Electron. Commun. Probab. **19** (2014), no. 41, 1–14.
DOI: 10.1214/ECP.v19-3266
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Large deviations for weighted sums of stretched exponential random variables*

Nina Gantert† Kavita Ramanan‡ Franz Rembart§

Abstract

We consider the probability that a weighted sum of n i.i.d. random variables $X_j, j = 1, \ldots, n$, with stretched exponential tails is larger than its expectation and determine the rate of its decay, under suitable conditions on the weights. We show that the decay is subexponential, and identify the rate function in terms of the tails of X_j and the weights. Our result generalizes the large deviation principle given by Kiesel and Stadtmüller [9] as well as the tail asymptotics for sums of i.i.d. random variables provided by Nagaev [10, 11]. As an application of our result, motivated by random projections of high-dimensional vectors, we consider the case of random, self-normalized weights that are independent of the sequence $\{X_j\}_{j\in\mathbb{N}}$, identify the decay rate for both the quenched and annealed large deviations in this case, and show that they coincide. As another application we consider weights derived from kernel functions that arise in nonparametric regression.

Keywords: Large deviations; weighted sums; subexponential random variables; stretched exponential random variables; self-normalized weights; quenched and annealed large deviations; kernels; nonparametric regression.
AMS MSC 2010: 60F10; 62G32.
Submitted to ECP on January 18, 2014, final version accepted on June 16, 2014.

1 Introduction

Let $\{X_j\}_{j\in\mathbb{N}}$ be a sequence of independent and identically distributed (i.i.d.) random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that take values in the real line $\mathbb{R}$ and have finite expectation $m := \mathbb{E}[X_1] < \infty$. For $n \in \mathbb{N}$, let $S_n := \sum_{j=1}^n X_j$ denote the partial sum, and let $\bar{S}_n := S_n/n$ denote the empirical mean. The strong law of large numbers implies that $\bar{S}_n \to m$ almost surely. Cramér's Theorem on large deviations tells us that if the X_j have finite exponential moments, that is, there exists $t > 0$ such that

$$M(t) := \mathbb{E}[\exp(tX_1)] < \infty, \tag{1.1}$$

then for any $x > m$, the probability $\mathbb{P}\left(\bar{S}_n \geq x\right)$ decays exponentially. More precisely,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\bar{S}_n \geq x\right) = -\Lambda^*(x),$$

*This research was supported in part by NSF CMMI-1234100 and AFOSR FA9550-12-1-0399
†Technische Universität München, Germany. E-mail: gantert@ma.tum.de
‡Brown University, USA. E-mail: Kavita_Ramanan@brown.edu
§University of Oxford, UK. E-mail: franz.rembart@stats.ox.ac.uk

where $\Lambda^*(x) := \sup_{t\geq 0}\{tx - \log M(t)\} > 0$. We will refer to this case as the "light-tailed" case. It is well known that if $M(t) = +\infty$ for all $t > 0$, the probabilities $\mathbb{P}\left(\bar{S}_n \geq x\right)$ decay slower than exponentially. The reason is that, in contrast to when (1.1) holds, a "deviation" of the type $\bar{S}_n \geq x$ is produced by the event that *just one* of the random variables takes a large value. For instance, if there is $r \in (0,1)$ and $c > 0$ such that $\mathbb{P}(X_1 \geq t) = c\exp(-t^r)$ for t large enough, then

$$\lim_{n\to\infty} \frac{1}{n^r} \log \mathbb{P}\left(\bar{S}_n \geq x\right) = -(x-m)^r, \quad \forall x > m. \tag{1.2}$$

The result in (1.2) goes back to Theorem 3 in [10] and it will also follow from our main result, Theorem 1. Cramér's Theorem was generalized in [9] to weighted sums of i.i.d. random variables; see Section 2 below for a precise statement of their results. Our main result, Theorem 1, gives a corresponding statement for weighted sums of i.i.d. random variables with stretched exponential tails, which arise in many applications. One motivation to consider weighted sums, which is elaborated upon in Section 5.1, comes from random projections of high-dimensional vectors, which are of relevance in asymptotic geometric analysis [5] and data analysis [2]. Another motivation stems from statistics (kernel functions, moving averages), see Section 5.2 for an example. The analogous example for the light-tailed case was considered in [9].

This article is organized as follows: We first present the result and the regularity conditions from [9] in Section 2. Our main result, Theorem 1, is given in Section 3, and its proof is presented in Section 4. Finally, in Section 5.1, we give an application to random weights, and in Section 5.2, we consider weights derived from kernel functions that arise in non-parametric regression.

2 The Light-Tailed Case

For $n \in \mathbb{N}$, let $\{a_j(n)\}_{j\in\mathbb{N}}$ be a sequence of real numbers which we will call weights. For $n \in \mathbb{N}$, define the weighted sum

$$\bar{S}_n := \sum_{j=1}^{n} a_j(n) X_j, \tag{2.1}$$

and let μ_n be the distribution of $\bar{S}_n$, that is, the measure on $\mathcal{B}(\mathbb{R})$, the set of Borel sets in $\mathbb{R}$, given by

$$\mu_n(A) := \mathbb{P}\left(\bar{S}_n \in A\right), \quad A \in \mathcal{B}(\mathbb{R}). \tag{2.2}$$

When the $\{X_j\}_{j\in\mathbb{N}}$ have finite exponential moments, that is, the moment generating function $M(t)$ defined in (1.1) is finite for all $t \in \mathbb{R}$, a large deviation principle for the sequence of weighted sums $\{\bar{S}_n\}_{n\in\mathbb{N}}$ was established in [9] under suitable assumptions on the weights, see Assumption A below. The "classical" case of Cramér's theorem corresponds to $a_j(n) = 1/n, j = 1, 2, \ldots, n$, $n \in \mathbb{N}$.

Assumption A. *(A.1) There exists a sequence of real numbers $\{s_\nu\}_{\nu\in\mathbb{N}}$ such that $s_\nu \neq 0$ for all $\nu \in \mathbb{N}$, the limit $s := \lim_{\nu\to\infty} \sqrt[\nu]{|s_\nu|}$ exists and*

$$\sum_{j=1}^{n} (a_j(n))^\nu = \frac{s_\nu}{n^{\nu-1}} R(\nu, n) \text{ for all } \nu \text{ and } n \in \mathbb{N}, \tag{2.3}$$

for some function $R : \mathbb{N}^2 \to \mathbb{R}$ that satisfies, for every $\nu \in \mathbb{N}$, $R(\nu, n) \to 1$ as $n \to \infty$.

