, n_0-1$ and assume for some constants W_b, $W > 0$*

a′) $$\nu_{k,i} = E\left(\frac{|X_i - EX_i|}{s_n}\right)^k \leq \frac{1}{2} \max_{j \leq n_0} \sigma_j^2 W^{k-2} k!, \quad i = 0, \ldots, n_0-1, k \geq 1, \tag{38}$$

b′) $E\left|\frac{b_n}{s_n}\right|^k \leq \frac{1}{2} \sigma_b^2 W_b^{r-2} k!, \qquad n \geq n_0, k \geq 1,$

as well as conditions c), d) from Theorem 3.4. Then there exists some $\widetilde{\sigma} \geq \sigma := \max_{j \leq n_0} \sigma_j$ *such that*

$$P(Y_n \geq t) \leq \exp\left(- \frac{t^2}{\widetilde{\sigma}^2 (2 + \frac{\sqrt{2}}{\widetilde{\sigma}} t)} \right), \qquad \forall t > 0. \tag{39}$$

Proof: As in the proof of Theorem 3.4 we obtain in the first step the estimate

$$E|b^{(n)}|^k \leq \frac{1}{2} \bar{\sigma}_b^2 \overline{W}_b^{k-2} k! \tag{40}$$

for some constants $\overline{W}_b$ and $\bar{\sigma}_b$ which can be given in explicit form. Then we establish the induction step and obtain after some calculations similar to those in the proof of Theorem 3.4

$$E|Y_n|^q \leq q! \frac{1}{2} \widetilde{\sigma}^2 (\widetilde{W})^{q-2} \tag{41}$$

with some explicit constants $\widetilde{c}$, $\widetilde{W}$ and in fact $\widetilde{\sigma} = 2\widetilde{W}$. These estimates allow to apply the Bernstein inequality in (37) to the r.v.s (Y_n) to obtain

$$P(Y_n \geq s\widetilde{\sigma}) \leq \exp\left(- \frac{s^2}{2 + 2\frac{\widetilde{W}}{\widetilde{\sigma}} s} \right) \tag{42}$$

which yields the result replacing s by $\frac{t}{s}$ and with $\widetilde{W} = \frac{\widetilde{\sigma}}{\sqrt{2}}$. For details of the argument see Schopp (2005). □

4 Exponential tail bounds; the moment method

In this section we derive exponential tail bounds for recursive sequences using a characterization of general asymptotic exponential tails by the asymptotics of moments. This method uses essentially the following lemma which is based on a theorem of Kasahara (1978).

Define the functions $f(x)$, $g(x)$ on $\mathbb{R}_+$,

$$\begin{aligned} f(x) &\leq_{\text{as}} g(x) && \text{if} \quad \overline{\lim_{x\to\infty}} \frac{f(x)}{g(x)} \leq 1 \quad \text{and} \\ f(x) &\sim g(x) && \text{if} \quad \lim_{x\to\infty} \frac{f(x)}{g(x)} = 1 \end{aligned} \tag{43}$$

Lemma 4.1 *Let X be a random variable, p, $a > 0$ and* $b := \left(\frac{1}{pea}\right)^{1/p}$.

1) For $X \geq 0$ are equivalent:

a)
$$-\ln P(X > x) \geq_{as} ax^p \tag{44}$$

b)
$$(EX^k)^{1/k} \leq_{as} bk^{1/p} \quad \textit{for } k \in 2\mathbb{N} \tag{45}$$

2) *For general X and $p > 1$ a) is further equivalent to:*

c)
$$\ln Ee^{tX} \leq_{as} ct^q \tag{46}$$
where $c = q^{-1}(pa)^{-(q-1)}$ and $\frac{1}{p} + \frac{1}{q} = 1$.

3) *The statements in 1), 2) remain valid also if $\leq_{as}$ is replaced by asymptotic equivalence $\sim$.*

Remark 4.2 *In the paper of Kasahara (1978) the statement of Lemma 4.1 was given for the asymptotic equivalence case (as in part 3)). The method of proof in that paper however also allows to cover the $\leq_{as}$-bounds as in parts 1), 2) of Lemma 4.1. For details see Schopp (2005). The characterization of asymptotic exponential tails in part 3) of Lemma 4.1 was used in Janson and Chassaing (2004) to derive asymptotic exponential tails for the Wiener index of simply generated trees.*

In the following theorem we apply Lemma 4.1 to recurrences as in (1).

