UPPER BOUNDS FOR THE MAXIMUM OF A RANDOM WALK WITH NEGATIVE DRIFT

JOHANNES KUGLER AND VITALI WACHTEL

ABSTRACT. Consider a random walk $S_n = \sum_{i=0}^{n} X_i$ with negative drift. This paper deals with upper bounds for the maximum $M = \max_{n \geq 1} S_n$ of this random walk in different settings of power moment existences. As it is usual for deriving upper bounds, we truncate summands. Therefore we use an approach of splitting the time axis by stopping times into intervals of random but finite length and then choose a level of truncation on each interval. Hereby we can reduce the problem of finding upper bounds for M to the problem of finding upper bounds for $M_\tau = \max_{n \leq \tau} S_n$. In addition we test our inequalities in the heavy traffic regime in the case of regularly varying tails.

1. INTRODUCTION AND STATEMENT OF RESULTS

Let $\{S_n, n \geq 0\}$ denote the random walk with increments X_i, that is,

$$S_0 := 0, \ S_n := \sum_{i=1}^{n} X_i, \ n \geq 1.$$

We shall assume that $X_1, X_2, \ldots$ are independent copies of a random variable X with distribution function F and $a := -\mathbf{E}[X] > 0$. The random walk S_n drifts to $-\infty$ and the total maximum $M := \max_{k \geq 0} S_k$ is finite almost surely. The random variable M plays a crucial role in a number of applications. For example, its distribution coincides with the stationary distribution of the queue-length in simple queueing systems. Another important application comes from the insurance mathematics: Under some special restrictions on X the quantity $\mathbf{P}(M > u)$ is equal to the ruin probability in the so-called Poisson model.

The tail-behaviour of M has been studied extensively in the literature. The first result goes back, apparently, to Cramer and Lundberg (see, for example, Asmussen [2]): If

$$\mathbf{E}[e^{h_0 X}] = 1 \quad \text{for some } h_0 > 0, \tag{1}$$

and, in addition, $\mathbf{E}[Xe^{h_0 X}] < \infty$, then there exists a constant $c_0 \in (0, 1)$ such that

$$\mathbf{P}(M > x) \sim c_0 e^{-h_0 x} \quad \text{as } x \to \infty. \tag{2}$$

The case $\mathbf{E}[Xe^{h_0 X}] = \infty$ has been considered recently by Korshunov [10].

If (1) is not fulfilled, then one assumes that the distribution of X is regular in some sense. To specify what regular means we recall some definitions and known properties. For their proofs we refer to Asmussen [2]. Consider a distribution

function B on $\mathbb{R}$ and let $\overline{B}(x) = 1 - B(x)$ be the right tail of B. A distribution function B with support $\mathbb{R}_+$ is called subexponential, if $\overline{B}(x) > 0$ for all x and

$$\lim_{x\to\infty} \frac{\overline{B^{*n}}(x)}{\overline{B}(x)} = n \tag{3}$$

for all $n \geq 2$, where $\overline{B^{*n}}(x)$ is the n-fold convolution of B with itself. For the subexponentiality it is sufficient to verify the equation (3) in the case $n = 2$. All subexponential distributions are heavy-tailed, i.e. $\mathbf{E}[\exp(\epsilon X)] = \infty$ for all $\epsilon > 0$, hence subexponential distributions do not satisfy (1). If (1) is not fulfilled, the most classical result for the asymptotics of M is due to Veraverbeke [16], who showed that if the integrated tail $G(x) := \int_x^\infty \overline{F}(u)du$ is subexponential, then

$$\mathbf{P}(M > x) \sim \frac{1}{a}G(x) \quad \text{as } x \to \infty. \tag{4}$$

In many situations one needs non-asymptotic properties of the distribution of M. Since the exact form of that distribution is known in some special cases only, good estimates are required. Under condition (1) one has for all $x > 0$ the inequality

$$\mathbf{P}(M > x) \leq e^{-h_0 x}. \tag{5}$$

In the case when (1) is not fulfilled, upper bounds for $\mathbf{P}(M > x)$ have been derived by Kalashnikov [9] and by Richards [15]. The approach in these papers is based on the representation of M as a geometric sum of independent random variables:

$$\mathbf{P}(M > x) = \sum_{k=0}^{\infty} q(1-q)^k \mathbf{P}(\chi_1^+ + \chi_2^+ \ldots + \chi_k^+ > x), \tag{6}$$

where $\{\chi_k^+\}$ are independent random variables and $q = \mathbf{P}(M = 0)$. The main difficulty in this approach is the fact that one has to know the distribution of χ_k^+ and the parameter q. In some special cases this information can be obtained from the initial data. But in general one has to obtain appropriate estimates for q and $\mathbf{P}(\chi_1^+ > x)$.

The main purpose of the present paper is to derive upper bounds for $\mathbf{P}(M > x)$ assuming the existence of power moments of X only. Thereby we want to avoid the representation via geometric sums and use a supermartingale-construction instead.

As it is usual for deriving upper bounds, we are going to truncate summands and to use inequalities, which are based on truncated exponential moments. But the problem is that we have infinitely many X_i's. So we can not truncate all of them at the same level. Thus, we have to split the time axis into intervals of finite length and then choose a level of truncation on each of these intervals. One can take, for example, a deterministic strictly increasing sequence k_n with $k_0 = 0$ and consider the intervals $I_n := (k_n, k_{n+1}]$:

$$\mathbf{P}(M \geq x) = \mathbf{P}\left(\bigcup_{k\geq 0}\{S_k \geq x\}\right) \leq \sum_{n=0}^{\infty}\mathbf{P}\left(\bigcup_{k\in I_n}\{S_k \geq x\}\right)$$

$$\leq \sum_{n=0}^{\infty}\mathbf{P}\left(\max_{k\leq k_{n+1}}(S_k - ka) \geq x - k_n a\}\right). \tag{7}$$

Now, one can apply the Fuk-Nagaev inequalities, see [13], to every probability in the last line. It is clear that replacing $\sup_{k\in I_n}(S_k - ka)$ by $\sup_{k\leq k_{n+1}}(S_k - ka)$ is

not too rough if and only if k_{n+1} and $k_{n+1} - k_n$ are comparable. Thus, one has to take k_n exponentially growing. Using this approach with $k_n = x2^n$, Borovkov [3] obtained a version of the Markov inequality for M.