(A.2) There exist sequences $\{r_\nu\}_{\nu\in\mathbb{N}}$ *and* $\{\delta_n\}_{n\in\mathbb{N}}$ *such that* $\limsup_{\nu\to\infty} \sqrt[\nu]{r_\nu} \le 1$, $\lim_{n\to\infty}\delta_n = 0$ *and the error term satisfies*

$$|R(\nu,n)-1| \le r_\nu \frac{(1+\delta_n)^\nu}{n} \text{ for all } \nu \text{ and } n. \tag{2.4}$$

Now, let Λ denote the cumulant (or log moment) generating function of X_1, and let $\{c_\nu\}_{\nu\in\mathbb{N}}$ be the sequence of coefficients that arise in the power series expansion for Λ: that is, given $M(t)$ as in (1.1),

$$\Lambda(t) := \log M(t) = \sum_{\nu=1}^{\infty} \frac{c_\nu}{\nu!} t^\nu, \quad t\in\mathbb{R}. \tag{2.5}$$

Also, for $t>0$, let $\chi(t) := \sum\limits_{\nu=1}^{\infty} \frac{s_\nu c_\nu}{\nu!} t^\nu$, and let χ^* denote its Legendre-Fenchel transform:

$$\chi^*(t) := \sup_{t\in\mathbb{R}}\{tx - \chi(t)\}. \tag{2.6}$$

It was shown in [9] that under Assumption A the sequence of measures $\{\mu_n\}_{n\in\mathbb{N}}$ on $\mathcal{B}(\mathbb{R})$ defined in (2.2) satisfies a large deviation principle with speed n and rate function χ^*. Recall that this means that

$$-\inf_{x\in A^\circ}\chi^*(x) \le \liminf_{n\to\infty}\frac{1}{n}\mu_n(A^\circ) \le \limsup_{n\to\infty}\frac{1}{n}\mu_n(\bar{A}) \le -\inf_{x\in\bar{A}}\chi^*(x), \quad \forall A\in\mathcal{B}(\mathbb{R}),$$

where A° and $\bar{A}$, respectively, represent the interior and the closure of the set A.

Remark 2.1. In fact, [9] provides a more general result that considers an infinite sum and refers to a general scale within the regularity conditions (cf. Assumption A), that is, they prove large deviations for the family of weighted sums of the form $A(\lambda) := \sum_{j=1}^{\infty} a_j(\lambda)X_j$, where $\lambda\in I$ and either $I=\mathbb{N}$ or $I=[0,\infty]$.

Our goal will be to relax the finiteness assumption (1.1) on the moment generating function $M(\cdot)$.

3 Main Result

In order to present our large deviation result for weighted sums of stretched exponential random variables, we will use slightly different assumptions on the weights from those used in [9]. We will restrict our considerations to non-negative weights. As we show in Lemma 3.1 below, in this case, our assumptions are weaker than those used in [9].

Assumption B. *(B.1) There exists a real number* $s_1 \neq 0$ *such that the sequence* $\{R(1,n)\}_{n\in\mathbb{N}}$ *of real numbers defined by*

$$\sum_{j=1}^{n} a_j(n) = s_1 R(1,n), \text{ for all } n\in\mathbb{N},$$

satisfies $R(1,n)\to 1$ *as* $n\to\infty$.

(B.2) There exists a real number s *such that for* $a_{max}(n) := \max_{1\le j\le n} a_j(n)$,

$$\lim_{n\to\infty} n\cdot a_{max}(n) = s. \tag{3.1}$$

Examples of weight sequences that satisfy both Assumption A and Assumption B include Valiron means, see [9], as well as kernel functions (see Section 5.1).

Recall that a function $\ell : (0,\infty) \to (0,\infty)$ is called **slowly varying** (at infinity) if for every $a > 0$,

$$\lim_{x\to\infty} \frac{\ell(ax)}{\ell(x)} = 1. \tag{3.2}$$

We now state our main result.

Theorem 1 (Large Deviations for Weighted Sums, Stretched Exponential Tails)**.** *Let* $\{X_j\}_{j\in\mathbb{N}}$ *be a sequence of i.i.d. random variables on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with*

$$\mathbb{E}[|X_1|^k] < \infty \quad \forall k \in \mathbb{N}, \tag{3.3}$$

and let $m := \mathbb{E}[X_1]$*. Suppose that there exist a constant* $r \in (0,1)$ *and slowly varying functions* b*,* c_1*,* $c_2 : (0,\infty) \to (0,\infty)$ *and a constant* $t^* > 0$ *such that for* $t \geq t^*$*,*

$$c_1(t) \exp\left(-b(t)t^r\right) \leq \mathbb{P}\left(X_1 \geq t\right) \leq c_2(t) \exp\left(-b(t)t^r\right). \tag{3.4}$$

Let $\{a_j(n)\}_{j\in\mathbb{N}}$*,* $n \in \mathbb{N}$*, be an infinite array of non-negative real numbers that satisfy Assumption B with associated constants* $s_1, s \in \mathbb{R}$*, and let* $\{\bar{S}_n\}_{n\in\mathbb{N}}$ *be the sequence of weighted sums defined in* (2.1)*. Then*

$$\lim_{n\to\infty} \frac{1}{b(n)n^r} \log \mathbb{P}\left(\bar{S}_n \geq x\right) = -\left(\frac{x}{s} - \frac{s_1}{s} m\right)^r, \quad \forall x > s_1 m. \tag{3.5}$$

Remark 3.1. The non-negativity assumption on the weights cannot be relaxed without additional information about the lower tail of the $\{X_j\}$, that is, about the probabilites $\mathbb{P}(X_1 \leq -t)$ for $t > 0$. Consider the following example: $a_j(n) = 1/n, j = 1, \ldots, \lfloor 2n/3 \rfloor$, $a_j(n) = -1/n, j = \lfloor 2n/3 \rfloor + 1, \ldots, n$ (where, for $z \in \mathbb{R}$, $\lfloor z \rfloor$ represents the greatest integer less than or equal to z). Then Assumption B is satisfied with $s_1 = 1/3$ and $s = 1$. Take i.i.d. random variables $\{X_j\}_{j\in\mathbb{N}}$ with mean m that satisfy (3.3), (3.4) and the lower tail bound $\mathbb{P}(X_1 \leq -t) = \exp(-t^\alpha)$ for some α with $0 < \alpha < r$, and t large enough. Then, applying Theorem 1 to $\{-X_j\}_{j\in\mathbb{N}}$ with $a_j(n) = 1/n$, and for any $\varepsilon > 0$, noting on the one hand that, as $n \to \infty$, $\mathbb{P}(X_1 + \ldots + X_{\lfloor 2n/3 \rfloor} \geq 2n(m+\varepsilon)/3)$ is negligible in comparison with $\mathbb{P}(-X_{\lfloor 2n/3 \rfloor+1} - \ldots - X_n \geq n(x - 2(m+\varepsilon)/3))$, and on the other hand that $\mathbb{P}(X_1 + \ldots X_{\lfloor 2n/3 \rfloor} \geq 2n(m-\varepsilon)/3)$ converges to 1 by the strong law of large numbers, it can be shown that for every $x > m/3$, we have with $\gamma_x = (x - m/3)^\alpha > 0$ that

$$\lim_{n\to\infty} \frac{1}{n^\alpha} \log \mathbb{P}\left(\bar{S}_n \geq x\right) = -\gamma_x < 0.$$

However, we cannot recover α, and hence, γ_x from the assumptions in Theorem 1.

Remark 3.2. For the same reason as in the last remark, namely that the only assumption on the lower tail of $\{X_j\}_{j\in\mathbb{N}}$ is (3.3), the result in (3.4) cannot be strengthened to a large deviation principle without imposing further assumptions. For $x < s_1 m$, the decay of $\mathbb{P}(\bar{S}_n \leq x)$ is determined by the lower tail of the $\{X_j\}$. For example, if the $\{X_j\}_{j\in\mathbb{N}}$ are bounded below, Cramér's Theorem implies that $\mathbb{P}(\bar{S}_n \leq x)$ decays exponentially in n. If, on the other hand, $\mathbb{P}(X_1 \leq -t) = \exp(-t^\alpha)$ with $0 < \alpha < r$, then as in Remark 3.1, we can show that $-\infty < \lim_{n\to\infty} n^{-\alpha} \log \mathbb{P}(\bar{S}_n \leq x) < 0$.