Theorem 4.3 *Let (X_n) satisfy the recursive equation (1) and assume that for some nonnegative constants c, b and $p > 0$*

a) $\left(E\left(\frac{X_i}{s_i}\right)^r\right)^{1/r} \leq cr^{1/p}(1 + o(1)), \quad i = 0, \dots, n_0-1$, *as $r \to \infty$,*

b) $(E|b_n|^r)^{1/r} \leq br^{1/p}(1 + o(1))$, *uniformly for $n \geq n_0$, as $r \to \infty$,*

c) $\frac{1}{s_n} + \sum_{r=1}^{M} \frac{s(I_r^{(n)})}{s_n}|A_r(n)| \leq 1, \quad \forall n \geq n_0$.

Then
$$\left(E\left(\frac{|X_n|}{s_n}\right)^r\right)^{1/r} \leq \max\{b, c\}r^{1/p}(1 + o(1)), \text{ uniformly for } n \in \mathbb{N}_0 \text{ as } r \to \infty \tag{47}$$

and
$$P\left(\frac{|X_n|}{s_n} > x\right) \leq e^{-ax^p(1+o(1))} \text{ as } x \to \infty, \text{ uniformly for } n \in \mathbb{N}_0, \tag{48}$$

where $a := \frac{1}{pe\max\{b,c\}^p}$.

Proof: The proof is by induction in n. To take care of the uniformity, we choose the following notation for $i < n_0$ and $n \geq n_0$

$$\left(E\left(\frac{|X_i|}{s_i}\right)^r\right)^{1/r} \leq cr^{1/p}(1 + a_i(r)), \quad (E|b_n|^r)^{1/r} \leq br^{1/p}(1 + a_b(r))$$

where $a_i(r) \to 0, i = 0, \dots, n_0-1,\ a_b(r) \to 0$ as $r \to \infty$. Let $a^*(r) := \max\{a_b(r),\ a_i(r),\ i = 0, \dots, n_0-1\}$. By condition a)

$$\left(E\left(\frac{|X_n|}{s_n}\right)^r\right)^{1/r} \leq \max\{b, c\}r^{1/p}(1 + a^*(r))$$

holds for $0 \le i \le n_0 - 1$. For the induction step we denote by Υ_n the distribution of $(A(n), I^{(n)}, b_n)$ and get

$$
\begin{aligned}
E|Y_n|^q &\le E\Big(\sum_{r=1}^M |A_r(n)| \frac{s(I_r^{(n)})}{s_n} \big|Y^{(r)}_{I_r^{(n)}}\big| + \frac{|b_n|}{s_n}\Big)^q \\
&= \int E\Big(\sum_{r=1}^M |a_r| \frac{s_{i_r}}{s_n} \big|Y^{(r)}_{i_r}\big| + \frac{|\beta|}{s_n}\Big)^q d\Upsilon_n(a,i,\beta) \\
&\le \int \sum_{\ell=0}^q \binom{q}{\ell} \Big(\sum_{r=1}^M \frac{s_{i_r}}{s_n} \big|a_r Y^{(r)}_{i_r}\big|\Big)^\ell \Big(\frac{|\beta|}{s_n}\Big)^{q-\ell} d\Upsilon_n(a,i,\beta) \\
(*) \quad &\le \sum_{\ell=0}^q \binom{q}{\ell} \Big(1 - \frac{1}{s_n}\Big)^\ell \Big(c\ell^{1/p}(1+a^*(q))\Big)^\ell \Big(\frac{1}{s_n}\Big)^{q-\ell} \Big(b(q-\ell)^{\frac{1}{p}}(1+a_b(q))\Big)^{q-\ell} \\
&\le \Big(\max\{b,c\} q^{1/p}(1+a^*(q))\Big)^q E\Big(\frac{1}{s_n} + 1 - \frac{1}{s_n}\Big)^q \\
&= \max\{b,c\} q^{1/p}(1+a^*(q))^q.
\end{aligned}
$$

