GUMBEL AND FRÉCHET CONVERGENCE OF THE MAXIMA OF INDEPENDENT RANDOM WALKS

THOMAS MIKOSCH,* AND
JORGE YSLAS,* ** *University of Copenhagen*

Abstract

We consider point process convergence for sequences of independent and identically distributed random walks. The objective is to derive asymptotic theory for the largest extremes of these random walks. We show convergence of the maximum random walk to the Gumbel or the Fréchet distributions. The proofs depend heavily on precise large deviation results for sums of independent random variables with a finite moment generating function or with a subexponential distribution.

Keywords: Large deviation; subexponential distribution; regular variation; extreme value theory; Gumbel distribution; Fréchet distribution; maximum random walk.

2010 Mathematics Subject Classification: Primary 60F10
Secondary 60F05; 60G50; 60G55; 60G70

1. Introduction

Let (X_i) be an independent and identically distributed (i.i.d.) sequence of random variables with generic element X, distribution F, and right tail $\overline{F} = 1 - F$. Define the corresponding partial sum process

$$S_0 = 0, \quad S_n = X_1 + \cdots + X_n, \quad n \geq 1.$$

Consider i.i.d. copies $(S_{ni})_{i=1,2,...}$ of S_n. We also introduce an integer sequence (p_n) such that $p = p_n \to \infty$ as $n \to \infty$. We are interested in the limiting behavior of the k largest values among $(S_{ni})_{i=1,...,p}$, in particular in the possible limit laws of the maximum $\max_{i=1,...,p} S_{ni}$. More generally, writing ε_x for the Dirac measure at x, we are interested in the limiting behavior of the point processes

$$N_p = \sum_{i=1}^{p} \varepsilon_{c_n^{-1}(S_{ni}-d_n)} \xrightarrow{d} N, \quad n \to \infty, \tag{1.1}$$

for suitable constants $c_n > 0$ and $d_n \in \mathbb{R}$ toward a Poisson random measure N with Radon mean measure μ (we write PRM(μ)).

Our main motivation for this work comes from random matrix theory, in particular when dealing with sample covariance matrices. Their entries are dependent random walks. However, in various situations the theory can be modified in such a way that it suffices to study independent random walks. We refer to Section 4.6 for a discussion.

Relation (1.1) is equivalent to the following limit relations for the tails:

$$p_n \, \mathbb{P}(c_n^{-1}(S_n - d_n) \in (a, b]) \to \mu(a, b],$$

for any $a < b$ provided that $\mu(a, b] < \infty$; see Resnick [29, Theorem 5.3]. These conditions involve precise large deviation probabilities for the random walk (S_n); in Section 3 we provide some results which are relevant in this context.

We distinguish between two types of precise large deviation results:

- normal approximation,

- subexponential approximation.

The normal approximation can be understood as an extension of the central limit theorem for $(S_n/\sqrt{n})$ toward increasing intervals. This approximation causes the maxima of $(S_{ni}/\sqrt{n})$ to behave like the maxima of an i.i.d. normal sequence, that is, these maxima converge in distribution to the Gumbel distribution. This is in contrast to the subexponential approximation, which requires that F is a so-called *subexponential* distribution; see Section 2.1. In particular, $\overline{F}$ is heavy-tailed in the sense that the moment generating function does not exist. This fact implies that $\mathbb{P}(S_n > x_n) \sim n\,\overline{F}(x_n)$ for sufficiently fast increasing sequences $x_n \to \infty$. Hence $n\,\overline{F}(x_n)$ dominates $\mathbb{P}(S_n > x_n)$ at sufficiently high levels x_n and, as in limit theory for the maxima of an i.i.d. sequence, $\overline{F}$ determines the type of the limit distribution of the maxima of (S_{ni}) as well as the normalizing and centering constants. In this case we also assume that F belongs to the maximum domain of attraction (MDA) of the Gumbel or Fréchet distributions, and we borrow the known normalizing and centering constants from these MDAs. Thus, in the case of the MDA of the Gumbel distribution, the maxima of (S_{ni}) may converge to the Gumbel distribution due to two distinct mechanisms: the normal approximation at medium-high thresholds or the subexponential approximation at high-level thresholds. In the case of the MDA of the Fréchet distribution, two distinct approximations are possible: Gumbel approximation at medium-high thresholds and Fréchet approximation at high-level thresholds provided the distribution has finite second moment. If this condition is not satisfied, only the Fréchet approximation is possible.

The paper is organized as follows. In Section 2 we introduce the necessary notions for this paper: subexponential and regularly varying distributions (Section 2.1), maximum domain of attraction and relevant distributions in it (Section 2.2), point process convergence of triangular arrays toward Poisson random measures (Section 2.3), and precise large deviations (Section 2.4). Due to the importance of the latter topic we devote Section 3 to it and collect some of the known precise large deviation results in the case when the moment generating function is finite in some neighborhood of the origin and for subexponential distributions. The main results of this paper are formulated in Section 4. Based on the large deviation results of Section 3, we give sufficient conditions for the point process convergence relation (1.1) to hold and we clarify which rates of growth are possible for $p_n \to \infty$. In particular, we consider the case when p_n in (1.1) is replaced by $k_n = [n/r_n]$ for some integer sequence $r_n \to \infty$ and n is replaced by r_n. This means that we are interested in (1.1) when $S_{ni} = S_{r_n i} - S_{r_n(i-1)}$, $i = 1, \ldots, k_n$, are i.i.d. block sums. We also discuss extensions of these results to stationary regularly varying sequences (Section 4.3.3) and i.i.d. multivariate regularly varying sequences (Section 4.3.4).

2. Preliminaries I

2.1. Subexponential and regularly varying distributions

We are interested in the class $\mathcal{S}$ of *subexponential distributions* F, that is, it is a distribution supported on $[0, \infty)$ such that, for any $n \geq 2$,

$$\mathbb{P}(S_n > x) \sim n\,\overline{F}(x), \quad x \to \infty.$$

For an encyclopedic treatment of subexponential distributions, see Foss, Korshunov, and Zachary [10]. In insurance mathematics, $\mathcal{S}$ is considered a natural class of heavy-tailed distributions. In particular, F does not have a finite moment generating function; see Embrechts, Klüppelberg, and Mikosch [8, Lemma 1.3.5].

The regularly varying distributions are another class of heavy-tailed distributions supported on $\mathbb{R}$. We say that X and its distribution F are *regularly varying with index* $\alpha > 0$ if there are a slowly varying function L and constants $p_\pm$ such that $p_+ + p_- = 1$ and

$$F(-x) \sim p_- \, x^{-\alpha} \, L(x) \quad \text{and} \quad \overline{F}(x) \sim p_+ \, x^{-\alpha} \, L(x), \quad x \to \infty. \tag{2.1}$$

A non-negative regularly varying X is subexponential (see [8, Corollary 1.3.2]).

2.2. Maximum domains of attraction

We call a non-degenerate distribution H an *extreme value distribution* if there exist constants $c_n > 0$ and $d_n \in \mathbb{R}$, $n \geq 1$, such that the maxima $M_n = \max(X_1, \ldots, X_n)$ satisfy the limit relation

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} Y \sim H, \quad n \to \infty. \tag{2.2}$$

In the context of this paper we deal with two standard extreme value distributions: the Fréchet distribution $\Phi_\alpha(x) = \exp(-x^{-\alpha})$, $x > 0$, and the Gumbel distribution $\Lambda(x) = \exp(-\exp(-x))$, $x \in \mathbb{R}$. As a matter of fact, the third type of extreme value distribution – the Weibull distribution – cannot appear since (2.2) is only possible for X with finite right endpoint but a random walk is not bounded from above by a constant. We say that the distribution F of X is in the *maximum domain of attraction* of the extreme value distribution H ($F \in \mathrm{MDA}(H)$).

Example 2.1. A distribution $F \in \mathrm{MDA}(\Phi_\alpha)$ for some $\alpha > 0$ if and only if

$$\overline{F}(x) = \frac{L(x)}{x^\alpha}, \quad x > 0$$

(see [8, Section 3.3.1]). Then

$$c_n^{-1} M_n \xrightarrow{d} Y \sim \Phi_\alpha, \quad n \to \infty,$$

where (c_n) can be chosen such that $n \, \mathbb{P}(X > c_n) \to 1$.

Example 2.2. A distribution F with infinite right endpoint obeys $F \in \mathrm{MDA}(\Lambda)$ if and only if there exists a positive function $a(x)$ with derivative $a'(x) \to 0$ as $x \to \infty$ such that

$$\lim_{u \to \infty} \frac{\overline{F}(u + a(u)\,x)}{\overline{F}(u)} = \mathrm{e}^{-x}, \quad x \in \mathbb{R}$$

(see [8, Section 3.3.3]). Then

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} Y \sim \Lambda, \quad n \to \infty,$$

where (d_n) can be chosen such that $n \, \mathbb{P}(X > d_n) \to 1$ and $c_n = a(d_n)$.