Stretched exponential distributions have been proposed as a complement to the frequently used power law distributions to model many naturally occurring heavy-tailed distributions, see e.g. [6] for applications. Any distribution that satisfies (3.4) and is bounded below also satisfies (3.3). A concrete example is the Weibull distribution with shape parameter lying in the interval $(0,1)$. Before proceeding to the proof of Theorem 1, we comment on the relationship between Assumptions A and B. Specifically, for a non-negative sequence of weights, we show in Lemma 3.1 that Assumption B is weaker than Assumption A. To see that it is strictly weaker, consider the sequence of weights defined by $a_j(n) = n^{-1} + n^{-(1+\varepsilon)}$, $j = 1, ..., n$, for some $\varepsilon \in (0, \frac{1}{2})$. It is easy to show that this sequence satisfies Assumption B, but does not satisfy condition (A.2).

Lemma 3.1 (Relationship between Assumptions A and B)**.** *Let* $\{a_j(n)\}_{j\in\mathbb{N}}$*,* $n \in \mathbb{N}$*, be an infinite array of non-negative real numbers that satisfy Assumption A. Then* $\{a_j(n)\}_{j\in\mathbb{N}}$*,* $n \in \mathbb{N}$ *also satisfies Assumption B.*

Proof. Given weights $\{a_j(n)\}_{j\in\mathbb{N}}$, $n \in \mathbb{N}$, that satisfy Assumption A, clearly (B.1) follows immediately from (A.1). It only remains to show (B.2). First, note that by Assumption (A.2), $R(\nu, n)$ satisfies the inequality

$$1 - r_\nu \frac{(1+\delta_n)^\nu}{n} \le R(\nu, n) \le 1 + r_\nu \frac{(1+\delta_n)^\nu}{n}. \tag{3.6}$$

Moreover, for any $\varepsilon > 0$, we can find $\nu^*(\varepsilon) \in \mathbb{N}$ and $n^*(\varepsilon) \in \mathbb{N}$ such that

$$0 \le r_\nu \le (1+\varepsilon)^\nu, \quad \forall \nu \ge \nu^*(\varepsilon), \quad \text{and} \quad 0 \le \delta_n \le \varepsilon, \quad \forall n \ge n^*(\varepsilon). \tag{3.7}$$

By using the inequality $(a_{max}(n))^\nu \le \sum_{j=1}^n (a_j(n))^\nu$, (A.1) and (A.2), we see that for $\nu, n \in \mathbb{N}$,

$$na_{max}(n) \le n \left(\sum_{j=1}^n (a_j(n))^\nu \right)^{\frac{1}{\nu}} = n(s_\nu R(\nu, n))^{\frac{1}{\nu}} \cdot (n^{1-\nu})^{\frac{1}{\nu}} \le n^{\frac{1}{\nu}} (s_\nu)^{\frac{1}{\nu}} \left(1 + r_\nu \frac{(1+\delta_n)^\nu}{n} \right)^{\frac{1}{\nu}}.$$

Together with (3.7), this implies that for $\varepsilon > 0$, and $\nu \ge \nu^*(\varepsilon)$, $n \ge n^*(\varepsilon)$,

$$na_{max}(n) \le (s_\nu)^{\frac{1}{\nu}} \left(n(1+\varepsilon)^{2\nu} + (1+\varepsilon)^{2\nu} \right)^{\frac{1}{\nu}} = (n+1)^{\frac{1}{\nu}} (s_\nu)^{\frac{1}{\nu}} (1+\varepsilon)^2.$$

Setting $\nu = n$, for $n \ge \max\{\nu^*(\varepsilon), n^*(\varepsilon)\}$, we have

$$na_{max}(n) \le \sqrt[n]{n+1}\, \sqrt[n]{s_n} (1+\varepsilon)^2.$$

Since $s = \lim_{n\to\infty} \sqrt[n]{s_n}$ by (A.1), taking first the limit superior as $n \to \infty$ and then as $\varepsilon \downarrow 0$, we see that

$$\limsup_{n\to\infty} na_{max}(n) \le \lim_{\varepsilon\downarrow 0} s(1+\varepsilon)^2 = s. \tag{3.8}$$

Next, to bound $na_{max}(n)$ from below, we will make use of the fact that $(na_{max}(n))^\nu \ge n^{\nu-1} \sum_{j=1}^n (a_j(n))^\nu$. Indeed, then for $\varepsilon > 0$, by (2.3), (2.4) and (3.7), for $\nu \ge \nu^*(\varepsilon)$ and $n \ge n^*(\varepsilon)$, we have

$$na_{max}(n) \ge (s_\nu R(\nu, n))^{\frac{1}{\nu}} \ge (s_\nu)^{\frac{1}{\nu}} \left(1 - r_\nu \frac{(1+\delta_n)^\nu}{n} \right)^{\frac{1}{\nu}} \ge (s_\nu)^{\frac{1}{\nu}} \left(1 - \frac{(1+\varepsilon)^{2\nu}}{n} \right)^{\frac{1}{\nu}}.$$

Taking limits as $n \to \infty$ and noting that $(1 - \frac{(1+\varepsilon)^{2\nu}}{n})^n \sim \exp\{-(1+\varepsilon)^{2\nu}\}$ and $n\nu \to \infty$ as $n \to \infty$, we obtain

$$\liminf_{n\to\infty} na_{max}(n) \ge (s_\nu)^{\frac{1}{\nu}} \liminf_{n\to\infty} \left(\left(1 - \frac{(1+\varepsilon)^{2\nu}}{n} \right)^n \right)^{\frac{1}{n\nu}} \ge (s_\nu)^{\frac{1}{\nu}}, \quad \forall \nu \ge \nu^*(\varepsilon).$$

Sending $\nu \to \infty$ and recalling from (A.1) that $s = \lim_{\nu\to\infty} \sqrt[\nu]{s_\nu}$, we conclude that

$$\liminf_{n\to\infty} na_{max}(n) \geq s. \tag{3.9}$$

Combining (3.8) and (3.9), we see that the weights $\{a_j(n)\}_{j\in\mathbb{N}}$ satisfy (B.2), and thus Assumption B. □

4 Proof of Theorem 1

We will prove a slightly stronger statement than Theorem 1, namely we show in Section 4.2 that if (3.3) holds for only $k = 1, 2$ and the first inequality in (3.4) holds, then the lower bound

$$\liminf_{n\to\infty} \frac{1}{b(n)\, n^r} \log \mathbb{P}\left(\bar{S}_n \geq x\right) \geq -\left(\frac{x}{s} - \frac{s_1}{s} m\right)^r, \quad \forall x > s_1 m, \tag{4.1}$$

holds; and in Section 4.3 we show that (3.3) and the second inequality in (3.4) imply the upper bound

$$\limsup_{n\to\infty} \frac{1}{b(n)\, n^r} \log \mathbb{P}\left(\bar{S}_n \geq x\right) \leq -\left(\frac{x}{s} - \frac{s_1}{s} m\right)^r, \quad \forall x > s_1 m. \tag{4.2}$$

First, in Section 4.1, we summarize some relevant properties of slowly varying functions. Throughout the section, the notation $f(x) \sim g(x)$ as $x \to \infty$ for two functions $f, g : \mathbb{R} \to \mathbb{R}$ means that $\lim_{x\to\infty} f(x)/g(x) = 1$. Also, given a set A, $\mathbb{1}_A$ will denote the indicator function of A, which equals 1 on A and 0 on the complement.