Our strategy is quite different and consists in splitting $[0, \infty)$ into random intervals defined by a sequence of stopping times. More precisely, we introduce the stopping time

$$\tau_z := \min\{k \geq 0 : S_k \leq -z\}, \quad z \leq x.$$

Let $M_\tau = \max_{1 \leq k \leq \tau_z} S_k$. We split the tail probability

$$\mathbf{P}\left(M > x\right) \leq \mathbf{P}\left(M_{\tau_z} > x\right) + \mathbf{P}\left(\max_{k \geq \tau_z} S_k > x\right) \tag{8}$$

and can consider the continuation of the process (S_k) beyond τ_z as a probabilistic replica of the entire process. By $S_{\tau_z} \leq -z$ a.s. follows

$$\mathbf{P}\left(\max_{k \geq \tau_z} S_k > x\right) \leq \mathbf{P}\left(M > x + z\right).$$

As a result, we have

$$\mathbf{P}\left(M > x\right) \leq \mathbf{P}\left(M_\tau > x\right) + \mathbf{P}\left(M > x + z\right),$$

and inductively we conclude

$$\mathbf{P}\left(M > x\right) \leq \sum_{j=0}^{\infty} \mathbf{P}\left(M_\tau > x + jz\right). \tag{9}$$

It is worth mentioning that the difference between (7) and (9) is the same as between Riemann and Lebesgue integrals: We do not fit the random walk S_n into a fixed splitting of the time, but choose the splitting depending on the paths of the random walk.

A decomposition similar to (8) has been used by Denisov [5] for deriving the asymptotics of $\mathbf{P}(M_{\tau_0} > x)$ from that of $\mathbf{P}(M \in [x, x - S_{\tau_0}))$. In the present paper we use the opposite approach: We obtain estimates for $\mathbf{P}(M > x)$ from the ones for $\mathbf{P}(M_{\tau_z} > x)$.

We now state our results on M_τ.

Theorem 1. *Assume that $A_t := \mathbf{E}[|X|^t] < \infty$ for some $t \in (1, 2]$. For all y satisfying $y^{t-1} \geq (e-1)A_t a^{-1}$ we have the following inequality:*

$$\mathbf{P}\left(M_\tau > x\right) \leq \frac{A_t^{x/y}}{a^{x/y-1}} \mathbf{E}[\tau_z] y^{-1-(t-1)x/y} \log\left(1 + ay^{t-1}/A_t\right)$$

$$+ \left(1 + \frac{A_t^{x/y}}{a^{x/y}} y^{-(t-1)x/y}\right) \mathbf{E}[\tau_z] \mathbf{P}(X > y). \tag{10}$$

Remark 2. We show in the proof that (10) remains true, if one replaces a and A_t by $-\mathbf{E}[X, |X| \leq y]$ and $A_t(y) = \mathbf{E}[|X|^t, |X| \leq y]$ respectively. In this case the restriction $y^{t-1} > (e-1)a^{-1}A_t$ should be replaced by $\mathbf{E}[X, |X| \leq y] < 0$. The use of truncated moments is more convenient in theoretical applications, but for deriving concrete estimates for M it is easier to use full moments. $\diamond$

Fix $\alpha \in (0, 1)$ and put $\beta = 1 - \alpha$.

Theorem 3. *Assume $\mathbf{Var}(X) < \infty$ and $A_{t,+} := \mathbf{E}[X^t, X > 0] < \infty$ for some $t > 2$.*

(i) *If y satisfies the condition*

$$\frac{2\alpha a}{e^t \mathbf{Var}(X)} \le \frac{1}{y} \log\left(1 + \frac{\beta a}{A_{t,+}} y^{t-1}\right),$$

then

$$\mathbf{P}\left(M_\tau > x\right) \le \frac{2\alpha a^2}{e^t \mathbf{Var}(X)} \mathbf{E}[\tau_z] \left(\exp\left\{\frac{2\alpha a x}{e^t \mathbf{Var}(X)}\right\} - 1\right)^{-1}$$

$$+ \left(1 + \left(\exp\left\{\frac{2\alpha a x}{e^t \mathbf{Var}(X)}\right\} - 1\right)^{-1}\right) \mathbf{E}[\tau_z]\mathbf{P}(X > y). \qquad (11)$$

(ii) *If y satisfies the condition*

$$\frac{2\alpha a}{e^t \mathbf{Var}(X)} \ge \frac{1}{y} \log\left(1 + \frac{\beta a}{A_{t,+}} y^{t-1}\right), \qquad (12)$$

then

$$\mathbf{P}\left(M_\tau > x\right) \le \frac{A_{t,+}^{x/y} \beta^{-x/y}}{a^{x/y-1}} \mathbf{E}[\tau_z] y^{-1-(t-1)x/y} \log\left(1 + \beta a y^{t-1}/A_{t,+}\right)$$

$$+ \left(1 + \frac{A_{t,+}^{x/y} \beta^{-x/y}}{a^{x/y}} y^{-(t-1)x/y}\right) \mathbf{E}[\tau_z]\mathbf{P}(X > y). \qquad (13)$$

Remark 4. Analogously to Theorem 1 one can replace $\mathbf{Var}(X)$ and $A_{t,+}$ by the corresponding truncated expectations $B^2(-\infty, x) = \mathbf{E}[X^2, X \le y]$ and $A_{t,+}(y) = \mathbf{E}[X^t, X \in (0, y]]$ respectively in Theorem 3. $\diamond$

Corollary 5. *Assume that* $\mathbf{P}(|X| > x) = L(x)x^{-r}$ *for some* $r > 1$ *and*

$$\mathbf{P}(X > x))/\mathbf{P}(|X| > x) \to p \in (0, 1) \quad \text{as } x \to \infty.$$

Then, it follows from (10) and (13) that

$$\limsup_{x \to \infty} \frac{\mathbf{P}\left(M_\tau > x\right)}{\mathbf{P}(X > x)} \le \mathbf{E}[\tau_z]$$

for every $z > 0$.

But it follows from the results of Asmussen [1] (see also Denisov [5]), that

$$\lim_{x \to \infty} \frac{\mathbf{P}\left(M_\tau > x\right)}{\mathbf{P}(X > x)} = \mathbf{E}[\tau_z]$$

under the condition that the tail of F is regularly varying. This means that the inequalities (10) and (13) are asymptotically precise in the case of regularly varying tails.