The standard normal distribution $\Phi \in \mathrm{MDA}(\Lambda)$ and satisfies

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} Y \sim \Lambda, \quad n \to \infty, \tag{2.3}$$

where $c_n = 1/d_n$ and

$$d_n = \sqrt{2 \log n} - \frac{\log \log n + \log 4\pi}{2(2 \log n)^{1/2}}. \tag{2.4}$$

Since $d_n \sim \sqrt{2 \log n}$, we can replace c_n in (2.3) by $1/\sqrt{2 \log n}$, while d_n cannot be replaced by $\sqrt{2 \log n}$.

The standard lognormal distribution (i.e. $X = \exp(Y)$ for a standard normal random variable Y) is also in MDA(Λ). In particular, one can choose

$$c_n = d_n / \sqrt{2 \log n} \quad \text{and} \quad d_n = \exp\left(\sqrt{2 \log n} - \frac{\log \log n + \log 4\pi}{2(2 \log n)^{1/2}}\right) \qquad (2.5)$$

(see [8, page 156]).

The standard Weibull distribution has tail $\overline{F}(x) = \exp(-x^{-\tau})$, $x > 0$, $\tau > 0$. We consider a distribution F on $(0, \infty)$ with a Weibull-type tail $\overline{F}(x) \sim c\, x^\beta \exp(-\lambda x^\tau)$ for constants $c, \beta, \lambda, \tau > 0$. Then $F \in$ MDA(Λ) and one can choose

$$c_n = (\lambda \tau)^{-1} s_n^{1/\tau - 1} \quad \text{and} \quad d_n = s_n^{1/\tau} + \frac{1}{\tau} s_n^{1/\tau - 1}\left(\frac{\beta}{\lambda \tau} \log s_n + \frac{\log c}{\lambda}\right), \qquad (2.6)$$

where $s_n = \lambda^{-1} \log n$ (see [8, page 155]).

2.3. Point process convergence of independent triangular arrays

For further use we will need the following point process limit result (Resnick [29, Theorem 5.3]).

Proposition 2.1. *Let* $(X_{ni})_{n=1,2,\ldots;i=1,2,\ldots}$ *be a triangular array of row-wise i.i.d. random elements on some state space* $E \subset \mathbb{R}^d$ *equipped with the Borel* σ-*field* $\mathcal{E}$. *Let* μ *be a Radon measure on* $\mathcal{E}$. *Then*

$$\widetilde{N}_p = \sum_{i=1}^{p} \varepsilon_{X_{ni}} \xrightarrow{d} N, \quad n \to \infty,$$

holds for some PRM(μ) N *if and only if*

$$p\, \mathbb{P}(X_{n1} \in \cdot) \xrightarrow{v} \mu(\cdot), \quad n \to \infty,$$

where $\xrightarrow{v}$ *denotes vague convergence on E.*

2.4. Large deviations

Our main goal is to prove the point process convergence (1.1) for i.i.d. sequences (S_{ni}) of partial sum processes ($\mathbb{R}$- or $\mathbb{R}^d$-valued), properly normalized and centered. It follows from Proposition 2.1 that this means proving relations of the type

$$p\, \mathbb{P}(c_n^{-1}(S_n - d_n) \in (a, b]) \to \mu(a, b] \quad \text{or} \quad p\, \mathbb{P}(c_n^{-1}(S_n - d_n) > a) \to \mu(a, \infty),$$

provided $\mu(a, b] + \mu(a, \infty) < \infty$. Since $p = p_n \to \infty$, this means that $\mathbb{P}(c_n^{-1}(S_n - d_n) > a) \to 0$ as $n \to \infty$. We will refer to these vanishing probabilities as *large deviation probabilities*. In Section 3 we consider some of the well-known precise large deviation results in heavy- and light-tailed situations.

3. Preliminaries II: precise large deviations

In this section we collect some precise large deviation results in the light- and heavy-tailed cases.

3.1. Large deviations with normal approximation

We assume $\mathbb{E}[X] = 0$, $\mathrm{var}(X) = 1$ and write Φ for the standard normal distribution. We start with a classical result when X has finite exponential moments.

Theorem 3.1. (Petrov's theorem [26], Theorem 1, Chapter VIII.) *Assume that the moment generating function* $\mathbb{E}[\exp(hX)]$ *is finite in some neighborhood of the origin. Then the following tail bound holds for* $0 \leq x = \mathrm{o}(\sqrt{n})$:

$$\frac{\mathbb{P}(S_n/\sqrt{n} > x)}{\overline{\Phi}(x)} = \exp\left(\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)\right)\left[1 + O\left(\frac{x+1}{\sqrt{n}}\right)\right], \quad n \to \infty.$$

where $\lambda(t) = \sum_{k=0}^{\infty} a_k t^k$ *is the Cramér series whose coefficients* a_k *depend on the cumulants of* X, *and* $\lambda(t)$ *converges for sufficiently small values* $|t|$.

Under the conditions of Theorem 3.1, uniformly for $x = \mathrm{o}(n^{1/6})$,

$$\frac{\mathbb{P}(S_n/\sqrt{n} > x)}{\overline{\Phi}(x)} \to 1, \quad n \to \infty. \tag{3.1}$$

Theorem 7 in Chapter VIII of Petrov [26] considers the situation of Theorem 3.1 under the additional assumption that the cumulants of order $k = 3, \ldots, r+2$ of X vanish for some positive integer r. Then the coefficients $a_0, \ldots, a_{r-1}$ in the series $\lambda(t)$ vanish, and it is not difficult to see that (3.1) holds uniformly for $0 \leq x = \mathrm{o}(n^{(r+1)/(2(r+3))})$.

In Section VIII.3 of [26] we also find necessary and sufficient conditions for (3.1) to hold in certain intervals. The following result was proved by S. V. Nagaev [21] for $x \in (0, \sqrt{(s/2 - 1)\log n})$ and improved by Michel [18] for $x \in (0, \sqrt{(s-2)\log n})$. The statement of the proposition is sharp under the given moment condition; see Theorem 3.2 below.

Proposition 3.1. *Assume that* $\mathbb{E}[|X|^s] < \infty$ *for some* $s > 2$. *Then* (3.1) *holds uniformly for* $0 \leq x \leq \sqrt{(s-2)\log n}$.

3.2. Large deviations with normal/subexponential approximations

Cline and Hsing [4] (in an unpublished article) discovered that the subexponential class $\mathcal{S}$ of distributions exhibits a completely different kind of large deviation behavior.

Proposition 3.2. (Cline and Hsing [4].) *We consider a distribution* F *on* $(0, \infty)$ *with infinite right endpoint. Then the following statements hold.*

(1) $F \in \mathcal{S}$ *if and only if*

$$\lim_{x \to \infty} \frac{\overline{F}(x+y)}{\overline{F}(x)} = 1, \quad \textit{for any real } y, \tag{3.2}$$

and there exists a sequence $\gamma_n \to \infty$ *such that*

$$\lim_{n \to \infty} \sup_{x > \gamma_n} \frac{\mathbb{P}(S_n > x)}{n\,\overline{F}(x)} \leq 1.$$

(2) *If* $F \in \mathcal{S}$, *then there exists a sequence* $\gamma_n \to \infty$ *such that*

$$\lim_{n \to \infty} \sup_{x > \gamma_n} \left|\frac{\mathbb{P}(S_n > x)}{n\,\overline{F}(x)} - 1\right| = 0. \tag{3.3}$$

Remark 3.1. If F satisfies (3.2) we say that F is *long-tailed*; we write $F \in \mathcal{L}$. It is well known that $F \in \mathcal{S}$ implies $F \in \mathcal{L}$; see Embrechts *et al.* [8, Lemma 1.3.5, page 41]. The converse is not true.

Proposition 3.2 shows that the subexponential class is the one for which heavy-tailed large deviations are reasonable to study. Given that we know that F is long-tailed, F is subexponential if and only if a uniform large deviation relation of the type (3.3) holds.

Subexponential and normal approximations to large deviation probabilities have been studied in detail in various papers. Among them, large deviations for i.i.d. regularly varying random variables are perhaps studied best. S. V. Nagaev [25] formulated a seminal result about the large deviations of a random walk (S_n) in the case of regularly varying X with finite variance. He dedicated this theorem to his brother A. V. Nagaev, who had started this line of research in the 1960s; see for example [22] and [23].