4.1 Properties of Slowly Varying Functions

We will need the following preliminaries on slowly varying functions. Proposition (4.1) corresponds to Proposition 1.3.6 in [1], where Lemma (4.2) refers to (1.4) in [7].

Proposition 4.1 (Properties of Slowly Varying Functions)**.** *Let $\ell : (0,\infty) \to (0,\infty)$ be a slowly varying function (at infinity). Then*

(i) $\lim_{x\to\infty} \dfrac{\log \ell(x)}{\log x} = 0.$
(ii) For any $\alpha \in \mathbb{R}$, the function $f(x) = (\ell(x))^\alpha, x \in \mathbb{R}$, is slowly varying.
(iii) For any $\alpha > 0$, $x^\alpha l(x) \to \infty$ and $x^{-\alpha} l(x) \to 0$ as $x \to \infty$.

Furthermore, if $m : (0,\infty) \to (0,\infty)$ is another slowly varying function then

(iv) the functions $f(x) = \ell(x)m(x)$ and $g(x) = \ell(x) + m(x)$, $x \in \mathbb{R}$, are slowly varying.
(v) if $m(x) \to \infty$ as $x \to \infty$, then the function $f(x) = \ell(m(x)), x \in \mathbb{R}$, is slowly varying.

Lemma 4.2 (Representation for Slowly Varying Functions)**.** *A function $\ell : (0,\infty) \to (0,\infty)$ is slowly varying if and only if there exist $a > 0$, $\bar{\eta} \in \mathbb{R}$ and bounded measurable functions $\eta(\cdot)$ and $\varepsilon(\cdot)$ with $\eta(x) \to \bar{\eta}$, $\varepsilon(x) \to 0$ as $x \to \infty$ such that, for $x \geq a$, ℓ can be written in the form*

$$\ell(x) = \exp\left\{\eta(x) + \int_a^x \frac{\varepsilon(u)}{u} du\right\}. \tag{4.3}$$

As a direct consequence of Lemma 4.2, we have the following result.

Lemma 4.3. *Let $\ell : (0,\infty) \to (0,\infty)$ be a slowly varying function and let $g : (0,\infty) \to (0,\infty)$ be another function such that $g(x) \to c$ for some $c \in (0,\infty)$ as $x \to \infty$. Then we have*

$$\lim_{x\to\infty} \frac{\ell\left(g(x)x\right)}{\ell(x)} = 1. \tag{4.4}$$

4.2 The Lower Bound

For $n \in \mathbb{N}$, let $j^*(n) := \inf\{1 \le j \le n : a_j(n) = a_{max}(n)\}$. For any fixed $\varepsilon > 0$, since the $\{X_j\}_{j\in\mathbb{N}}$ are i.i.d.,

$$\begin{aligned}
\mathbb{P}(\bar{S}_n \ge x) &= \mathbb{P}\left(\sum_{j=1}^n a_j(n)(X_j - m) \ge x - \sum_{j=1}^n a_j(n)m\right) \\
&\ge \mathbb{P}\left(a_{\max}(n)(X_{j^*(n)} - m) \ge x - \sum_{j=1}^n a_j(n)m + \varepsilon, \sum_{j\in\{1,\dots,n\}, j\neq j^*(n)} a_j(n)(X_j - m) \ge -\varepsilon\right) \\
&= \mathbb{P}\left(X_1 \ge t_1(n)\right) \mathbb{P}\left(\sum_{j\in\{1,\dots,n\}, j\neq j^*(n)} a_j(n)(X_j - m) \ge -\varepsilon\right),
\end{aligned}$$

where $t_1(n) = t_1^\varepsilon(n)$ is defined by

$$t_1(n) := \frac{1}{na_{\max}(n)}\left[n\left(x - \sum_{j=1}^n a_j(n)m + a_{\max}(n)m + \varepsilon\right)\right], \quad n \in \mathbb{N}. \tag{4.5}$$

Applying the lower bound of (3.4) with $t = t_1(n)$, we obtain

$$\mathbb{P}\left(\bar{S}_n \ge x\right) \ge c_1\left(t_1(n)\right) \exp\left\{-b\left(t_1(n)\right)\left(t_1(n)\right)^r\right\} \cdot \mathbb{P}\left(\sum_{j\in\{1,\dots,n\}, j\neq j^*(n)} a_j(n)(X_j - m) \ge -\varepsilon\right). \tag{4.6}$$

Note that by Assumption B, $t_1(n) \sim \left(\frac{x}{s} - \frac{s_1}{s}m + \frac{\varepsilon}{s}\right) n$ as $n \to \infty$. Since $c_1(\cdot)$ and $b(\cdot)$ are slowly varying functions, Lemma 4.3 implies that $c_1\left(t_1(n)\right) \sim c_1(n)$ and $b\left(t_1(n)\right) \sim b(n)$ as $n \to \infty$. Moreover, note that for some fixed $\delta \in (0, r)$, we can express $\log c_1(n)/b(n)n^r = (\log c_1(n)/\log n)(\log n/n^\delta)(b(n)n^{r-\delta})^{-1}$, which goes to zero as $n \to \infty$ by properties (i) and (iii) of Proposition 4.1. Furthermore, since the $\{X_j\}$ have finite second moments by (3.3), and (B.2) implies that $\sum_{j=1, j\neq j^*(n)}^n a_j(n)^2 \le n(a_{\max}(n))^2 \to 0$ as $n \to \infty$, it follows that $\sum_{j\in\{1,\dots,n\}, j\neq j^*(n)} a_j(n)(X_j - m)$ converges to zero in $\mathbb{L}^2$. In turn, this implies that $\lim_{n\to\infty} \mathbb{P}(\sum_{j\in\{1,\dots,n\}, j\neq j^*(n)} a_j(n)(X_j - m) \ge -\varepsilon) = 1$. Thus, taking logarithms of both sides of (4.6), then dividing by $b(n)n^r$ and sending first $n \to \infty$, and then $\varepsilon \downarrow 0$, we obtain the lower bound (4.1).