In all these inequalities we have $\mathbf{E}[\tau_z]$ on the right hand side. It is really hard to get an exact expression for this value via initial data, but there are good upper bounds in the literature: Since $\mathbf{E}[\tau_z] < \infty$ (see, for example, Feller [6]) by Wald's Identity,

$$\mathbf{E}[\tau_z] = \frac{z + \mathbf{E}[R_z]}{a}, \qquad (14)$$

where $R_z = -z - S_{\tau_z}$ denotes the overshoot in τ_z. Hence, we get upper bounds for $\mathbf{E}[\tau_z]$ by the inequality of Lorden [11]: For $\mathbf{E}[X] < 0$ and $\mathbf{E}[(X^-)^2] < \infty$,

$$\mathbf{E}[R_z] \le \frac{\mathbf{E}[(X^-)^2]}{a} \qquad (15)$$

and the one from Mogul'skii [12]: For $\mathbf{E}[X] \leq 0$ and $\mathbf{E}[|X|^3] < \infty$,

$$\mathbf{E}[R_z] \leq A \frac{3}{2} \frac{\mathbf{E}[|X|^3]}{\mathbf{E}[X^2]}, \tag{16}$$

where A is a certain constant, $A \leq 2$. The disadvantage of these bounds is, that we have to assume the existence of the second or even the third moment. We give another bound, which only requires the finiteness of the moment of order t, $t \in (1, 2]$.

Proposition 6. *Assume that* $A_{t,-} := \mathbf{E}[(X^-)^t] < \infty$ *for some* $t \in (1, 2]$, *then, for every* $z > 0$,

$$\mathbf{E}[R_z] \leq \frac{t^{t/(t-1)} A_{t,-}^{1/(t-1)}}{(t-1)a^{t/(t-1)}} \left(\mathbf{E}[-X, X < 0] + \frac{z^{2-t}}{t} A_{t,-} \right). \tag{17}$$

Combining (14) with (15), (16) or (17) we obtain upper bounds for $\mathbf{E}[\tau_z]$. Plugging these bounds into the inequalities in Theorems 1 and 3 we get bounds for $\mathbf{P}(M_\tau > x)$, which contain information on X only. So, they can be used for concrete calculations.

We now come back to the global maximum.

Theorem 7. *Fix some* $\theta \in (0, 1)$ *and define*

$$c_1 := \frac{3A_t^{1/\theta}\theta^{-(t-1)/\theta}}{(t-1)a^{1/\theta-1}}, \quad c_2 := \frac{3A_{t,+}^{1/\theta}\theta^{-(t-1)/\theta}}{(t-1)a^{1/\theta-1}}.$$

(i) *Assume that* $A_t < \infty$ *for some* $t \in (1, 2]$. *Then, for every* x *satisfying* $x^{t-1} \geq \theta^{1-t}(e-1)A_t a^{-1}$ *and* $x \geq z(t-1)\theta^{-1}$, *we have*

$$\mathbf{P}(M > x) \leq c_1 \frac{\mathbf{E}[\tau_z]}{z} \log\left(1 + \frac{a\theta^{t-1}x^{t-1}}{A_t}\right) x^{-(t-1)/\theta}$$

$$+ \left(1 + \left(\frac{A_t}{\theta^{t-1}ax^{t-1}}\right)^{1/\theta}\right) \mathbf{E}[\tau_z] \left(\frac{1}{\theta z}G(\theta x) + \mathbf{P}(X > \theta x)\right). \tag{18}$$

(ii) *Assume that* $\mathbf{Var}(X) < \infty$ *and* $A_{t,+} < \infty$ *for some* $t > 2$. *Then, for every* x *satisfying* (12) *for* $y = \theta x$, $x^{t-1} \geq \theta^{1-t}(e^\theta - 1)A_{t,+}\beta^{-1}a^{-1}$ *and* $x \geq z(t-1)\theta^{-1}$, *we have*

$$\mathbf{P}(M > x) \leq c_2\beta^{-1/\theta}\frac{\mathbf{E}[\tau_z]}{z} \log\left(1 + \frac{\beta a\theta^{t-1}x^{t-1}}{A_{t,+}}\right) x^{-(t-1)/\theta}$$

$$+ \left(1 + \left(\frac{A_{t,+}}{\beta\theta^{t-1}ax^{t-1}}\right)^{1/\theta}\right) \mathbf{E}[\tau_z] \left(\frac{1}{\theta z}G(\theta x) + \mathbf{P}(X > \theta x)\right). \tag{19}$$

Corollary 8. *If the assumptions of Corollary 5 hold, then it follows from Theorem 7 that*

$$\limsup_{x \to \infty} \frac{\mathbf{P}(M > x)}{G(x)} \leq \frac{\mathbf{E}[\tau_z]}{z}\theta^{-r}.$$

Since the left-hand side does not depend on θ and z, we can let $\theta \to 1$ and $z \to \infty$. Noting that each of (15) and (16) combined with (14) yields

$$\frac{\mathbf{E}[\tau_z]}{z} \to \frac{1}{a} \quad \text{as } z \to \infty,$$

we conclude

$$\limsup_{x\to\infty} \frac{\mathbf{P}(M > x)}{G(x)} \leq \frac{1}{a}.$$

Comparing this with (4), we see that the inequalities in Theorem 7 are asymptotically precise. This even remains valid, if we bound $\mathbf{E}[\tau_z]$ in these inequalities by combining (15) or (16) with (14).

The reason why we are able to obtain asymptotically precise bounds is, that we may choose z arbitrary large. That possibility seems to be a quite important advantage of our method compared to geometric sums. If the distribution of χ_1^+ is subexponential, then it follows easily from (6) that

$$\mathbf{P}(M > x) \sim \left(\frac{1}{q} - 1\right) \mathbf{P}(\chi_1^+ > x) \quad \text{as } x \to \infty.$$

Therefore, in order to obtain an upper bound for the maximum we need to control the quantity $1/q$. It is well known that $1/q = \mathbf{E}[-S_{\tau_0}] = \mathbf{E}[R_0]$. Thus, we may apply (15), (16) or (17) with $z = 0$. But corresponding inequalities for M will not be asymptotically precise. Summarising, the approach via geometric sums can only lead to asymptotically precise bounds if q is known.