Theorem 3.2. (Nagaev's theorem [22, 25].) *Consider an i.i.d. sequence (X_i) of random variables with $\mathbb{E}[X] = 0$, $\mathrm{var}(X) = 1$ and $\mathbb{E}[|X|^{2+\delta}] < \infty$ for some $\delta > 0$. Assume that $\overline{F}(x) = x^{-\alpha} L(x), x > 0,$ for some $\alpha > 2$ and a slowly varying function L. Then, for $x \geq \sqrt{n}$ as $n \to \infty$,*

$$\mathbb{P}(S_n > x) = \overline{\Phi}(x/\sqrt{n})\,(1 + o(1)) + n\,\overline{F}(x)\,(1 + o(1)).$$

In particular, if X satisfies (2.1) with constants $p_\pm$, then for any positive constant $c_1 < \alpha - 2$,

$$\sup_{1 < x/\sqrt{n} < \sqrt{c_1 \log n}} \left| \frac{\mathbb{P}(\pm S_n > x)}{\overline{\Phi}(x/\sqrt{n})} - 1 \right| \to 0, \quad n \to \infty, \tag{3.4}$$

and for any constant $c_2 > \alpha - 2$,

$$\sup_{x/\sqrt{n} > \sqrt{c_2 \log n}} \left| \frac{\mathbb{P}(\pm S_n > x)}{n\,\mathbb{P}(|X| > x)} - p_\pm \right| \to 0, \quad n \to \infty.$$

Remark 3.2. If X is regularly varying with index α, $\mathbb{E}[|X|^s]$ is finite (infinite) for $s < \alpha$ $(s > \alpha)$. Therefore the normal approximation (3.4) is in agreement with Proposition 3.1.

In the infinite variance regularly varying case this result is complemented by an analogous statement. It can be found in Cline and Hsing [4] and Denisov, Dieker, and Shneer [6].

Theorem 3.3. *Consider an i.i.d. sequence (X_i) of regularly varying random variables with index $\alpha \in (0, 2]$ satisfying (2.1). Assume $\mathbb{E}[X] = 0$ if this expectation is finite. Choose (a_n) such that*

$$n\,\mathbb{P}(|X| > a_n) + \frac{n}{a_n^2}\mathbb{E}[X^2\,\mathbf{1}(|X| \leq a_n)] = 1, \quad n = 1, 2, \dots,$$

and (γ_n) such that $\gamma_n/a_n \to \infty$ as $n \to \infty$. For $\alpha = 2$, also assume for sufficiently small $\delta > 0$,

$$\lim_{n \to \infty} \sup_{x > \gamma_n} \frac{n}{x^2}\,\frac{\mathbb{E}[X^2\,\mathbf{1}(|X| \leq x)]}{[n\,\mathbb{P}(|X| > x)]^\delta} = 0.$$

Choose (d_n) such that

$$d_n = \begin{cases} 0 & \alpha \in (0, 1) \cup (1, 2], \\ n\,\mathbb{E}[X\,\mathbf{1}(|X| \leq a_n)] & \alpha = 1. \end{cases} \tag{3.5}$$

Then the following large deviation result holds:

$$\sup_{x > \gamma_n} \left| \frac{\mathbb{P}(\pm (S_n - d_n) > x)}{n \, \mathbb{P}(|X| > x)} - p_\pm \right| \to 0, \quad n \to \infty.$$

Remark 3.3. The normalization (a_n) is chosen such that $a_n^{-1}(S_n - d_n) \xrightarrow{d} Y_\alpha$ for an α-stable random variable Y_α, $\alpha \in (0, 2]$. Therefore $\gamma_n^{-1}(S_n - d_n) \xrightarrow{\mathbb{P}} 0$. In the case $\alpha < 2$, in view of Karamata's theorem (see Bingham, Goldie, and Teugels [3]), it is possible to choose (a_n) according as $n \, \mathbb{P}(|X| > a_n) \to 1$. The case $\alpha = 2$ is delicate: in this case var(X) can be finite or infinite. In the former case (a_n) is proportional to $\sqrt{n}$, and in the latter case $(a_n/\sqrt{n})$ is a slowly varying sequence; see Feller [9] or Ibragimov and Linnik [15, Section II.6].

Normal and subexponential approximations to large deviation probabilities also exist for subexponential distributions that have all moments finite. Early on, this was observed by A. V. Nagaev [22, 23, 24]. Rozovskii [31] did not use the name of subexponential distribution, but the conditions on the tails of the distributions he introduced are 'close' to subexponentiality; he also allowed for distributions F supported on the whole real line. In particular, A. V. Nagaev and Rozovskii discovered that, in general, the x-regions where the normal and subexponential approximations hold are separated from each other. To make this precise, we call two sequences (ξ_n) and (ψ_n) *separating sequences for the normal and subexponential approximations to large deviation probabilities* if, for an i.i.d. sequence (X_i) with variance 1,

$$\sup_{x < \xi_n} \left| \frac{\mathbb{P}(S_n - \mathbb{E}[S_n] > x)}{\overline{\Phi}(x/\sqrt{n})} - 1 \right| \to 0,$$

$$\sup_{x > \psi_n} \left| \frac{\mathbb{P}(S_n - \mathbb{E}[S_n] > x)}{n \, \mathbb{P}(X > x)} - 1 \right| \to 0, \quad n \to \infty.$$

A. V. Nagaev and Rozovskii gave conditions under which (ψ_n) and (ξ_n) cannot have the same asymptotic order; that is, we necessarily have $\psi_n/\xi_n \to \infty$. In particular, in the x-region (ξ_n, ψ_n) neither the normal nor the subexponential approximation holds; Rozovskii [31] also provided large deviation approximations for $\mathbb{P}(S_n > x)$ for these regions involving $\overline{\Phi}(x/\sqrt{n})$ and a truncated Cramér series. Explicit expressions for (ψ_n) and (ξ_n) are in general hard to get. We focus on two classes of subexponential distributions where the separating sequences are known.

- *Lognormal-type tails.* We write $F \in \text{LN}(\gamma)$: for some constants $\beta, \xi \in \mathbb{R}$, $\gamma > 1$ and $\lambda, c > 0$,
$$\overline{F}(x) \sim c \, x^\beta \, (\log x)^\xi \, \exp(-\lambda \, (\log x)^\gamma), \quad x \to \infty.$$
In the notation $\text{LN}(\gamma)$ we suppress the dependence on β, ξ, λ, c.

- *Weibull-type tails.* We write $F \in \text{WE}(\tau)$: for some $\beta \in \mathbb{R}$, $\tau \in (0, 1)$, $\lambda, c > 0$,
$$\overline{F}(x) \sim c \, x^\beta \, \exp(-\lambda \, x^\tau), \quad x \to \infty.$$
In the notation $\text{WE}(\tau)$ we suppress the dependence on β, λ, c.

The name 'Weibull-type tail' is motivated by the fact that the Weibull distribution F with shape parameter $\tau \in (0, 1)$ belongs to $\text{WE}(\tau)$. Indeed, in this case $\overline{F}(x) = \exp(-\lambda x^\tau)$, $x > 0$,

TABLE 1: Separating sequences (ξ_n) and (ψ_n) for the normal and subexponential approximations of $\mathbb{P}(S_n - \mathbb{E}[S_n] > x)$. We also assume $\mathrm{var}(X) = 1$. Here (h_n), $(\widetilde{h}_n)$ are any sequences converging to infinity. For completeness, we also include the regularly varying class $\mathrm{RV}(\alpha)$. The table is taken from Mikosch and Nagaev [19].

$F \in$	ξ_n	ψ_n
$\mathrm{RV}(\alpha)$, $\alpha > 2$	$((\alpha - 2)n \log n)^{1/2}$	$((\alpha - 2)n \log n)^{1/2}$
$\mathrm{LN}(\gamma)$, $1 < \gamma < 2$	$(n (\log n)^\gamma)^{1/2}$	$(n (\log n)^\gamma)^{1/2}$
$\mathrm{LN}(\gamma)$, $\gamma \geq 2$	$(n (\log n)^\gamma)^{1/2}/\widetilde{h}_n$	$n^{1/2}(\log n)^{\gamma-1} h_n$
$\mathrm{WE}(\tau)$, $0 < \tau \leq 0.5$	$n^{1/(2-\tau)}/\widetilde{h}_n$	$n^{1/(2-2\tau)} h_n$
$\mathrm{WE}(\tau)$, $0.5 < \tau < 1$	$n^{2/3}/\widetilde{h}_n$	$n^{1/(2-2\tau)} h_n$

for positive parameters λ. Similarly, the lognormal distribution F belongs to $\mathrm{LN}(2)$. This is easily seen by an application of Mill's ratio: for a standard normal random variable Y,

$$\overline{F}(x) = \mathbb{P}(Y > \log x) \sim \frac{\exp(-(\log x)^2/2)}{\sqrt{2\pi} \log x}, \quad x \to \infty.$$

These classes of distributions have rather distinct tail behavior. It follows from the theory in Embrechts *et al.* [8, Sections 1.3 and 1.4] that membership of F in $\mathrm{RV}(\alpha)$, $\mathrm{LN}(\gamma)$ or $\mathrm{WE}(\tau)$ implies $F \in \mathcal{S}$. The case $\mathrm{WE}(\tau)$, $0 < \tau < 1$, was already considered by A. V. Nagaev [23, 24].