4.3 The Upper Bound

Let $t_2(n) := n\left(\frac{x}{s} - \frac{s_1}{s}m\right)$. Then, we can write

$$\mathbb{P}\left(\bar{S}_n \ge x\right) \le A_1^n + A_2^n, \tag{4.7}$$

where, for $n \in \mathbb{N}$,

$$A_1^n := \mathbb{P}\left(\max_{1\le j\le n} X_j \ge t_2(n)\right), \qquad A_2^n := \mathbb{P}\left(\bar{S}_n \ge x, \max_{1\le j\le n} X_j < t_2(n)\right).$$

The union bound and the upper tail bound for X_1 in (3.4) imply that

$$A_1^n \le n\mathbb{P}(X_1 \ge t_2(n)) \le nc_2\left(t_2(n)\right) \cdot \exp\left\{-b\left(t_2(n)\right)\left(t_2(n)\right)^r\right\}.$$

Since b is slowly varying, $b\left(t_2(n)\right) \sim b\left(n\right)$ as $n \to \infty$, and properties (i) and (iii) of Proposition 4.1 show that $\lim_{n\to\infty} \log n/b(n)n^r = \lim_{n\to\infty} \log c_2(t_2(n))/b(n)n^r = 0$. Together with the last display, this implies that

$$\limsup_{n\to\infty} \frac{1}{b\left(n\right)n^r} \log A_1^n \le \limsup_{n\to\infty} \frac{-(t_2(n))^r}{n^r} = -\left(\frac{x}{s} - \frac{s_1}{s}m\right)^r. \tag{4.8}$$

Next, we turn to A_2^n. Applying the exponential Chebyshev inequality with a positive real parameter $\beta_\zeta(n)/s$ (to be specified later), and using the i.i.d. property of the sequence $\{X_j\}_{j\in\mathbb{N}}$, we obtain

$$A_2^n \le \exp\left\{-\beta_\zeta(n)\frac{x}{s}\right\} \cdot \prod_{j=1}^n \mathbb{E}\left[\exp\left\{\beta_\zeta(n)\frac{a_j(n)}{s}X_j\right\} \cdot \mathbb{1}_{\{X_j<t_2(n)\}}\right]. \tag{4.9}$$

Now, for $\zeta > 0$, define

$$\beta_\zeta(n) := \zeta n^r b\left(n\left(\frac{x}{s} - \frac{s_1}{s}m\right)\right) = \zeta n^r b(t_2(n)). \tag{4.10}$$

Then, since $b(\cdot)$ is slowly varying, $\lim_{n\to\infty} \beta_\zeta(n)/(b(n)n^r) = \zeta$. Together with (4.9) this implies that

$$\limsup_{n\to\infty} \frac{1}{b(n)n^r} \log A_2^n \le -\zeta\frac{x}{s} + \limsup_{n\to\infty} \frac{1}{b(n)n^r} \sum_{j=1}^n \Lambda_\zeta^j(n), \tag{4.11}$$

where, for $j = 1,\dots,n$, $n \in \mathbb{N}$, and $\zeta > 0$, we define

$$\Lambda_\zeta^j(n) := \log \mathbb{E}\left[\exp\left\{\beta_\zeta(n)\frac{a_j(n)}{s}X_j^{(n)}\right\}\right], \quad \text{where } X_j^{(n)} := X_j\mathbb{1}_{\{X_j<t_2(n)\}}. \tag{4.12}$$

We now show that the upper bound (4.2) is satisfied if the following proposition holds.

Proposition 4.4 (Boundedness of the remainder). *For every* $\zeta < \left(\frac{x}{s} - \frac{s_1}{s}m\right)^{r-1}$,

$$\limsup_{n\to\infty} \frac{1}{b\,(n)\,n^r} \sum_{j=1}^n \Lambda_\zeta^j(n) \le \zeta m \frac{s_1}{s}. \tag{4.13}$$

Indeed, given Proposition 4.4, we can substitute (4.13) into (4.11) and send $\zeta \uparrow \left(\frac{x}{s} - \frac{s_1}{s}m\right)^{r-1}$ to conclude that

$$\limsup_{n\to\infty} \frac{1}{b(n)n^r} \log A_2^n \le -\left(\frac{x}{s} - \frac{s_1}{s}m\right)^r.$$

Together with (4.7), and the analogous bound (4.8) for A_1^n, we obtain the upper bound (4.2).

Thus, to prove the upper bound, it only remains to prove Proposition 4.4. We use similar techniques as in [8].

Proof of Proposition 4.4. Fix $\zeta < (\frac{x}{s} - \frac{s_1}{s}m)^{r-1}$ and denote $\beta_\zeta(n)$ and Λ_ζ^j simply as $\beta(n)$ and Λ^j. For the fixed $r \in (0,1)$, we also choose $k \in \mathbb{N}$ such that $r < k/(k+1)$. Then, by the definition (4.12) of Λ^j, the estimates $\log x \le x - 1$ for $x > 0$ and $e^x - 1 \le x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + ... + \frac{1}{(k+1)!}x^{k+1}e^x$, finiteness of the moments of X_j due to (3.3), and the fact that $\beta(n)/(b(n)n^r) \to \zeta$ and $\sum_{j=1}^n a_j \to s_1$ as $n \to \infty$, we have

$$\limsup_{n\to\infty} \frac{1}{b\,(n)\,n^r} \sum_{j=1}^n \Lambda^j(n) \quad \le \quad \limsup_{n\to\infty} \frac{1}{b(n)n^r} \left(\sum_{j=1}^n \sum_{i=1}^k \frac{\mathbb{E}\left[\left(\beta(n)\frac{a_j(n)}{s}X_j^{(n)}\right)^i\right]}{i!} \right) + \frac{B_0}{(k+1)!},$$

with

$$B_0 := \limsup_{n\to\infty} \frac{1}{b(n)n^r} \sum_{j=1}^n \left(\beta(n)\frac{a_j(n)}{s}\right)^{k+1} \cdot \mathbb{E}\left[\left(X_j^{(n)}\right)^{k+1} \exp\left(\beta(n)\frac{a_j(n)}{s}X_j^{(n)}\right)\right].$$

Now, note that due to (3.3) and Assumption B, $\lim_{n\to\infty} \frac{1}{b(n)n^r} \sum_{j=1}^n \mathbb{E}[(\beta(n)\frac{a_j(n)}{s}X_j^{(n)})^i]$ is equal to $\zeta m\frac{s_1}{s}$ if $i=1$, and is equal to zero if $i \neq 1$. This implies that

$$\limsup_{n\to\infty} \frac{1}{b\,(n)\,n^r} \sum_{j=1}^n \Lambda^j(n) \le \zeta m \frac{s_1}{s} + \frac{B_0}{(k+1)!}.$$

To complete the proof of Proposition 4.4, it suffices to show that $B_0 = 0$. In this regard, we distinguish between the cases $X_j^{(n)} < t^*$ and $X_j^{(n)} \ge t^*$, where we recall that for $t \ge t^*$, (3.4) is satisfied. Specifically, we bound B_0 by $\limsup_{n\to\infty}(B_1(n)+B_2(n))$, where

$$B_1(n) := \frac{1}{b(n)n^r} \sum_{j=1}^n \left(\beta(n)\frac{a_j(n)}{s}\right)^{k+1} \cdot (t^*)^{k+1} \exp\left(\beta(n)\frac{a_j(n)}{s}t^*\right), \tag{4.14}$$

$$B_2(n) := \frac{1}{b(n)n^r} \sum_{j=1}^n \left(\beta(n)\frac{a_j(n)}{s}\right)^{k+1} \cdot \mathbb{E}\left[\left(X_j^{(n)}\right)^{k+1} \exp\left(\beta(n)\frac{a_j(n)}{s}X_j^{(n)}\right) \mathbb{1}_{\left\{X_j^{(n)} \ge t^*\right\}}\right]. \tag{4.15}$$

We now show that both $B_1(n)$ and $B_2(n)$ converge to 0 as $n \to \infty$. Note that (B.2), the definition of $\beta(n)$ in (4.10) and, since $r < k/(k+1)$, property (iii) of Proposition 4.1 imply that

$$\lim_{n\to\infty} n\left(\beta(n)\frac{a_{\max}(n)}{s}\right)^{k+1} = \lim_{n\to\infty}\left(\frac{a_{\max}(n)n}{s}\right)^{k+1}\left(\zeta n^{r-\frac{k}{k+1}}b(n)\right)^{k+1} = 0 \quad \text{and} \quad \lim_{n\to\infty}\left(\beta(n)\frac{a_{\max}(n)}{s}\right) = 0. \tag{4.16}$$

Combined with (4.14) and recalling that $a_{max}(n) := \max_{1\le j\le n} a_j(n)$, this shows that $B_1(n) \to 0$ as $n \to \infty$.