We next test our inequalities in the heavy-traffic regime. Let $\{S^{(a)}, a \geq 0\}$ be a family of random walks with $\mathbf{E}[X^{(a)}] = -a$. We shall assume that $X^{(a)} = X^{(0)} - a$ for all $a > 0$. Let $M^{(a)}$ denote the corresponding maximum. It is known that if $X^{(0)}$ belongs to the domain of attraction of a stable law, then there exists a regularly varying function $g(a)$ such that $g(a)M^{(a)}$ converges weakly as $a \to 0$. It turns out that our inequalities may be applied to large deviation problems in the heavy-traffic convergence mentioned above. More precisely, they give asymptotically precise bounds for the probabilities $\mathbf{P}(M^{(a)} > x_a)$ if $x_a \gg 1/g(a)$. In the case of $\sigma^2 := \mathbf{Var}(X^{(0)})$ being finite, one has $g(a) = a$ and the weak limit of $aM^{(a)}$ is the exponential distribution with parameter $2/\sigma^2$.

Theorem 9. *Assume that $\sigma^2 < \infty$ and the right tail of the distribution function of $X^{(0)}$ is regularly varying with index $r > 2$, that is, $\mathbf{P}(X^{(0)} > u) = u^{-r}L(u)$, where L is slowly varying. If*

$$\liminf_{a\to 0} \frac{x_a}{a^{-1} \log a^{-1}} > e^r \frac{(r-2)}{2} \sigma^2, \tag{20}$$

then

$$\mathbf{P}(M^{(a)} > x_a) \sim \frac{x_a^{-r+1} L(x_a)}{(r-1)a} \quad \text{as } a \to 0. \tag{21}$$

Olvera-Cravioto, Blanchet and Glynn [14] have shown that for an $M/G/1$ queue the relation (21) holds under the condition

$$\liminf_{a\to 0} \frac{x_a}{a^{-1} \log a^{-1}} > \frac{(r-2)}{2} \sigma^2.$$

We believe that the latter should be sufficient for the validity of (21) also in the general case. The extra factor e^r appears in (20) only as a consequence of the technique we used, and can be removed by adoption of (19) to the heavy-traffic setting.

Theorem 10. *Assume that $\mathbf{E}[(\min\{0, X^{(0)}\})^2] < \infty$ and $\mathbf{P}(X^{(0)} > u) = u^{-r}L(u)$ with $r \in (1, 2)$. If*

$$\liminf_{a \to 0} g(a)x_a = \infty, \tag{22}$$

then

$$\mathbf{P}(M^{(a)} > x_a) \sim \frac{x_a^{-r+1}L(x_a)}{(r-1)a} \quad as \ a \to 0. \tag{23}$$

We have imposed the condition $\mathbf{E}[(\min\{0, X^{(0)}\})^2] < \infty$ just to use the Lorden inequality for the overshoot. If one replaces that condition by $\mathbf{E}[(\min\{0, X^{(0)}\})^t] < \infty$ with $t \in (1, 2)$, then, using Proposition 6, one can show that (23) holds for $x_a \gg a^{-t/(t-1)^2}$ only. The reason is the roughness of Proposition 6 for small values of a. Indeed, if we use (17) even with $t = 2$, then we get the bound $\mathbf{E}[R_z] \leq Ca^{-2}$, which is much worse then the Lorden inequality.

2. Proofs

2.1. **Proofs of Theorems 1 and 3.** We set for brevity $\tau = \tau_z$.

Lemma 11. *For all h satisfying*

$$\mathbf{E}[e^{hX}, X \leq y] \leq 1 \tag{24}$$

we have the inequality

$$\mathbf{P}(M_\tau > x) \leq \left(1 + \frac{1}{e^{hx} - 1}\right)\mathbf{E}[\tau]\mathbf{P}(X > y) + \mathbf{E}[\tau]\frac{ah}{e^{hx} - 1}. \tag{25}$$

Proof. Our strategy is to truncate the random variables X_i in the level y:

$$\mathbf{P}\left(M_\tau > x\right) \leq \mathbf{P}\left(M_\tau > x, \max_{1 \leq k \leq \tau} X_k \leq y\right) + \mathbf{P}\left(\max_{1 \leq k \leq \tau} X_k > y\right)$$

$$= \mathbf{P}\left(M_\tau \mathbf{1}_{\{\max_{1 \leq k \leq \tau} X_k \leq y\}} > x\right) + \mathbf{P}\left(\max_{1 \leq k \leq \tau} X_k > y\right). \tag{26}$$

From the Wald identity follows

$$\mathbf{P}\left(\max_{1 \leq k \leq \tau} X_k > y\right) \leq \mathbf{E}\left[\sum_{k=1}^{\tau} \mathbf{1}_{\{X_k > y\}}\right] = \mathbf{E}[\tau]\mathbf{P}(X > y). \tag{27}$$

To examine the first term on the right-hand-side of (26) we introduce the process $\{W_k\}$ defined by

$$W_0 := 1, \quad W_k := \prod_{i=1}^{k} e^{hX_i}\mathbf{1}_{\{X_i \leq y\}}, \ k \geq 1.$$

It is clear that if h satisfies (24), $\{W_k\}$ is a positive supermartingale. Define

$$\sigma_y := \min\{k \geq 1 : X_k > y\}, \ t_x := \min\{k \geq 1 : S_k > x\} \text{ and } T := \min\{\sigma_y, t_x, \tau\}.$$

Applying the Optional Stopping Theorem to the supermartingale $\{W_{k \wedge T}\}$ leads us

$$1 = W_0 \geq \mathbf{E}[W_T] = \mathbf{E}\left[W_T\mathbf{1}_{\{t_x < \tau, t_x < \sigma_y\}}\right] + \mathbf{E}\left[W_T\mathbf{1}_{\{\tau < t_x, \tau < \sigma_y\}}\right].$$