For the heaviest tails when $F \in \mathrm{LN}(\gamma)$, $1 < \gamma < 2$, we can still choose $\xi_n = \psi_n$. This means that one threshold sequence separates the normal and subexponential approximations to the right tail $\mathbb{P}(S_n - \mathbb{E}[S_n] > x)$. Rozovskii [31] discovered that the classes $\mathrm{LN}(\gamma)$, $\gamma \geq 2$, and $\mathrm{LN}(\gamma)$, $1 < \gamma < 2$ have rather distinct large deviation properties. In the case $\gamma \geq 2$ we cannot choose (ξ_n) and (ψ_n) the same. The class $\mathrm{LN}(\gamma)$ with $1 < \gamma < 2$ satisfies the conditions of Theorem 3b in Rozovskii [31], which implies that

$$\mathbb{P}(S_n - \mathbb{E}[S_n] > x) = \left[\overline{\Phi}(x/\sqrt{n})\mathbf{1}(x < \gamma_n) + n\overline{F}(x)\mathbf{1}(x > \gamma_n)\right](1 + \mathrm{o}(1))$$

uniformly for x, where $\gamma_n = (\lambda 2^{-\gamma+1})^{1/2} n^{1/2} (\log n)^{\gamma/2}$. For $\gamma = 2$ the conditions of Theorem 3a in [31] are satisfied: with $g(x) = \lambda(\log x)^2 - (\beta + 2) \log x - \xi \log(\log x) - \log c$ and as $n \to \infty$,

$$\mathbb{P}(S_n - \mathbb{E}[S_n] > x) = \left[\overline{\Phi}(x/\sqrt{n}) \mathbf{1}(x < \gamma_n) + n\overline{F}(x)\mathrm{e}^{(n(g'(x))^2)/2}\mathbf{1}(x > \gamma_n)\right](1 + \mathrm{o}(1)).$$

Direct calculation shows that $\mathbb{P}(S_n - \mathbb{E}[S_n] > \gamma_n) \sim \exp(\lambda)\, n\, \overline{F}(\gamma_n)$ while, uniformly for $x > \gamma_n h_n$, $h_n \to \infty$, we have $\mathbb{P}(S_n - \mathbb{E}[S_n] > x) \sim n\overline{F}(x)$.

It is interesting to observe that all but one class of subexponential distributions considered in Table 1 have the property that $c\, n \in (\psi_n, \infty)$ for any $c > 0$. The exception is $\mathrm{WE}(\tau)$ for $\tau \in (0.5, 1)$. This fact turns the investigation of the tail probabilities $\mathbb{P}(S_n - \mathbb{E}[S_n] > c\, n)$ into a complicated technical problem. The exponential ($\mathrm{WE}(1)$) and superexponential ($\mathrm{WE}(\tau)$), $\tau > 1$, classes do not contain subexponential distributions. The corresponding partial sums exhibit the light-tailed large deviation behavior of Petrov's Theorem 3.1. As a historical remark, Linnik [17] and S. V. Nagaev [21] determined lower separating sequences (ξ_n) for the normal approximation to the tails $\mathbb{P}(S_n - \mathbb{E}[S_n] > x)$ under the assumption that $\overline{F}$ is dominated by the tail of a regular subexponential distribution from the table.

Denisov *et al.* [6] and Cline and Hsing [4] considered a unified approach to subexponential large deviation approximations for general subexponential and related distributions. In particular, they identified separating sequences (ψ_n) for the subexponential approximation of the tails

$\mathbb{P}(S_n - \mathbb{E}[S_n] > x)$ for general subexponential distributions. Denisov *et al.* [6] also considered local versions, i.e. approximations to the tails $\mathbb{P}(S_n \in [x, x + T])$ for $T > 0$ as $x \to \infty$.

4. Main results

4.1. Gumbel convergence via normal approximations to large deviation probabilities for small x.

We assume that $\mathbb{E}[X] = 0$ and $\mathrm{var}(X) = 1$ and the large deviation approximation to the standard normal distribution Φ holds: for some $\gamma_n \to \infty$,

$$\sup_{0 \leq x < \gamma_n} \left| \frac{\mathbb{P}(S_n/\sqrt{n} > x)}{\overline{\Phi}(x)} - 1 \right| \to 0, \quad n \to \infty. \tag{4.1}$$

We recall that $\Phi \in \mathrm{MDA}(\Lambda)$ and (2.3) holds. An analogous relation holds for the maxima of i.i.d. random walks $S_{n1}/\sqrt{n}, \ldots, S_{np}/\sqrt{n}$, as follows from the next result.

Theorem 4.1. *Assume that* (4.1) *is satisfied for some* $\gamma_n \to \infty$. *Then*

$$p \, \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > d_p + x/d_p \right) \to \mathrm{e}^{-x}, \quad n \to \infty, \quad x \in \mathbb{R}, \tag{4.2}$$

holds for any integer sequence $p_n \to \infty$ *such that* $p_n < \exp(\gamma_n^2/2)$ *and* (d_p) *is defined in* (2.4). *Moreover, for the considered* (p_n), (4.2) *is equivalent to either of the following limit relations.*

(1) *For* $\Gamma_i = E_1 + \cdots + E_i$ *and an i.i.d. standard exponential sequence* (E_i), *the following point process convergence holds on the state space* $\mathbb{R}$:

$$N_p = \sum_{i=1}^{p} \varepsilon_{d_p \, (S_{ni}/\sqrt{n} - d_p)} \overset{d}{\to} N = \sum_{i=1}^{\infty} \varepsilon_{- \log \Gamma_i}, \tag{4.3}$$

where N *is* $\mathrm{PRM}(- \log \Lambda)$ *on* $\mathbb{R}$.

(2) *Gumbel convergence of the maximum random walk*:

$$d_p \max_{i=1,\ldots,p} (S_{ni}/\sqrt{n} - d_p) \overset{d}{\to} Y \sim \Lambda, \quad n \to \infty.$$

Proof. In view of Proposition 2.1, it suffices for $N_p \overset{d}{\to} N$ to show that

$$p \, \mathbb{P}\left(d_p \left(\frac{S_n}{\sqrt{n}} - d_p \right) > x \right) = p \, \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > d_p + x/d_p \right) \sim p \, \overline{\Phi}(d_p + x/d_p) \to \mathrm{e}^{-x}, \quad x \in \mathbb{R}.$$

But this follows from (4.1) and the definition of (d_p) if we assume that $d_p + x/d_p < \gamma_n$, i.e. $p_n < \exp(\gamma_n^2/2)$ such that $p_n \to \infty$.

If $N_p \overset{d}{\to} N$, a continuous mapping argument implies that

$$\mathbb{P}(N_p(x, \infty) = 0) = \mathbb{P}\left(\max_{i=1,\ldots,p} d_p(S_{ni}/\sqrt{n} - d_p) \leq x \right)$$
$$\to \mathbb{P}(N(x, \infty) = 0)$$
$$= \Lambda(x), \quad x \in \mathbb{R}, \quad n \to \infty.$$

On the other hand, for $x \in \mathbb{R}$ as $n \to \infty$,

$$\mathbb{P}\Big(\max_{i=1,\dots,p} d_p(S_{ni}/\sqrt{n} - d_p) \leq x \Big) = \Big(1 - \frac{p\, \mathbb{P}(d_p(S_{n1}/\sqrt{n} - d_p) > x)}{p} \Big)^p$$
$$\to \exp(-e^{-x}),$$

if and only if (4.2) holds. $\square$

Remark 4.1. If we replace the quantities $(S_{ni}/\sqrt{n})_{i=1,\dots,p}$ in (4.3) with i.i.d. standard normal random variables, then this limit relation remains valid. This means that, under (4.1), for example under the assumption of a finite moment generating function in some neighborhood of the origin (see Section 3), the central limit theorem makes the tails of $(S_{ni}/\sqrt{n})_{i=1,\dots,p}$ almost indistinguishable from those of the standard normal distribution. This is in stark contrast to subexponential distributions, where the characteristics of $\overline{F}(x)$ show up in the tail $\mathbb{P}(S_{ni}/\sqrt{n} > x)$ for large values of x.

4.1.1. *The extreme values of i.i.d. random walks.* Write

$$S_{n,(p)} \leq \cdots \leq S_{n,(1)}$$

for the ordered values of $S_{n1}, \dots, S_{np}$ The following result is immediate from Theorem 4.1.

Corollary 4.1. *Assume that the conditions of Theorem 4.1 hold. Then*

$$\sqrt{2 \log p} \Big(\frac{S_{n,(1)}}{\sqrt{n}} - d_p, \dots, \frac{S_{n,(k)}}{\sqrt{n}} - d_p \Big) \xrightarrow{d} (-\log \Gamma_1, \dots, -\log \Gamma_k), \quad n \to \infty. \quad (4.4)$$

Moreover, if there is $\gamma_n \to \infty$ such that

$$\sup_{0 \leq x < \gamma_n} |\mathbb{P}(\pm S_n/\sqrt{n} > x)/\overline{\Phi}(x) - 1| \to 0 \quad \text{as } n \to \infty,$$

then we have

$$\mathbb{P}\Big(\max_{i=1,\dots,p} d_p(S_{ni}/\sqrt{n} - d_p) \leq x, \min_{i=1,\dots,p} d_p(S_{ni}/\sqrt{n} + d_p) \leq y \Big)$$
$$\to \Lambda(x)(1 - \Lambda(-y)), \quad x, y \in \mathbb{R}, \quad n \to \infty. \quad (4.5)$$

Proof. We observe that $d_p/\sqrt{2 \log p} \to 1$. Then (4.3) and the continuous mapping theorem imply that (4.4) holds for any fixed $k \geq 1$.