Next, to bound $B_2(n)$, first note that by Hölder's inequality, for any $\varepsilon > 0$ we have

$$\begin{aligned}&\mathbb{E}\left[\left(X_1^{(n)}\right)^{k+1} \exp\left(\beta(n)\frac{a_{max}(n)}{s}X_1^{(n)}\right)\mathbb{1}_{\{X_1^{(n)} \ge t^*\}}\right] \\ &\le \mathbb{E}\left[\left(X_1^{(n)}\right)^{(k+1)\cdot\frac{1+\varepsilon}{\varepsilon}}\mathbb{1}_{\{X_1^{(n)}\ge t^*\}}\right]^{\frac{\varepsilon}{1+\varepsilon}} \cdot \mathbb{E}\left[\exp\left((1+\varepsilon)\beta(n)\frac{a_{max}(n)}{s}X_1^{(n)}\right)\mathbb{1}_{\{X_1^{(n)}\ge t^*\}}\right]^{\frac{1}{1+\varepsilon}}.\end{aligned} \tag{4.17}$$

Due to the finiteness of the moments of X_1 assumed in (3.3), the first limit in (4.16) yields

$$\limsup_{n\to\infty} n \cdot \left(\beta(n)\frac{a_{max}(n)}{s}\right)^{k+1} \mathbb{E}\left[\left(X_1^{(n)}\right)^{(k+1)\cdot\frac{1+\varepsilon}{\varepsilon}}\mathbb{1}_{\{X_1^{(n)}\ge t^*\}}\right]^{\frac{\varepsilon}{1+\varepsilon}} = 0.$$

When combined with (4.15) and (4.17), to prove the convergence of $B_2(n)$ to zero, it clearly suffices to show that

$$\limsup_{n\to\infty} \frac{1}{b(n)n^r}\mathbb{E}\left[\exp\left((1+\varepsilon)\beta(n)\frac{a_{max}(n)}{s}X_1^{(n)}\right)\mathbb{1}_{\{X_1^{(n)}\ge t^*\}}\right]^{\frac{1}{1+\varepsilon}} < \infty \tag{4.18}$$

for $\zeta < (1+\varepsilon)^{-1}\left(\frac{x}{s} - \frac{s_1}{s}m\right)^{r-1}$ and the claim follows as $\varepsilon \to 0$. To derive an upper bound for the expectation in (4.18) we will use the following integration-by-parts formula.

Lemma 4.5 (Integration by parts)**.** *For any random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and any $\alpha > 0$, q_1, $q_2 \in \mathbb{R}$ with $q_1 < q_2$ the following relation holds:*

$$\mathbb{E}\left[\exp(\alpha X)\mathbb{1}_{\{q_1\le X\le q_2\}}\right] = \alpha\int_{q_1}^{q_2}\exp(\alpha z)\,\mathbb{P}(X\ge z)\,dz + \exp(\alpha q_1)\,\mathbb{P}(X\ge q_1) - \exp(\alpha q_2)\,\mathbb{P}(X>q_2).$$

Recalling that $X_j^{(n)} = X_j \mathbb{1}_{\{X_j < t_2(n)\}}$, applying Lemma 4.5 with $q_1 = t^*$ and $q_2 = t_2(n)$, we deduce that

$$\begin{aligned} &\frac{1}{b(n)n^r}\mathbb{E}\left[\exp\left((1+\varepsilon)\beta(n)\frac{a_{max}(n)}{s}X_1^{(n)}\right)\mathbb{1}_{\{X_1^{(n)}\geq t^*\}}\right] \\ &\quad\leq \frac{1}{b(n)n^r}\int_{t^*}^{t_2(n)}(1+\varepsilon)\beta(n)\frac{a_{max}(n)}{s}\exp\left((1+\varepsilon)\beta(n)\frac{a_{max}(n)}{s}z\right)\mathbb{P}\left(X_1\geq z\right)dz \\ &\quad\quad + \frac{1}{b(n)n^r}\exp\left((1+\varepsilon)\beta(n)\frac{a_{max}(n)}{s}t^*\right). \end{aligned} \tag{4.19}$$

Since $b(n)n^r \to \infty$, the second term on the right-hand side of (4.19) converges to zero by the second limit in (4.16). Now, let $\zeta^* := \zeta \cdot \left(\frac{x}{s} - \frac{s_1}{s}m\right)$. Inserting the upper bound (3.4) on the tail of X_1, substituting $y := (t_2(n))^{-1}z$ and recalling the definition of $\beta(n)$ from (4.10), we see that the first term on the right-hand side of (4.19) is bounded above by

$$(1+\varepsilon)\zeta^*\frac{b(t_2(n))}{b(n)}\frac{na_{max}(n)}{s} \quad \cdot \int_{\frac{t^*}{t_2(n)}}^{1} I_n(y)dy, \tag{4.20}$$

where the integrand $I_n(\cdot)$ is given by

$$I_n(y) := c_2\left(t_2(n)y\right)\exp\left\{n^r b\left(t_2(n)\right)\left((1+\varepsilon)\zeta^*\frac{na_{max}(n)}{s}y - \frac{b(t_2(n)y)}{b(t_2(n))}\left(\frac{x}{s}-\frac{s_1}{s}m\right)^r y^r\right)\right\}, \quad y\in(0,1].$$

Since $b(\cdot)$ is slowly varying and condition (B.2) holds, we see that the coefficient in front of the integral in (4.20) converges to $(1+\varepsilon)\zeta^*$ as $n\to\infty$. It now remains to show that, for every $\zeta^* < (1+\varepsilon)^{-1}\left(\frac{x}{s}-\frac{s_1}{s}m\right)^r$, the integral in (4.20) stays bounded as $n\to\infty$. By the assumption that $b(\cdot)$ is slowly varying and since $r<1$, for any fixed $y\in(0,1]$ and any $\zeta^* < (1+\varepsilon)^{-1}\left(\frac{x}{s}-\frac{s_1}{s}m\right)^r$, it follows that $I_n(y)\to 0$ as $n\to\infty$. Therefore, we need to examine the lower limit of integration $y_n := t^*/(t_2(n))$ and show that $I_n(y_n)$ stays bounded as $n\to\infty$. Recalling that $t_2(n) = n(\frac{x}{s}-\frac{s_1}{s}m)$ and $\zeta^* = \zeta(\frac{x}{s}-\frac{s_1}{s}m)$, note that

$$I_n(y_n) = c_2(t^*)\exp\left\{n^{r-1}b(t_2(n))(1+\varepsilon)\zeta\frac{na_{\max}(n)}{s}t^* - b(t^*)(t^*)^r\right\}.$$

Since $na_{max}(n) \sim s$, $b(t_2(n)) \sim b(n)$ and $n^{r-1}b(n) \to 0$ as $n\to\infty$, it follows that $\limsup_{n\to\infty} I_n(y_n) < \infty$.