We analyse the two terms on the right-hand-side separately:

$$\mathbf{E}\left[W_T\mathbf{1}_{\{t_x < \tau, t_x < \sigma_y\}}\right] \geq e^{hx}\mathbf{P}(t_x < \tau < \sigma_y) = e^{hx}\mathbf{P}\left(M_\tau \mathbf{1}_{\{\max_{1 \leq k \leq \tau} X_k \leq y\}} > x\right)$$

and

$$\mathbf{E}\left[W_T \mathbf{1}_{\{\tau < t_x, \tau < \sigma_y\}}\right] = \mathbf{E}\left[e^{hS_\tau}\right] - \mathbf{E}\left[e^{hS_\tau} \mathbf{1}_{\{M_\tau > x\} \cup \{\max_{1 \le k \le \tau} X_k > y\}}\right]$$

$$\ge \mathbf{E}\left[e^{hS_\tau}\right] - e^{-hz}\left(\mathbf{P}\left(M_\tau \mathbf{1}_{\{\max_{1 \le k \le \tau} X_k \le y\}} > x\right) + \mathbf{P}\left(\max_{1 \le k \le \tau} X_k > y\right)\right).$$

Consequently,

$$\mathbf{P}\left(M_\tau \mathbf{1}_{\{\max_{1 \le k \le \tau} X_k \le y\}} > x\right) \le \frac{1 - \mathbf{E}\left[e^{hS_\tau}\right] + \mathbf{P}\left(\max_{1 \le k \le \tau} X_k > y\right)}{e^{hx} - 1}$$

and hence by applying (27),

$$\mathbf{P}\left(M_\tau \mathbf{1}_{\{\max_{1 \le k \le \tau} X_k \le y\}} > x\right) \le \frac{1 - \mathbf{E}\left[e^{hS_\tau}\right] + \mathbf{E}[\tau]\mathbf{P}\left(X > y\right)}{e^{hx} - 1}.$$

It is easy to see that

$$\mathbf{E}\left[e^{hS_\tau}\right] \ge \mathbf{E}\left[1 + hS_\tau\right] = 1 + h\mathbf{E}[S_\tau]$$

and as a result we have

$$\mathbf{P}\left(M_\tau \mathbf{1}_{\{\max_{1 \le k \le \tau} X_k \le y\}} > x\right) \le \mathbf{E}[\tau]\frac{ha + \mathbf{P}\left(X > y\right)}{e^{hx} - 1}. \tag{28}$$

Applying (27) and (28) to the summands in (26) finishes the proof. $\qquad \square$

To prove Theorems 1 and 3 we need to choose a specific h for which (24) holds. The optimal choice would be the positive solution of the equation $\mathbf{E}[e^{hX}, X \le y] = 1$, which is in the spirit of the Cramer-Lundberg condition. But it is not clear how to solve this equation. For this reason we replace $\mathbf{E}[e^{hX}, X \le y] = 1$ by the equation $\phi(h, y) = 1$, where $\phi(h, y)$ is an appropriate upper bound for $\mathbf{E}[e^{hX}, X \le y]$.

If $A_t < \infty$, we may use a bound from the proof of Theorem 2 from [8], which says

$$\mathbf{E}[e^{hX}, X \le y] \le 1 + h\mathbf{E}[X, |X| \le y] + \frac{e^{hy} - 1 - hy}{y^t}A_t. \tag{29}$$

Using the Markov inequality we also obtain

$$\mathbf{E}[X, |X| \le y] \le -a - \mathbf{E}[X, X \le -y] \le -a + \frac{A_t}{y^{t-1}},$$

and therefore

$$\mathbf{E}[e^{hX}, X \le y] \le 1 - ha + \frac{e^{hy} - 1}{y^t}A_t.$$

Put $h_0 := \frac{1}{y}\log\left(1 + ay^{t-1}/A_t\right)$. It is easy to see that

$$-h_0 a + \frac{e^{h_0 y} - 1}{y^t}A_t \le 0$$

for all y such that $y^{t-1} \ge (e-1)A_t/a$ and this implies that h_0 satisfies (24). Using (25) with $h = h_0$ and applying the inequality

$$(1 + u)^{x/y} \ge 1 + u^{x/y}, \; x \ge y,$$

we obtain

$$\mathbf{P}(M_\tau > x) \le \frac{A_t^{x/y}}{a^{x/y-1}} \mathbf{E}[\tau] y^{-1-(t-1)x/y} \log\left(1 + ay^{t-1}/A_t\right)$$

$$+ \left(1 + \frac{A_t^{x/y}}{a^{x/y}} y^{-(t-1)x/y}\right) \mathbf{E}[\tau]\mathbf{P}(X > y).$$

Thus, the proof of Theorem 1 is complete.

In order to show that one can replace $\mathbf{E}[X]$ and A_t by the corresponding truncated moments, see Remark 2, we first note that analogously to (29) and by additionally using $e^x - 1 \le xe^x$,

$$\mathbf{E}[e^{hX}, X \le y] \le 1 + h\mathbf{E}[X, |X| \le y] + h\frac{e^{hy} - 1}{y^{t-1}}\mathbf{E}[|X|^t, |X| \le y].$$

If $\mathbf{E}[X, |X| \le y] < 0$, then

$$h_0 := \frac{1}{y} \log\left(1 + \frac{|\mathbf{E}[X, |X| \le y]|y^{t-1}}{\mathbf{E}[|X|^t, |X| \le y]}\right)$$

is strictly positive and solves

$$h\mathbf{E}[X, |X| \le y] + h\frac{e^{hy} - 1}{y^{t-1}}\mathbf{E}[|X|^t, |X| \le y] = 0.$$

Therefore, we may use Lemma 11 with $h = h_0$ and get an inequality with truncated moments.

To bound $\mathbf{E}[e^{hX}, X \le y]$ under the conditions of Theorem 3 we proceed similar to the proof of Theorem 3 from [13] and get

$$\mathbf{E}[e^{hX}, X \le y] \le 1 - ha + e^t\mathbf{Var}(X)\frac{h^2}{2} + \frac{e^{hy} - 1 - hy}{y^t}A_{t,+}.$$

Following further the method from the proof of this Theorem, we split this upper bound into two parts:

$$-\alpha ha + e^t\mathbf{Var}(X)\frac{h^2}{2} =: f_1(h),$$

$$-\beta ha + \frac{e^{hy} - 1 - hy}{y^t}A_{t,+} =: f_2(h).$$

We consider f_1 and f_2 separately. It is clear that

$$h_1 := \frac{2\alpha a}{e^t\mathbf{Var}(X)}$$

is the positive solution of the equation $f_1(h) = 0$. Moreover, $f_1(h) < 0$ for all $h \in (0, h_1)$.