We observe that

$$\mathbb{P}\Big(\max_{i=1,\dots,p} d_p(S_{ni}/\sqrt{n} - d_p) \leq x, \min_{i=1,\dots,p} d_p(S_{ni}/\sqrt{n} + d_p) \leq y \Big)$$
$$= \mathbb{P}\Big(\max_{i=1,\dots,p} d_p(S_{ni}/\sqrt{n} - d_p) \leq x \Big)$$
$$\quad - \mathbb{P}\Big(\max_{i=1,\dots,p} d_p(S_{ni}/\sqrt{n} - d_p) \leq x, \min_{i=1,\dots,p} d_p(S_{ni}/\sqrt{n} + d_p) > y \Big)$$
$$= P_1(x, y) - P_2(x, y).$$

TABLE 2: Upper bounds for p.

Example 4.1	Petrov case	$\exp(o(n^{1/3}))$
Example 4.2	$\mathbb{E}[|X|^s] < \infty$, $s > 2$	$n^{(s-2)/2}$
Example 4.3	RV(α), $\alpha > 2$, $c < \alpha - 2$	$n^{c/2}$
Example 4.4	LN(γ), $\gamma > 1$	$\exp(o((\log n)^\gamma))$
Example 4.5	WE(τ), $\tau \le 0.5$	$\exp(o(n^{\tau/(2-\tau)}))$
Example 4.5	WE(τ), $\tau \in (0.5, 1)$	$\exp(o(n^{1/3}))$

Of course, $P_1(x, y) \to \Lambda(x)$. On the other hand,

$$
\begin{aligned}
P_2(x, y) &= \mathbb{P}\left(\bigcap_{i=1}^{p} \{S_{ni}/\sqrt{n} \le d_p + x/d_p, S_{ni}/\sqrt{n} > -d_p + y/d_p\} \right) \\
&= (\mathbb{P}(-d_p + y/d_p < S_{n1}/\sqrt{n} \le d_p + x/d_p))^p \\
&= \exp(p \log(1 - \mathbb{P}(S_{n1}/\sqrt{n} > d_p + x/d_p) - \mathbb{P}(S_{n1}/\sqrt{n} \le -d_p + y/d_p))) \\
&\to \exp(-(e^{-x} + e^y)) \\
&= \Lambda(x)\Lambda(-y).
\end{aligned}
$$

The last step follows from a Taylor expansion of the logarithm and Theorem 4.1. This proves (4.5). □

4.1.2. *Examples.* In this section we verify the assumptions of Theorem 4.1 for various classes of distributions F. We always assume $\mathbb{E}[X] = 0$ and var(X) = 1.

Example 4.1. Assume the existence of the moment generating function of X in some neighborhood of the origin. Petrov's Theorem 3.1 ensures (4.3) for $p \le \exp(o(n^{1/3}))$.

Example 4.2. Assume $\mathbb{E}[|X|^s] < \infty$ for some $s > 2$. Proposition 3.1 ensures that (4.3) for $p \le n^{(s-2)/2}$.

Example 4.3. Assume that X is regularly varying with index $\alpha > 2$. Then we can apply Nagaev's Theorem 3.2 with $\gamma_n = \sqrt{c \log n}$ for any $c < \alpha - 2$ and (4.3) holds for $p \le n^{c/2}$. This is in agreement with Example 4.2.

Example 4.4. Assume that X has a distribution in LN(γ) for some $\gamma > 1$. From Table 1, $\gamma_n = o((\log n)^{\gamma/2})$, and (4.3) holds for $p \le \exp(o((\log n)^\gamma))$.

Example 4.5. Assume that $F \in$ WE(τ), $0 < \tau < 1$. Table 1 yields $\gamma_n = o(n^{\tau/(2(2-\tau))})$ for $\tau \le 0.5$, hence $p \le \exp(o(n^{\tau/(2-\tau)}))$, and for $\tau \in (0.5, 1)$, $\gamma_n = o(n^{1/6})$ and $p \le \exp(o(n^{1/3}))$.

We summarize these examples in Table 2.

4.1.3. *The extremes of the blocks of a random walk.* We consider a random walk S_n with i.i.d. step sizes X_i with $\mathbb{E}[X] = 0$ and var(X) = 1, and with distribution F, and any integer sequence $r_n \to \infty$ such that $k_n = [n/r_n] \to \infty$ as $n \to \infty$. Set $S_{ni} = S_{r_n i} - S_{r_n(i-1)}$, i.e. this is the sum of the ith block $X_{r_n(i-1)+1}, \ldots, X_{r_n i}$. Then we are in the setting of Theorem 4.1 if we replace p_n

TABLE 3: Lower bounds on the block size r_n.

Example 4.1	Petrov case	$r_n/(\log n)^3 \to \infty$
Example 4.3	RV(α), $\alpha > 2$	$r_n > n^{2/(\alpha-\varepsilon)}$ any $\varepsilon \in (0, \alpha-2)$
Example 4.2	$\mathbb{E}[|X|^s] < \infty$, $s > 2$	$r_n > n^{2/s}$
Example 4.4	LN(γ), $\gamma > 1$	$r_n/\exp((2\log n)^{1/\gamma}) \to \infty$
Example 4.5	WE(τ), $\tau \le 0.5$	$r_n/(\log n)^{(2-\tau)/\tau} \to \infty$
Example 4.5	WE(τ), $\tau \in (0.5, 1)$	$r_n/(\log n)^3 \to \infty$

by k_n and n by r_n. We are interested in the following result for the point process of the block sums of S_n with length r_n (see (4.3)):

$$N_{k_n} = \sum_{i=1}^{k_n} \varepsilon_{d_{k_n}((S_{r_ni}-S_{r_n(i-1)})/\sqrt{r_n}-d_{k_n})} \xrightarrow{d} N = \sum_{i=1}^{\infty} \varepsilon_{-\log \Gamma_i}. \qquad (4.6)$$

This means we are looking for (r_n) such that $n/r_n < \exp(\gamma_{r_n}^2/2)$. This amounts to the conditions on (r_n) shown in Table 3.

This table shows convincingly that, the heavier the tails, the larger we have to choose the block length r_n. Otherwise, the normal approximation does not function sufficiently well simultaneously for the block sums $S_{r_ni} - S_{r_n(i-1)}$, $i = 1, \ldots, k_n$. In particular, in the regularly varying case we always need that r_n grows polynomially.

Notice that we have from (4.6) in particular

$$\frac{d_{k_n}}{\sqrt{r_n}} \max_{i=1,\ldots,k_n} (S_{r_ni} - S_{r_n(i-1)} - \sqrt{r_n} d_{k_n}) \xrightarrow{d} -\log \Gamma_1 \sim \Lambda, \quad n \to \infty.$$

The normalization $d_{k_n}/\sqrt{r_n}$ is asymptotic to $\sqrt{(2\log k_n)/r_n}$.

4.2. Gumbel convergence via the subexponential approximation to large deviation probabilities for very large x

In this section we will exploit the subexponential approximation to large deviation probabilities for subexponential distributions F, that is,

$$\sup_{x > \gamma_n} \left| \frac{\mathbb{P}(S_n - \mathbb{E}[S_n] > x)}{n\,\mathbb{P}(X > x)} - 1 \right| \to 0, \qquad (4.7)$$

and we will also assume that $F \in \text{MDA}(\Lambda)$; see Example 2.2 for the corresponding MDA conditions and the definition of the centering constants (d_n) and the normalizing constants (c_n). Then, in particular, X has all moments finite. In this case, the Gumbel approximation of the point process of the (S_{ni}) is also possible.

Theorem 4.2. *Assume that $F \in \text{MDA}(\Lambda) \cap \mathcal{S}$, the subexponential approximation (4.7) holds, and for sufficiently large n and an integer sequence $p_n \to \infty$,*

$$d_{np} + x\,c_{np} > \gamma_n, \quad \text{for any } x < 0, \qquad (4.8)$$

where (d_{np}) and (c_{np}) are the subsequences of (d_n) and (c_n), respectively, evaluated at np. Then we have

$$p\,\mathbb{P}(S_n - \mathbb{E}[S_n] > d_{np} + x\,c_{np}) \to \mathrm{e}^{-x}, \quad x \in \mathbb{R}, \quad n \to \infty. \tag{4.9}$$

Moreover, (4.9) is equivalent to either of the following limit relations.

(1) *Point process convergence to a Poisson process on the state space $\mathbb{R}$:*

$$N_p = \sum_{i=1}^{p} \varepsilon_{c_{np}^{-1}(S_{ni} - \mathbb{E}[S_n] - d_{np})} \xrightarrow{d} N, \quad n \to \infty, \tag{4.10}$$

where $N \sim \mathrm{PRM}(-\log \Lambda)$; see Theorem 4.1.