Thus, we have shown that B_2^n converges to zero as $n\to\infty$ and hence, that $B_0 = 0$. This completes the proof of Proposition 4.4, and hence, the upper bound (4.2) and Theorem 1 follow. □

5 Examples

5.1 Example 1: Random Weights

We consider a sequence of strictly positive i.i.d. random variables $\{\theta_j\}_{j\in\mathbb{N}}$ on $(\Omega,\mathcal{F},\mathbb{P})$ and assume that they are $\mathbb{P}$-almost surely uniformly bounded, that is, their essential supremum is finite:

$$M^* := \inf\left\{a\in\mathbb{R} : \mathbb{P}\left(\theta_1 > a\right) = 0\right\} < \infty. \tag{5.1}$$

Furthermore, define the triangular array of weights $\{a_j(n,\theta_1,...,\theta_n), j=1,\dots,n\}_{n\in\mathbb{N}}$ by

$$a_j(n,\theta_1,...,\theta_n) := \frac{\theta_j}{\sum\limits_{i=1}^{n}\theta_i}, \quad j=1,\dots,n, n\in\mathbb{N}, \tag{5.2}$$

let $a_{max}(n,\theta_1,\dots,\theta_n) = \max_{j=1,\dots,n} a_j(n,\theta_1,\dots,\theta_n)$ and let $\{\bar{S}_n\}_{n\in\mathbb{N}}$ be the corresponding sequence of weighted sums:

$$\bar{S}_n := \sum_{j=1}^{n} a_j(n,\theta_1,...,\theta_n)X_j = \sum_{j=1}^{n}\frac{\theta_j}{\sum\limits_{i=1}^{n}\theta_i}X_j, \quad n\in\mathbb{N}. \tag{5.3}$$

We prove a large deviation theorem for the sequence of random weighted sums $\{\bar{S}_n\}_{n\in\mathbb{N}}$, both in the "quenched" (i.e., conditioned on the weight sequence $\{\theta_j\}_{j\in\mathbb{N}}$), and "annealed" (i.e., averaged over the weight sequence) cases. Note that $\bar{S}_n$ is a random convex combination of the data $\{X_i\}$. If, instead, we set $a_j(n,\theta_1,\cdots,\theta_n) = \theta_j/\sqrt{\sum_{i=1}^n \theta_i^2}$, then $a(n) := (a_1(n),\dots,a_n(n))$ is a unit vector in $\mathbb{R}^n$ and $\bar{S}_n$ can be viewed as a one-dimensional random projection of the data vector $(X_1,\dots,X_n)$. The latter case is more involved and will be considered in a more general setting in forthcoming work.

Theorem 2 (Large Deviations for Random Weights, Stretched Exponential Tails)**.** *Let $\{X_j\}_{j\in\mathbb{N}}$ be a sequence of i.i.d. random variables such as in Theorem 1 and let $\{\theta_j\}_{j\in\mathbb{N}}$ be a sequence of i.i.d. random variables which is independent of the sequence $\{X_j\}_{j\in\mathbb{N}}$, and is almost surely uniformly bounded by M^* as specified in* (5.1). *Define $\bar{S}_n$ by* (5.3). *Then, for $x>m$, we have*

$$\lim_{n\to\infty}\frac{1}{b(n)\,n^r}\log\mathbb{P}\left(\bar{S}_n\geq x\middle|\,\theta_1,\theta_2,...\right) = -\left[\left(\frac{\mathbb{E}[\theta_1]}{M^*}\right)(x-m)\right]^r \quad \mathbb{P}\text{-a.s.}, \tag{5.4}$$

and

$$\lim_{n\to\infty}\frac{1}{b(n)\,n^r}\log\mathbb{P}\left(\bar{S}_n\geq x\right) = -\left[\left(\frac{\mathbb{E}[\theta_1]}{M^*}\right)(x-m)\right]^r. \tag{5.5}$$

Proof. The proof of (5.4) is a direct application of Theorem 1. First of all, note that for every $n\in\mathbb{N}$, $\sum_{j=1}^n a_j(n,\theta_1,...,\theta_n)=1$ almost surely, and hence $s_1=1$, where s_1 is the quantity defined in (B.1). Furthermore,

$$n\cdot a_{max}(n,\theta_1,...,\theta_n) = \frac{n\cdot\max\{\theta_j : 1\leq j\leq n\}}{\sum\limits_{i=1}^{n}\theta_i} = \frac{\max\{\theta_j : 1\leq j\leq n\}}{\frac{1}{n}\sum\limits_{i=1}^{n}\theta_i}. \tag{5.6}$$

It is easy to check that almost surely, $\max\{\theta_j : 1\leq j\leq n\}\to M^*$ as $n\to\infty$. By the strong law of large numbers, it follows that almost surely, $n\cdot a_{max}(n,\theta_1,...,\theta_n)\to s := M^*/\mathbb{E}[\theta_1]$ as $n\to\infty$. By Theorem 1 we conclude that, for $x>m$, the quenched asymptotics (5.4) are valid.

We now turn to the proof of (5.5). Note that we have

$$\mathbb{P}\left(\bar{S}_n\geq x\right) = \mathbb{P}\left(\frac{\frac{1}{n}\sum\limits_{j=1}^{n}\theta_jX_j}{\frac{1}{n}\sum\limits_{i=1}^{n}\theta_i}\geq x\right). \tag{5.7}$$

Now, $\frac{1}{n}\sum_{i=1}^n\theta_i\to\mathbb{E}[\theta_1]$, $\mathbb{P}$-almost surely, and the probability of a deviation decays exponentially in n, due to Cramér's Theorem (recall that the $\{\theta_i\}$ are uniformly bounded!).

We will now show that

$$\lim_{n\to\infty}\frac{1}{b\,(n)\,n^r}\log\mathbb{P}\left(\bar S_n\geq x\right)\approx\lim_{n\to\infty}\frac{1}{b\,(n)\,n^r}\log\mathbb{P}\left(\frac{1}{n}\sum_{j=1}^{n}\theta_jX_j\geq\mathbb{E}[\theta_1]x\right),\tag{5.8}$$

in the sense explained in (5.9) and (5.10) below. Fix $\delta>0$ and consider the events $F_n:=\{\frac{1}{n}\sum_{i=1}^{n}\theta_i\geq(1-\delta)\mathbb{E}[\theta_1]\}$ and their complements F_n^c for $n\in\mathbb{N}$. Then, $\mathbb{P}\left(\bar S_n\geq x\right)\leq\mathbb{P}(\frac{1}{n}\sum_{j=1}^{n}\theta_jX_j\geq(1-\delta)\mathbb{E}[\theta_1]x)+\mathbb{P}(F_n^c)$, and since $\mathbb{P}(F_n^c)$ decays exponentially in n, it follows that for any $\delta>0$,

$$\limsup_{n\to\infty}\frac{1}{b\,(n)\,n^r}\log\mathbb{P}(\bar S_n\geq x)\leq\limsup_{n\to\infty}\frac{1}{b\,(n)\,n^r}\log\mathbb{P}\left(\frac{1}{n}\sum_{j=1}^{n}\theta_jX_j\geq(1-\delta)\mathbb{E}[\theta_1]x\right).\tag{5.9}$$