For the proof of $(*)$ we obtain from the induction hypothesis

$$
\begin{aligned}
E\Big(\sum_{r=1}^M \frac{s_{i_r}}{s_n} \big|a_r Y^{(r)}_{i_r}\big|\Big)^\ell &= \sum_{j_1+\dots+j_M=\ell} \binom{\ell}{j_1 \cdots j_M} \prod_{r=1}^M \Big(\frac{s_{i_r}}{s_n}|a_r|\Big)^{j_r} E\big|Y^{(r)}_{i_r}\big|^{j_r} \\
&\le \sum_{j_1+\dots+j_M=\ell} \binom{\ell}{j_1 \cdots j_M} \prod_{r=1}^M \Big(\frac{s_{i_r}}{s_n}|a_r|\Big)^{j_r} \Big(E\big|Y^{(r)}_{i_r}\big|^l\Big)^{j_r/l} \\
&\le \sum_{j_1+\dots+j_M=\ell} \binom{\ell}{j_1 \cdots j_M} \prod_{r=1}^M \Big(\frac{s_{i_r}}{s_n}|a_r|\Big)^{j_r} \big(cl^{1/p}(1+a^*(l))\big)^{j_r} \\
&\qquad \text{as } i_r < n \\
&\le \big(c\ell^{1/p}(1+a^*(l))\big)^\ell \sum_{j_1+\dots+j_M=\ell} \binom{\ell}{j_1 \cdots j_M} \prod_{r=1}^M \Big(\frac{s_{i_r}}{s_n}|a_r|\Big)^{j_r} \\
&= \big(c\ell^{1/p}(1+a^*(l))\big)^\ell \Big(\sum_{r=1}^M \frac{s_{i_r}}{s_n}|a_r|\Big)^\ell.
\end{aligned}
$$

This implies (47) and as consequence from Lemma 4.1 we obtain the exponential tail bounds in (48). The uniformity is clear, since $a^*(q)$ is independent of n_0. □

Acknowledgements

The authors are grateful to two reviewers for their careful reading of the paper, for many stylistic and organisational suggestions, and for pointing out several inaccuracies in a previous version of this paper.

References

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

Bennet, G. (1962). Probability inequalities for the sum of independent random variables. *American Statistical Association Journal 57*, 33–45.

Boucheron, S., G. Lugosi, and P. Massart (2000). A sharp concentration inequality with applications. *Random Struct. Algorithms 16*(3), 277–292.

Boucheron, S., G. Lugosi, and P. Massart (2003). Concentration inequalities using the entropy method. *Ann. Probab. 31*(3), 1583–1614.

Broutin, N. and L. Devroye (2005). Large deviations for the weighted height of an extended class of trees. `http://jeff.cs.mcgill.ca/~luc/devs.html`.

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

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

Fill, J. A. and S. Janson (2002). Quicksort asymptotics. *J. Algorithms 44*(1), 4–28.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Ann. Stat. Assoc. 58*, 13–30.

Hwang, H.-K. and R. Neininger (2002). Phase change of limit laws in the Quicksort recurrence under varying toll functions. *SIAM J. Comput. 31*(6), 1687–1722.

Janson, S. and P. Chassaing (2004). The center of mass of the ISE and the Wiener index of trees. *Electron. Comm. Probab. 9*, 178–187.

Karp, R. M. (1994). Probabilistic recurrence relations. *J. Assoc. Comput. Mach. 41*(6), 1136–1150.

Kasahara, Y. (1978). Tauberian theorems of exponential type. *J. Math. Kyoto. Univ. 18*, 209–219.

Lugosi, G. (2005). Concentration of measure inequalities. `http://www.econ.upf.es/~lugosi/anu.ps`.

Mahmoud, H. M. (1992). *Evolution of Random Search Trees*. Wiley-Interscience Series in Discrete Mathematics and Optimization. New York: John Wiley & Sons Inc. A Wiley-Interscience Publication.

McDiarmid, C. (1998). Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics*, Volume 16 of *Algorithms Combin.*, pp. 195–248. Berlin: Springer.

Motwani, R. and P. Raghavan (1995). *Randomized Algorithms*. Cambridge: Cambridge University Press.

Neininger, R. (2002). On binary search tree recursions with monomials as toll functions. *J. Comput. Appl. Math. 142*(1), 185–196.

Neininger, R. (2005). Recursive random variables with subgaussian distributions. *Statistics & Decisions 23*, 131–146.

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

Neininger, R. and L. Rüschendorf (2004b). On the contraction method with degenerate limit equation. *Ann. Probab. 32*(3B), 2838–2856.

Rachev, S. T. and L. Rüschendorf (1995). Probability metrics and recursive algorithms. *Adv. in Appl. Probab. 27*(3), 770–799.

Rösler, U. (1991). A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl. 25*(1), 85–100.

Rösler, U. (1992). A fixed point theorem for distributions. *Stochastic Process. Appl. 42*(2), 195–214.

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

Schopp, E.-M. (2005). *Stochastische Fixpunktgleichungen, exponentielle tail Abschätzungen und large deviation für rekursive Algorithmen*. Diplomarbeit, Universität Freiburg.

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

Szpankowski, W. (2001). *Average Case Analysis of Algorithms on Sequences*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York.