(2) *Gumbel convergence of the maximum random walk:*

$$\max_{i=1,\ldots,p} c_{np}^{-1}((S_{ni} - \mathbb{E}[S_n]) - d_{np}) \xrightarrow{d} Y \sim \Lambda, \quad n \to \infty. \tag{4.11}$$

Proof. If $d_{np} + x\,c_{np} > \gamma_n$ for every $x < 0$, then it holds for $x \in \mathbb{R}$. Therefore (4.7) applies. Since $F \in \mathrm{MDA}(\Lambda) \cap \mathcal{S}$ and by definition of (c_n) and (d_n), we have

$$p\,\mathbb{P}(S_n - \mathbb{E}[S_n] > d_{np} + x\,c_{np}) \sim p\,n\,\mathbb{P}(X > d_{np} + x\,c_{np}) \to \mathrm{e}^{-x}, \quad x \in \mathbb{R}, \quad n \to \infty,$$

proving (4.9). Proposition 2.1 yields the equivalence of (4.10) and (4.9). The equivalence of (4.10) and (4.11) follows from a standard argument. $\square$

Remark 4.2. Since $a(x)$ defined in Example 2.2 has density $a'(x) \to 0$ as $x \to \infty$, we have $a(x)/x \to 0$. On the other hand, $c_n = a(d_n)$ and $d_n \to \infty$ since $F \in \mathcal{S}$. Therefore, for any $x > 0$,

$$d_{np} + x\,c_{np} = d_{np}\left(1 + x\,\frac{a(d_{np})}{d_{np}}\right) \sim d_{np}.$$

Hence (4.8) holds if $d_{np} \geq (1 + \delta)\gamma_n$ for any small $\delta > 0$ and large n.

4.2.1. *The extreme values of i.i.d. random walks.* Relation (4.10) and a continuous mapping argument imply the following analog of Corollary 4.1. We use the same notation as in Section 4.1.1. One can follow the lines of the proof of Corollary 4.1.

Corollary 4.2. *Assume the conditions of Theorem 4.2. Then the following relation holds for $k \geq 1$:*

$$c_{np}^{-1}(S_{n,(1)} - \mathbb{E}[S_n] - d_{np}, \ldots, S_{n,(k)} - \mathbb{E}[S_n] - d_{np}) \xrightarrow{d} (-\log \Gamma_1, \ldots, -\log \Gamma_k)$$

as $n \to \infty$.

4.2.2. *Examples.* Theorem 4.2 applies to $F \in \mathrm{LN}(\gamma)$, $\gamma > 1$, and $F \in \mathrm{WE}(\tau)$, $0 < \tau < 1$; see the discussion in Section 3.2. However, the calculation of the constants (c_n) and (d_n) is rather complicated for these classes of subexponential distributions. For illustration of the theory we restrict ourselves to two parametric classes of distributions where these constants are known.

Example 4.6. We assume that X has a standard lognormal distribution. From (2.5), Table 1, and Remark 4.2 we conclude that we need to verify the condition $\exp\left(\sqrt{2\log(np)}\right) \geq h_n\sqrt{n}\log n$

for a sequence (h_n) increasing to infinity arbitrarily slowly. Calculation shows that it suffices to choose $p_n \to \infty$ such that $p > \exp((\log n)^2)$.

Example 4.7. We assume that X has a Weibull distribution with tail $\overline{F}(x) = \exp(-x^\tau)$ for some $\tau \in (0, 1)$. From (2.6) we conclude that $d_{np} \sim (\log np)^{1/\tau}$. In view of Remark 4.2 and Table 1 it suffices to verify that $(\log np)^{1/\tau} \ge h_n \, n^{1/(2-2\tau)}$ for a sequence $h_n \to \infty$ arbitrarily slowly. It holds if $p > n^{-1} \, \exp((h_n n^{1/(2-2\tau)})^\tau)$.

4.2.3. *The extremes of the blocks of a random walk.* We appeal to the notation in Section 4.1.3. We are in the setting of Theorem 4.2 if we replace p_n with k_n and n with r_n. We are interested in the following result for the point process of the block sums of S_n with length r_n (see (4.10)):

$$N_{k_n} = \sum_{i=1}^{k_n} \varepsilon_{c_n^{-1}\,(S_{r_n i} - S_{r_n(i-1)} - \mathbb{E}[S_{r_n}] - d_n)} \xrightarrow{d} N = \sum_{i=1}^{\infty} \varepsilon_{-\log \Gamma_i}.$$

We need to verify condition (4.8) which turns into $d_n + c_n x > \gamma_{r_n}$. In view of Remark 4.2 it suffices to prove that $d_n > h_n \gamma_{r_n}$ for a sequence $h_n \to \infty$ arbitrarily slowly; see Table 1 for some γ_n-values.

We start with a standard lognormal distribution; see (2.5) for the corresponding (c_n) and (d_n). In particular, we need to verify

$$d_n = \exp\left(\sqrt{2 \log n} - \frac{\log \log n + \log 4\pi}{2(2 \log n)^{1/2}}\right) \ge h_n \sqrt{r_n} \log r_n.$$

A sufficient condition is $\exp(2\sqrt{2 \log n}) > \widetilde{h}_n r_n$ for a sequence $\widetilde{h}_n \to \infty$ arbitrarily slowly. We observe that the left-hand expression is a slowly varying function.

Next we consider a standard Weibull distribution for $\tau \in (0, 1)$. The constants (c_n) and (d_n) are given in (2.6). In particular, we need to verify

$$d_n \sim (\log n)^{1/\tau} > h_n r_n^{1/(2-2\tau)}.$$

This holds if $(\log n)^{2(1-\tau)/\tau} h_n^{-2(1-\tau)} > r_n$. Again, this is a strong restriction on the growth of (r_n) and is in contrast to the regularly varying case where polynomial growth of (r_n) is possible; see Section 4.3.2.

4.3. Fréchet convergence via the subexponential approximations to large deviation probabilities for large x

In this section we assume that X is regularly varying with index $\alpha > 0$ in the sense of (2.1). Throughout we choose a normalizing sequence (a_n) such that $n \, \mathbb{P}(|X| > a_n) \to 1$ as $n \to \infty$. The following result is an analog of Theorems 4.1 and 4.2.

Theorem 4.3. *Assume that X is regularly varying with index $\alpha > 0$ and $\mathbb{E}[X] = 0$ if the expectation is finite. Choose a sequence (d_n) such that*

$$d_n = \begin{cases} 0 & \alpha \in (0, 1) \cup (1, \infty), \\ n \, \mathbb{E}[X \, \mathbf{1}(|X| \le a_n)] & \alpha = 1. \end{cases}$$

We assume that $p_n \to \infty$ is an integer sequence which satisfies the additional conditions

$$\begin{cases} a_{np} \ge \sqrt{(\alpha - 2 + \delta)n \log n} \text{ for some small } \delta > 0 & \text{if } \alpha > 2, \\ \lim_{n \to \infty} \sup_{x > a_{np}} p^\delta \dfrac{n}{x^2} \, \mathbb{E}[X^2 \, \mathbf{1}(|X| \le x)] = 0 \text{ for some small } \delta > 0 & \text{if } \alpha = 2. \end{cases} \tag{4.12}$$

Then we obtain the limit relation

$$p\, \mathbb{P}(\pm a_{np}^{-1}(S_n - d_n) > x) \to p_\pm x^{-\alpha}, \quad x > 0, \quad n \to \infty. \tag{4.13}$$

Moreover, (4.13) *is equivalent to*

$$N_p = \sum_{i=1}^{p} \varepsilon_{a_{np}^{-1}(S_{ni} - d_n)} \xrightarrow{d} N = \sum_{i=1}^{\infty} \varepsilon_{q_i \Gamma_i^{-1/\alpha}}, \tag{4.14}$$

where (Γ_i) *is defined in Theorem 4.1 and* (q_i) *is an i.i.d. sequence of Bernoulli variables with distribution* $\mathbb{P}(q_i = \pm 1) = p_\pm$ *independent of* (Γ_i).

Proof. We start by verifying (4.13). Assume $\alpha < 2$. Then, for any sequence $p_n \to \infty$, $a_{np}/a_n \to \infty$. Therefore Theorem 3.3 and the definition of (a_{np}) yield

$$p\, \mathbb{P}(\pm a_{np}^{-1}(S_n - d_n) > x) \sim p\, n\, \mathbb{P}(\pm X > a_{np}\, x) \sim p_\pm x^{-\alpha}, \quad n \to \infty.$$

If $\alpha > 2$, the same result holds in view of Theorem 3.2 since we assume condition (4.12). If $\alpha = 2$, we can again apply Theorem 3.3 with $\gamma_n = a_{np}$ and use (4.12).