Furthermore, it is easy to see that f_2 takes it's unique minimum in

$$h_2 := \frac{1}{y} \log\left(1 + a\frac{\beta y^{t-1}}{A_{t,+}}\right).$$

Since f_2 is convex, one has

$$f_2(h) < 0 \quad \text{for all } h \in (0, h_2].\tag{30}$$

The assumption in Theorem 3(i) means that $h_1 \le h_2$. In this case, taking into account (30), we obtain

$$f_1(h_1) + f_2(h_1) < 0.$$

From the latter inequality we conclude that h_1 satisfies (24) and by applying (25) with $h = h_1$ we obtain (11).

Under the conditions of Theorem 3 (ii) we have $h_2 \le h_1$. By the same arguments we get

$$f_1(h_2) + f_2(h_2) < 0.$$

Then, applying (25) with $h = h_2$ and using the inequality $(1 + u)^{x/y} \ge u^{x/y}$, we obtain (13).

2.2. **Proof of Proposition 6.** We want to use Theorem 2.1 from [4]. If we put $F := F_{-X}$ the conditions (G1)-(G3) of this Theorem are fulfilled in our setting. Hence we get

$$\mathbf{E}[R_z] \le c \int_0^\infty \mathbf{P}(-X > u)du + c \int_0^\infty \int_u^{u+z} \mathbf{P}(-X > v)dvdu, \qquad (31)$$

where

$$c = \frac{b^*(\epsilon a)}{a(1 - \epsilon)} \qquad (32)$$

with $b^*(u) = \min\{v : -\mathbf{E}[X, X < -v] \le u\}$ and $\epsilon \in (0, 1)$ arbitrary. By the Theorem of Fubini we obtain

$$\int_0^\infty \mathbf{P}(-X > u)du = \mathbf{E}[-X, X < 0]. \qquad (33)$$

Changing the order of integration gives us

$$\int_0^\infty \int_u^{u+z} \mathbf{P}(-X > v)dvdu = \int_0^z v\mathbf{P}(-X > v)dv + z \int_z^\infty \mathbf{P}(-X > v)dv$$

$$\le z^{2-t} \int_0^\infty v^{t-1}\mathbf{P}(-X > v)dv = \frac{z^{2-t}}{t}A_{t,-}. \qquad (34)$$

As you can easily see,

$$b^*(u) \le \left(\frac{A_{t,-}}{u} \right)^{1/(t-1)},$$

therefore by (32)

$$c \le \frac{A_{t,-}^{1/(t-1)}}{a^{t/(t-1)}\epsilon^{1/(1-t)}(1 - \epsilon)},$$

and by minimisation over $\epsilon \in (0, 1)$

$$c \le \frac{t^{t/(t-1)}A_{t,-}^{1/(t-1)}}{(t - 1)a^{t/(t-1)}}. \qquad (35)$$

Finally, combining (31), (33), (34) and (35) gives us the desired result.

2.3. **Proof of Theorem 7.** We prove (18) only. The proof of the second bound goes along the same line.

Using Theorem 1 with $y = \theta(x + jz)$, we obtain

$$\mathbf{P}(M_\tau > x + jz) \le \frac{A_t^{1/\theta}\theta^{-1-(t-1)/\theta}\mathbf{E}[\tau_z]}{a^{1/\theta-1}(x + jz)^{1+(t-1)/\theta}} \log \left(1 + \frac{a\theta^{t-1}(x + jz)^{t-1}}{A_t} \right)$$

$$+ \left(1 + \frac{A_t^{1/\theta}\theta^{-(t-1)/\theta}}{a^{1/\theta}}(x + jz)^{-(t-1)/\theta} \right) \mathbf{E}[\tau_z]\mathbf{P}(X > \theta(x + jz)),$$

and in view of (9),

$$\mathbf{P}(M > x) \leq \frac{A_t^{1/\theta}\theta^{-1-(t-1)/\theta}}{a^{1/\theta-1}}\mathbf{E}[\tau_z]\Sigma_1(x,z)$$
$$+ \left(1 + \frac{A_t^{1/\theta}}{a^{1/\theta}}x^{-(t-1)/\theta}\right)\mathbf{E}[\tau_z]\left(\mathbf{P}(X > \theta x) + \Sigma_2(x,z)\right), \quad (36)$$

where

$$\Sigma_1(x,z) := \sum_{j=0}^{\infty}\log\left(1 + \frac{a\theta^{t-1}(x+jz)^{t-1}}{A_t}\right)(x+jz)^{-1-(t-1)/\theta}$$

and

$$\Sigma_2(x,z) := \sum_{j=1}^{\infty}\mathbf{P}(X > \theta(x+jz)).$$

Define

$$\tilde{\Sigma}_1(x,z) := \sum_{j=1}^{\infty}\log\left(1 + \frac{a\theta^{t-1}(x+jz)^{t-1}}{A_t}\right)(x+jz)^{-1-(t-1)/\theta}.$$

The summands in this sum are strictly decreasing, so we conclude by the integral criteria for sums:

$$\tilde{\Sigma}_1(x,z) \leq \sum_{j=1}^{\infty}\int_{j-1}^{j}\log\left(1 + \frac{a\theta^{t-1}(x+uz)^{t-1}}{A_t}\right)(x+uz)^{-1-(t-1)/\theta}du$$
$$= \frac{1}{z}\int_x^{\infty}\log\left(1 + \frac{a\theta^{t-1}w^{t-1}}{A_t}\right)w^{-1-(t-1)/\theta}dw$$

and further by integration by parts,

$$\frac{1}{z}\int_x^{\infty}\log\left(1 + \frac{a\theta^{t-1}w^{t-1}}{A_t}\right)w^{-1-(t-1)/\theta}dw$$
$$\leq \frac{\theta}{z(t-1)}\log\left(1 + \frac{a\theta^{t-1}x^{t-1}}{A_t}\right)x^{-(t-1)/\theta} + \frac{\theta^2}{z(t-1)}x^{-(t-1)/\theta}.$$

Therefore, for all x sufficing $x^{t-1} \geq \theta^{1-t}(e^{\theta} - 1)A_t a^{-1}$ and $x \geq z(t-1)\theta^{-1}$,

$$\Sigma_1(x,z) \leq \frac{3\theta}{z(t-1)}\log\left(1 + \frac{a\theta^{t-1}x^{t-1}}{A_t}\right)x^{-(t-1)/\theta}.$$

Furthermore, it is easy to see that

$$\Sigma_2(x,z) \leq \sum_{j=1}^{\infty}\int_{j-1}^{j}\mathbf{P}(X > \theta(x+uz))du = \frac{1}{\theta z}G(\theta x). \quad (37)$$

and Theorem 7 is proved.