On the other hand, with $G_n:=\{\frac{1}{n}\sum_{i=1}^{n}\theta_i\leq(1+\delta)\mathbb{E}[\theta_1]\}$, we have $\mathbb{P}(\bar S_n\geq x)\geq\mathbb{P}(\{\bar S_n\geq x\}\cap G_n)\geq\mathbb{P}(\frac{1}{n}\sum_{j=1}^{n}\theta_jX_j\geq(1+\delta)\mathbb{E}[\theta_1]x)-\mathbb{P}(G_n^c)$, and since $\mathbb{P}(G_n^c)$ decays exponentially in n, we have

$$\liminf_{n\to\infty}\frac{1}{b\,(n)\,n^r}\log\mathbb{P}(\bar S_n\geq x)\geq\liminf_{n\to\infty}\frac{1}{b\,(n)\,n^r}\log\mathbb{P}\left(\frac{1}{n}\sum_{j=1}^{n}\theta_jX_j\geq(1+\delta)\mathbb{E}[\theta_1]x\right).\tag{5.10}$$

Looking at the right-hand sides of (5.9) and (5.10) we are in the situation of Theorem 1 with i.i.d. random variables θ_jX_j and weights $a_j(n)=\frac{1}{n},j=1,\dots,n$ that clearly satisfy Assumption B with $s=s_1=1$ and $R(\nu,1)=1$ for all $\nu\in\mathbb{N}$. Considering the tail of θ_1X_1, we see that due to (3.4), for $t\geq t^*$, $\mathbb{P}(\theta_1X_1\geq t)\leq\mathbb{P}(X_1\geq t/M^*)\leq c_2(t/M^*)\exp(-b(t/M^*)t^r(M^*)^{-r})$. On the other hand, for $t\geq t^*$, again by (3.4), $\mathbb{P}(\theta_1X_1\geq t)\geq\mathbb{P}(\theta_1\geq M^*-\delta)\mathbb{P}(X_1\geq t/(M^*-\delta))\geq\mathbb{P}(\theta_1\geq M^*-\delta)c_1(t/(M^*-\delta))\exp(-b(t/(M^*-\delta))t^r(M^*-\delta)^{-r})$. The proof is completed by applying the lower and upper bounds in (4.1) and (4.2), respectively, and then sending $\delta\downarrow 0$ to obtain (5.5). □

Remark 5.1. The equality of the quenched and annealed rate functions in (5.4) and (5.5), respectively, is characteristic of our regime; it is in sharp contrast to the case of light-tailed random variables X_j, that is, random variables X_j satisfying (1.1). In the light-tailed case, $\mathbb{P}\left(\bar S_n\geq x\middle|\,\theta_1,\theta_2,...\right)$ and $\mathbb{P}\left(\bar S_n\geq x\right)$ both decay exponentially in n, but the rate functions will in general not be the same. This was one of the motivations for the present paper, and will be treated in forthcoming work.

5.2 Example 2: Kernel Functions

Kernel functions are an important tool to smooth data. For example, they are used as weighting functions in non-parametric regression. Applications include the approximation of probability density functions and conditional expectations.

Definition 5.1 (Kernel). *A kernel is an integrable function $k:[-1,1]\to[0,\infty)$ satisfying the following two requirements:*

(i) $\int_{-1}^{1}k(u)du=1.$

(ii) $k(-u)=k(u)\quad\forall u\in[0,1].$

Define the triangular array of weights $\{a_j(n),j=1,\dots,n\}_{n\in\mathbb{N}}$ by

$$a_j(n):=\frac{1}{n}\cdot k\left(2\cdot\frac{j-n/2}{n}\right),\quad j=1,\dots,n,n\in\mathbb{N},\tag{5.11}$$

and let $\{\bar{S}_n\}_{n\in\mathbb{N}}$ be the corresponding sequence of weighted sums:

$$\bar{S}_n := \sum_{j=1}^{n} a_j(n) X_j = \frac{1}{n}\sum_{j=1}^{n} k\left(2\cdot\frac{j-n/2}{n}\right) X_j, \quad n\in\mathbb{N}. \tag{5.12}$$

Theorem 3 (Large Deviations for Kernel Weighted Sums, Stretched Exponential Tails)**.** *Let $\{X_j\}_{j\in\mathbb{N}}$ be a sequence of i.i.d. random variables such as in Theorem 1 and let $k:[-1,1]\to[0,\infty)$ be a kernel. Define $\bar{S}_n$ by* (5.12)*. Then, for $x>m$, we have*

$$\lim_{n\to\infty}\frac{1}{b(n)\,n^r}\log\mathbb{P}\left(\bar{S}_n\geq x\right) = -\left(\sup_{x\in[-1,1]} k(x)\right)^{-r}(x-m)^r. \tag{5.13}$$

Proof. The proof is a direct application of Theorem 1. Recall the definition of the quantities $\{s_\nu\}_{\nu\in\mathbb{N}}$ from Assumption B. It is straightforward to check that $s_\nu = \int_{-1}^{1}(k(u))^\nu du$ (in particular, $s_1=1$). Therefore,

$$s = \lim_{\nu\to\infty}\left(\int_{-1}^{1}(k(u))^\nu du\right)^{1/\nu}.$$

Since the p-norm converges to the supremum norm as $p\to\infty$, we conclude that $s = \sup_{x\in[-1,1]} k(x)$. □

Acknowledgments. N. Gantert and F. Rembart thank the Division of Applied Mathematics, Brown University, Providence, for its hospitality. N. Gantert further thanks ICERM, Providence, for an invitation to the program "Computational Challenges in Probability" where this work was initiated.

References

[1] Bingham, N., Goldie, C., and Teugels, J. (1987). *Regular Variation.* Cambridge University Press. MR-0898871

[2] Bingham, E., and Mannila, H. (2001). Random projection in dimensionality reduction: Application to image and text data. *Proc. of Seventh ACM SIGKDD International Conf. on Knowledge Discovery and Data Mining.*

[3] Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles*, 736:5–23.

[4] Dembo, A. and Zeitouni, O. (1993). *Large Deviation Techniques and Applications.* Jones and Bartlett, Boston, MA. MR-1202429

[5] Diaconis, P. and Freedman, D. (1984) Asymptotics of graphical projection pursuit. *Ann. Statist.* **12** 793–815. MR-0751274

[6] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling extremal events: for insurance and finance* Springer-Verlag, Berlin. MR-1458613

[7] Galambos, J. and Seneta, E. (1973). Regularly varying sequences. *Proceedings of the American Mathematical Society*, 41(1):110–116. MR-0323963

[8] Gantert, N. (1996). Large deviations for a heavy-tailed mixing sequence. *Unpublished.*

[9] Kiesel, R. and Stadtmüller, U. (2000). A large deviation principle for weighted sums of independent and identically distributed random variables. *Journal of Mathematical Analysis*, 251:929–939. MR-1794779

[10] Nagaev, S. V. (1969). Integral limit theorems taking large deviations into account when Cramér's condition does not hold, I. *Theory of Probability and its Applications*, 14(1):51–64.

[11] Nagaev, S. V. (1979). Large deviations for sums of independent random variables. *Annals of Probability*, 7:745–789. MR-0542129