We notice that the limit point process N is PRM(μ_α) with intensity

$$\mu_\alpha(\mathrm{d}x) = |x|^{-\alpha-1}(p_+ \mathbf{1}(x > 0) + p_- \mathbf{1}(x < 0))\, \mathrm{d}x. \tag{4.15}$$

An appeal to Proposition 2.1 shows that (4.13) and (4.14) are equivalent. $\qquad\square$

Remark 4.3. Assume $\alpha > 2$. Since $a_{np} = (np)^{1/\alpha}\ell(np)$ for a slowly varying function ℓ and $\ell(x) \geq x^{-\gamma/\alpha}$ for any small $\gamma > 0$ and sufficiently large x, (4.12) holds if $p \geq n^{(\alpha/2)-1+\gamma'}$ for any choice of $\gamma' > 0$. Assume $\alpha = 2$ and var$(X) < \infty$. Then $a_{np} \sim c\, \sqrt{np}$ and (4.12) is satisfied for any sequence $p_n \to \infty$ and $\delta < 1$. If var$(X) = \infty$, then $a_{np} = (np)^{1/2}\ell(np)$ for a slowly varying function ℓ and $\mathbb{E}[X^2(|X| \leq x)]$ is an increasing slowly varying function. Using Karamata bounds for slowly varying functions, we conclude that (4.12) holds if $p/n^\gamma \to \infty$ for any small $\gamma > 0$.

4.3.1. *The extreme values of i.i.d. random walks.* For simplicity, we assume $d_n = 0$. Let N_p^+ be the restriction of N_p to the state space $(0, \infty)$, and $S_{n,(1)}^+$ the maximum of $(S_{n1})_+, \ldots, (S_{np})_+$. We also write $\xi = \min\{i \geq 1 : q_i = 1\}$ and assume that ξ is independent of (Γ_i). Then (4.14) and the continuous mapping theorem imply that

$$\mathbb{P}(N_p^+(x, \infty) = 0) = \mathbb{P}(a_{np}^{-1} S_{n,(1)}^+ \leq x) \xrightarrow{d} \mathbb{P}(\Gamma_\xi^{-1/\alpha} \leq x) = \Phi_\alpha^{p+}(x). \tag{4.16}$$

Moreover, we have joint convergence of minima and maxima.

Corollary 4.3. *Assume the conditions of Theorem 4.3 and* $d_n = 0$. *Then*

$$\lim_{n \to \infty} \mathbb{P}\left(0 < a_{np}^{-1} \max_{i=1,\ldots,p} S_{ni} \leq x,\ -y < a_{np}^{-1} \min_{i=1,\ldots,p} S_{ni}\right) = \Phi_\alpha^{p+}(x)\Phi_\alpha^{p-}(y), \quad x, y > 0.$$

Proof. We have

$$\mathbb{P}\Big(a_{np}^{-1} \max_{i=1,\dots,p} S_{ni} \le x, \ -y < a_{np}^{-1} \min_{i=1,\dots,p} S_{ni}\Big) = \mathbb{P}(N_p((x,\infty) \cup (-\infty, -y]) = 0)$$

$$\to \mathbb{P}(N((x,\infty) \cup (-\infty, -y]) = 0)$$

$$= \exp\left(-(p_+ x^{-\alpha} + p_- y^{-\alpha})\right)$$

$$= \Phi_\alpha^{p_+}(x)\Phi_\alpha^{p_-}(y), \quad n \to \infty. \qquad \square$$

4.3.2. *The extremes of the blocks of a random walk.* We appeal to the notation of Section 4.1.3 and apply Theorem 4.3 in the case when n is replaced by some integer-sequence $r_n \to \infty$ such that $k_n = [n/r_n] \to \infty$ and p_n is replaced by k_n. We also assume for simplicity that $d_n = 0$. Observing that a_{np} turns into $a_{r_n k_n} \sim a_n$, (4.14) turns into

$$N_{k_n} = \sum_{i=1}^{k_n} \varepsilon_{a_n^{-1}(S_{r_n i} - S_{r_n(i-1)})} \xrightarrow{d} N = \sum_{i=1}^{\infty} \varepsilon_{q_i \Gamma_i^{-1/\alpha}}, \quad n \to \infty.$$

For simplicity, we assume $\alpha \ne 2$. If $\alpha < 2$, no further restrictions on (r_n) are required. If $\alpha > 2$, we have the additional growth condition $a_n > \sqrt{(\alpha - 2 + \delta)r_n \ \log r_n}$ for sufficiently large n. Since $a_n = n^{1/\alpha}\ell(n)$ for some slowly varying function ℓ, this amounts to showing that $n^{2/\alpha}\ell^2(n)/(\alpha - 2 + \delta) > r_n \log r_n$. Since any slowly varying function satisfies $\ell(n) \ge n^{-\varepsilon}$ for any $\varepsilon > 0$ and $n \ge n_0(\varepsilon)$, we get the following sufficient condition on the growth of (r_n): for any sufficiently small $\varepsilon > 0$, $n^{2/\alpha - \varepsilon} > r_n$. This condition ensures that (r_n) is significantly smaller than n, and the larger α the more stringent this condition becomes.

An appeal to (4.16) yields in particular

$$\mathbb{P}\Big(a_n^{-1} \max_{i=1,\dots,k_n} (S_{r_n i} - S_{r_n(i-1)})_+ \le x\Big) \xrightarrow{d} \mathbb{P}(\Gamma_\xi^{-1/\alpha} \le x) = \Phi_\alpha^{p_+}(x),$$

$$\mathbb{P}\Big(a_n^{-1} \max_{i=1,\dots,k_n} |S_{r_n i} - S_{r_n(i-1)}| \le x\Big) \xrightarrow{d} \mathbb{P}(\Gamma_1^{-1/\alpha} \le x) = \Phi_\alpha(x), \quad n \to \infty.$$

4.3.3. *Extension to a stationary regularly varying sequence.* In view of classical theory (e.g. Feller [9]), X is regularly varying with index $\alpha \in (0, 2)$ if and only if $a_n^{-1}(S_n - d_n) \xrightarrow{d} \xi_\alpha$ for an α-stable random variable ξ_α where one can choose (a_n) such that $n \mathbb{P}(|X| > a_n) \to 1$ and (d_n) as in (3.5). For the sake of argument we also assume $d_n = 0$; this is a restriction only in the case $\alpha = 1$.

If (r_n) is any integer sequence such that $r_n \to \infty$ and $k_n = [n/r_n] \to 0$, then

$$a_n^{-1} S_n = a_n^{-1} \sum_{i=1}^{k_n} (S_{r_n i} - S_{r_n(i-1)}) + o_{\mathbb{P}}(1) \xrightarrow{d} \xi_\alpha. \tag{4.17}$$

Moreover, since $a_n/a_{r_n} \to \infty$, Theorem 3.3 yields

$$\frac{\mathbb{P}(\pm a_n^{-1} S_{r_n} > x)}{r_n \mathbb{P}(|X| > a_n)} \sim \frac{\mathbb{P}(\pm X > x a_n)}{\mathbb{P}(|X| > a_n)} \to p_\pm x^{-\alpha}, \quad x > 0. \tag{4.18}$$

Classical limit theory for triangular arrays of the row-wise i.i.d. random variables $(S_{r_n i} - S_{r_n(i-1)})_{i=1,\dots,k_n}$ (e.g. Petrov [26, Theorem 8, Chapter IV]) yields that (4.17) holds if and only if

$$k_n \mathbb{P}(a_n^{-1} S_{r_n} \in \cdot) \xrightarrow{v} \mu_\alpha(\cdot), \tag{4.19}$$

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} k_n \operatorname{var}(a_n^{-1} S_{r_n} \mathbf{1}(|S_{r_n}| \le \delta a_n)) = 0, \tag{4.20}$$

where μ_α is defined in (4.15). We notice that (4.19) is equivalent to (4.18).

An alternative way of proving limit theory for the sum process (S_n) with an α-stable limit ξ_α would be to *assume* the relations (4.19) and (4.20). This would be rather indirect and complicated in the case of i.i.d. (X_i). However, this approach has some merits in the case when (X_i) is a strictly stationary sequence with a regularly varying dependence structure, that is, its finite-dimensional distributions satisfy a multivariate regular variation condition (see Davis and Hsing [5] or Basrak and Segers [1]), and we have a weak dependence assumption of the type

$$\mathbb{E}[\exp(a_n^{-1} i t S_n)] - (\mathbb{E}[\exp(a_n^{-1} i t S_{r_n})])^{k_n} \to 0, \quad t \in \mathbb{R}, \quad n \to \infty. \tag{4.21}$$

Then $a_n^{-1} S_n \xrightarrow{d} \xi_\alpha$ if and only if $a_n^{-1} \sum_{i=1}^{k_n} S_{ni} \xrightarrow{d} \xi_\alpha$, where $(S_{ni})_{i=1,\dots,k_n}$ is an i.i.d. sequence with the same distribution as S_{r_n}. Condition (4.21) is satisfied under mild conditions on (X_i), in particular under standard mixing conditions such as α-mixing. Thus we have to prove conditions (4.19) and (4.20). In the dependent case the limit measure μ_α has to be modified. The following analog of (4.18) holds: there exists a positive number θ_X such that

$$\frac{\mathbb{P}(\pm a_n^{-1} S_{r_n} > x)}{r_n \mathbb{P}(|X| > a_n)} \sim \theta_X \frac{\mathbb{P}(\pm X > x a_n)}{\mathbb{P}(|X| > a_n)} \to \theta_X p_\pm x^{-\alpha}, \quad x > 0.$$

The quantity θ_X has an explicit structure in terms of the so-called tail chain of the regularly varying sequence (X_i). It can be interpreted as a *cluster index* in the context of the partial sum operation acting on (X_i). For details we refer to Mikosch and Wintenberger [20] and the references therein.