2.4. **Proof of Theorem 9.** Foss, Korshunov and Zachary have shown, see Theorem 5.1 in [7], that for any random walk with the drift $-a$ and x_a with $x_a \to \infty$ as $a \to 0$ one has the following lower bound:

$$\liminf_{a\to 0}\frac{\mathbf{P}(M^{(a)} > x_a)}{a^{-1}G(x_a)} \geq 1.$$

It follows from the regular variation of $\mathbf{P}(X^{(0)} > u)$, that

$$G(x_a) \sim \frac{1}{r-1} x_a^{-r+1} L(x_a) \quad \text{as } a \to \infty, \tag{38}$$

therefore

$$\mathbf{P}(M^{(a)} \geq x_a) \geq (1 + o(1)) \frac{x_a^{-r+1} L(x_a)}{(r-1)a} \quad \text{as } a \to 0.$$

Thus, we only have to derive an upper bound. During this proof we assume a to be sufficiently small in every inequality. We want to apply Theorem 7 (ii) with $t < r$. It is clear that

$$A_{t,+}^{(a)} := \mathbf{E}[(X^{(a)})^t, X^{(a)} > 0] \leq \mathbf{E}[(X^{(0)})^t, X^{(0)} > 0],$$

therefore $A_{t,+}^{(a)}$ is finite for $t < r$ and

$$\lim_{a \to 0} A_{t,+}^{(a)} = A_{t,+}^{(0)} > 0.$$

Further, we have to show that (12) is fulfilled for $y = \theta x$ under our assumptions. Since the function $y^{-1} \log(1 + \beta a y^{t-1}/A_{t,+}^{(a)})$ is decreasing for $y \gg a^{1/(t-1)}$, we have the following bound for $x_a \geq ca^{-1} \log a^{-1}$:

$$\frac{1}{\theta x_a} \log \left(1 + \frac{\beta \theta^{t-1} a x_a^{t-1}}{A_{t,+}^{(a)}} \right) \leq \frac{a}{\theta c \log a^{-1}} \log \left(1 + \frac{\beta \theta^{t-1} c^{t-1}}{A_{t,+}^{(a)}} a^{2-t} \log^{t-1} a^{-1} \right)$$

$$= \frac{(t-2)}{\theta c} a(1 + o(1)).$$

This implies that if $c > e^r (r-2) \sigma^2 / 2$ and $\theta = (t-2)/(r-2)$, we can choose $\alpha < 1$ so close to 1 that x_a satisfies (12).

We take $z = z_a$ satisfying $a^{-1} \ll z \ll x_a$. Then, combining (14) and (15), we get

$$\frac{\mathbf{E}[\tau_z]}{z} \sim \frac{1}{a} \quad \text{as } a \to 0. \tag{39}$$

Since $(t-1)/\theta = (t-1)(r-2)/(t-2) > r - 1$,

$$\beta^{-1/\theta} \frac{\mathbf{E}[\tau_z]}{z} \log \left(1 + \frac{\beta \theta^{t-1} a x_a^{t-1}}{A_{t,+}^{(a)}} \right) x_a^{-(t-1)/\theta} = o \left(a^{-1} x_a^{-r+1} L(x_a) \right). \tag{40}$$

Further, it follows from the condition $z = o(x)$ and the regular variation of $\mathbf{P}(X^{(0)} > x)$ that

$$z \mathbf{P}(X^{(a)} > x_a) = o \left(x_a^{-r+1} L(x_a) \right). \tag{41}$$

Combining (39) with (41) and (38), we obtain

$$\left(1 + \left(\frac{A_{t,+}^{(a)}}{\beta \theta^{t-1} a x_a^{t-1}} \right)^{1/\theta} \right) \mathbf{E}[\tau_z] \left(\frac{1}{\theta z} G(\theta x_a) + \mathbf{P}(X^{(a)} > \theta x_a) \right)$$

$$\sim \theta^{-r} (r-1)^{-1} a^{-1} x_a^{-r+1} L(x_a) \tag{42}$$

and plugging (40) and (42) into (19) gives us

$$\limsup_{a \to 0} \frac{\mathbf{P}(M^{(a)} > x_a)}{a^{-1} x_a^{-r+1} L(x_a)} \leq \theta^{-r} (r-1)^{-1}.$$

To complete the proof it suffices to note, that we can choose θ arbitrary close to 1 by choosing t close to r. This implies that the previous inequality is valid even with $\theta = 1$.

2.5. **Proof of Theorem 10.** We again need an upper bound only. Let a be sufficiently small during this proof.

It follows from the assumptions in the theorem that $S_n^{(0)}/c_n$ converges weakly to a stable law of index r. The sequence c_n can be taken from the equation $c_n^{-r}L(c_n) = 1/n$. It is known that the function $g(a)$ in the heavy-traffic approximation can be defined by the relations

$$g(a) = 1/c_{n_a} \text{ and } an_a \sim c_{n_a}.$$

The latter can be rewritten as

$$c_{n_a} \sim a\frac{(c_{n_a})^r}{L(c_{n_a})}.$$

From this we infer that (22) is equivalent to

$$\frac{ax_a^{r-1}}{L(x_a)} \to \infty \quad \text{as } a \to 0. \tag{43}$$

We want to apply Theorem 1 with $-\mathbf{E}[X^{(a)}, |X^{(a)}| \le y]$ and $A_2(y)$ instead of a and A_2 respectively with $y = \theta x$. According to Remark 2 we have to show that $\mathbf{E}[X^{(a)}, |X^{(a)}| \le \theta x_a]$ is negative. Using the Markov inequality, we have

$$\mathbf{E}[X^{(a)}, |X^{(a)}| \le \theta x_a] \le -a + (\theta x_a)^{-1}\mathbf{E}[(\min\{0, X^{(0)}\})^2].$$

In view of (43), $ax_a \to \infty$. Therefore,

$$\mathbf{E}[X^{(a)}, |X^{(a)}| \le \theta x_a] \le -a(1 + o(1)).$$

Furthermore,

$$A_2(y) \sim \frac{r}{2-r}y^{2-r}L(y) \tag{44}$$

and consequently by $-\mathbf{E}[X^{(a)}, |X^{(a)}| \le \theta x_a] \sim a$,

$$A_2^{1/\theta}(\theta x_a)\mathbf{E}[\tau_z]\frac{(-\mathbf{E}[X^{(a)}, |X^{(a)}| \le \theta x_a])^{1-1/\theta}}{\theta^{1+1/\theta}x_a^{1+1/\theta}}\log\left(1 - \frac{\theta x_a\mathbf{E}[X^{(a)}, |X^{(a)}| \le \theta x_a]}{A_2(\theta x_a)}\right)$$

$$\le (1 + o(1))k_1\mathbf{E}[\tau_z]\mathbf{P}(X^{(a)} > x_a)\log\left(1 + k_2\frac{ax_a^{r-1}}{L(x_a)}\right)\left(\frac{ax_a^{r-1}}{L(x_a)}\right)^{-(1/\theta-1)} \tag{45}$$

with k_1 and k_2 appropriate. Then, (43) implies that

$$\log\left(1 + c_2\frac{ax_a^{r-1}}{L(x_a)}\right)\left(\frac{ax_a^{r-1}}{L(x_a)}\right)^{-(1/\theta-1)} = o(1). \tag{46}$$

Further,

$$\frac{A_2^{1/\theta}(\theta x_a)}{a^{1/\theta}}\theta^{-1/\theta}x_a^{-1/\theta} \sim k_3\left(\frac{ax_a^{r-1}}{L(x_a)}\right)^{-1/\theta}$$

with k_3 suitable and hence by (43),

$$\left(1 + \frac{A_2^{1/\theta}(\theta x_a)}{a^{1/\theta}}\theta^{-1/\theta}x_a^{-1/\theta}\right) = (1 + o(1)). \tag{47}$$

Then, combining (45), (46) and (47), Theorem 1 with $t = 2$ and $y = \theta x_a$ gives us

$$\mathbf{P}(M_\tau^{(a)} > x_a) \leq (1 + o(1))\theta^{-r}\mathbf{E}[\tau_z]\mathbf{P}(X^{(a)} > x_a),$$

where $\theta \in (0, 1)$ is arbitrary. Hence by $\theta \to 1$,

$$\mathbf{P}(M_\tau^{(a)} > x_a) \leq (1 + o(1))\mathbf{E}[\tau_z]\mathbf{P}(X^{(a)} > x_a).$$

By the summation formula (9) we get a bound for the total maximum:

$$\mathbf{P}(M^{(a)} > x_a) \leq (1 + o(1))\mathbf{E}[\tau_z]\sum_{j=0}^{\infty}\mathbf{P}(X^{(a)} > x_a + jz). \qquad (48)$$

Combining (37) and (38) with $a^{-1} \ll z \ll x_a$ gives us

$$\sum_{j=0}^{\infty}\mathbf{P}(X^{(a)} > x_a + jz) \leq (1 + o(1))\left(x_a^{-r}L(x_a) + \frac{x_a^{-r+1}L(x_a)}{z(r-1)}\right)$$

$$\sim (1 + o(1))\frac{x_a^{-r+1}L(x_a)}{z(r-1)}$$

and regarding (39) completes the proof.

REFERENCES

[1] Asmussen, S. Subexponential asymptotics for stochastic processes: extremal behaviour, stationary distributions and first passage probabilities. *Ann. Appl. Probab.*, 8:354–374, 1998.

[2] Asmussen, S. Ruin Probabilities. *World Scientific Publishing Co. Pte. Ltd.*, , 2000.

[3] Borovkov, A.A. Stochastic Processes in Queueing Theory. *Springer,* New York-Berlin, 1976.

[4] Borovkov, A.A. and Foss, S.G. Estimates for the excess of a random walk over an arbitrary boundary and their applications. *Theory Probab. Appl.,* 44:231–253, 1998.

[5] Denisov, D. A note on the asymptotics for the maximum on a random time interval of a random walk. *Markov Process. Related Fields*, 11:165–169 (2005).

[6] Feller, W. An introduction to probability theory and its applications, volume II, 2nd ed. *John Wiley & Sons, Inc.,* NewYork-London-Sydney, 1971.

[7] Foss, S., Korshunov, D. and Zachary, S. An Introduction to Heavy-Tailed and Subexponential Distributions. *Springer Science+Business Media*, 2011.

[8] Fuk, D.Kh. and Nagaev, S.V. Probability inequalities for sums of independent random variables. *Theory Probab. Appl.,* 4:643–660, 1971.

[9] Kalashnikov, V. Bounds for ruin probabilities in the presence of large claims and their comparison. *North Amer. Act. J.,* 3:116–129, 1999.

[10] Korshunov, D. The critical case of the Cramér-Lundberg theorem on the asymptotics of the distribution of the maximum of a random walk with negative drift. *Siberian Math. J.,* 46:1077–1081, 2005.

[11] Lorden, G. On excess over the boundary. *Ann. Math. Statist.,* 41:520–527, 1970.

[12] Mogul'skii, A.A. Absolute estimates for moments of certain boundary functionals. *Theory Probab. Appl.,* 18:340–347, 1973.

[13] Nagaev, S.V. Large deviations of sums of independent random variables. *Ann. Probab.,* 5:745–789, 1979.

[14] Olvera-Cravioto, M., Blanchet, J. and Glynn, P. On the transition from heavy traffic to heavy tails for the $M/G/1$ queue: the regularly varying case. *Ann. Appl. Prob.,* 21:645–668, 2011.

[15] Richards, A. On upper bounds for the tail distribution of geometric sums of subexponential random variables. *Queueing Syst.*, 62:229–242, 2009.

[16] Veraverbeke, N. Asymptotic behaviour of Wiener-Hopf factors of a random walk. *Stochastic Process. Appl.*, 5:27–37, 1977.

Mathematical Institute, University of Munich, Theresienstrasse 39, D-80333, Munich, Germany
E-mail address: kugler@math.lmu.de

Mathematical Institute, University of Munich, Theresienstrasse 39, D-80333, Munich, Germany
E-mail address: wachtel@math.lmu.de