4.3.4. *Extension to the multivariate regularly varying case.* Consider a sequence $(\mathbf{X}_i)$ of i.i.d. $\mathbb{R}^d$-valued random vectors with generic element $\mathbf{X}$, and define

$$\mathbf{S}_0 = \mathbf{0}, \quad \mathbf{S}_n = \mathbf{X}_1 + \cdots + \mathbf{X}_n, \quad n \ge 1.$$

We say that $\mathbf{X}$ is regularly varying with index $\alpha > 0$ and a Radon measure μ on $\mathbb{R}_0^d = \mathbb{R}^d \backslash \{\mathbf{0}\}$, and we write $\mathbf{X} \in \mathrm{RV}(\alpha, \mu)$ if the following vague convergence relation is satisfied on $\mathbb{R}_0^d$:

$$\frac{\mathbb{P}(x^{-1} \mathbf{X} \in \cdot)}{\mathbb{P}(|\mathbf{X}| > x)} \xrightarrow{v} \mu(\cdot), \quad x \to \infty, \tag{4.22}$$

and μ has the homogeneity property $\mu(t \cdot) = t^{-\alpha} \mu(\cdot)$, $t > 0$. We will also use the sequential version of regular variation: for a sequence (a_n) such that $n \mathbb{P}(|\mathbf{X}| > a_n) \to 1$, (4.22) is equivalent to

$$n \, \mathbb{P}(a_n^{-1} \mathbf{X} \in \cdot) \xrightarrow{v} \mu(\cdot), \quad n \to \infty.$$

For more reading on multivariate regular variation, we refer to Resnick [28, 29].

Hult, Lindskog, Mikosch, and Samorodnitsky [14] extended Nagaev's Theorem 3.2 to the multivariate case.

Theorem 4.4. (A multivariate Nagaev-type large deviation result.) *Consider an i.i.d. $\mathbb{R}^d$-valued sequence $(\mathbf{X}_i)$ with generic element $\mathbf{X}$. Assume the following conditions.*

(1) $\mathbf{X} \in \mathrm{RV}(\alpha, \mu)$.

(2) *The sequence of positive numbers (x_n) satisfies*

$$x_n^{-1}\mathbf{S}_n \overset{\mathbb{P}}{\to} \mathbf{0} \quad \text{as } n \to \infty, \tag{4.23}$$

and, in addition,

$$\begin{cases} \dfrac{x_n^2}{n\mathbb{E}[|\mathbf{X}|^2\mathbf{1}(|\mathbf{X}| \le x_n)]\log x_n} \to \infty & \alpha = 2 \text{ and } \mathbb{E}[|\mathbf{X}|^2] = \infty, \\ \dfrac{x_n^2}{n\log n} \to \infty & \alpha > 2 \text{ or } [\alpha = 2 \text{ and } \mathbb{E}[|\mathbf{X}|^2] < \infty]. \end{cases} \tag{4.24}$$

Then

$$\frac{\mathbb{P}(x_n^{-1}\mathbf{S}_n \in \cdot)}{n\mathbb{P}(|\mathbf{X}| > x_n)} \overset{v}{\to} \mu(\cdot), \quad n \to \infty.$$

Remark 4.4. Condition (4.23) requires that $n\mathbb{E}[\mathbf{X}]/a_{np} \to \mathbf{0}$ for $\alpha > 1$. It is always satisfied if $\mathbb{E}[\mathbf{X}] = \mathbf{0}$. Now assume that the latter condition is satisfied if the expectation of $\mathbf{X}$ is finite. If $\alpha \in (0, 2)$, we can choose any (p_n) such that $p_n \to \infty$. If $\alpha \ge 2$ and $(np)^{1/\alpha}/n^{0.5+\gamma/\alpha} \to \infty$, equivalently, $p/n^{\alpha/2-1+\gamma} \to \infty$ holds for any small $\gamma > 0$, then (4.24) is satisfied.

The following result extends Theorem 4.3 to the multivariate case.

Theorem 4.5. *Assume that $\mathbf{X}$ satisfies the conditions of Theorem 4.4. Consider an integer sequence $p = p_n \to \infty$ and, in addition for $\alpha \ge 2$, that $x_n = a_{np}$ satisfies (4.24). Then the following limit relation holds*:

$$N_p = \sum_{i=1}^p \varepsilon_{a_{np}^{-1}\mathbf{S}_{ni}} \overset{d}{\to} N,$$

where $(\mathbf{S}_{ni})$ are i.i.d. copies of $\mathbf{S}_n$ and N is $\mathrm{PRM}(\mu)$ on $\mathbb{R}_{\mathbf{0}}^d$.

Proof. In view of Proposition 2.1 it suffices to show that

$$p\,\mathbb{P}(a_{np}^{-1}\mathbf{S}_n \in \cdot) \overset{v}{\to} \mu(\cdot).$$

Assume $\alpha < 2$. Then, for any sequence $p_n \to \infty$, $a_{np}/a_n \to \infty$. Therefore Theorem 4.4 and the definition of (a_{np}) imply that, for any μ-continuity set $A \subset \mathbb{R}_{\mathbf{0}}^d$,

$$p\,\mathbb{P}(a_{np}^{-1}\mathbf{S}_n \in A) \sim p\,n\,\mathbb{P}(|\mathbf{X}| > a_{np})\,\mu(A) \to \mu(A), \quad n \to \infty.$$

If $\alpha \ge 2$, the same result holds by virtue of Theorem 4.4 and the additional condition (4.24). $\square$

Example 4.8. Write

$$\mathbf{S}_{ni} = (S_{ni}^{(1)}, \dots, S_{ni}^{(d)})^\top,$$
$$\mathbf{M}_n = \left(\max_{i=1,\dots,p} S_{ni}^{(1)}, \dots, \max_{i=1,\dots,p} S_{ni}^{(d)}\right)^\top = (M_n^{(1)}, \dots, M_n^{(d)})^\top.$$

For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ with non-negative components, we write $\mathbf{x} \le \mathbf{y}$ for the componentwise ordering, $[\mathbf{0}, \mathbf{x}] = \{\mathbf{y} : \mathbf{0} \le \mathbf{y} \le \mathbf{x}\}$ and $[\mathbf{0}, \mathbf{x}]^c = \mathbb{R}_+^d \setminus [\mathbf{0}, \mathbf{x}]$. We have by Theorem 4.5,

$$\mathbb{P}(\mathbf{0} \le a_{np}^{-1}\mathbf{M}_n \le \mathbf{x}) = \mathbb{P}(N_p([\mathbf{0}, \mathbf{x}]^c) = 0)$$
$$\to \mathbb{P}(N([\mathbf{0}, \mathbf{x}]^c) = 0)$$

$$= \exp\left(-\mu([\mathbf{0}, \mathbf{x}]^c)\right)$$
$$=: H(\mathbf{x}), \quad n \to \infty,$$

for the continuity points of the function $-\log H(\mathbf{x}) = \mu([\mathbf{0}, \mathbf{x}]^c)$. If $\mu(\mathbb{R}_+^d \setminus \{\mathbf{0}\})$ is not zero, H defines a distribution on $\mathbb{R}_+^d$ with the property $-\log H(t\mathbf{x}) = t^{-\alpha}(-\log H(\mathbf{x}))$, $t > 0$. The non-degenerate components of H are in the type of the Fréchet distribution; H is referred to as a multivariate Fréchet distribution with exponent measure μ.

4.3.5. *An extension to i.i.d. random sums.* In this section we consider an alternative random sum process:

$$S(t) = \sum_{i=1}^{\nu(t)} X_i, \quad t \geq 0,$$

where $(\nu(t))_{t \geq 0}$ is a process of integer-valued non-negative random variables independent of the i.i.d. sequence (X_i) with generic element X and finite expectation. Throughout we assume that $\lambda(t) = \mathbb{E}[\nu(t)]$, $t \geq 0$, is finite but $\lim_{t \to \infty} \lambda(t) = \infty$. We also define

$$m(t) = \mathbb{E}[S(t)] = \mathbb{E}[X]\,\lambda(t).$$

In addition, we assume some technical conditions on the process ν.

N1 $\nu(t)/\lambda(t) \xrightarrow{\mathbb{P}} 1$, $t \to \infty$.

N2 There exist ϵ, $\delta > 0$ such that

$$\lim_{t \to \infty} \sum_{k > (1+\delta)\lambda(t)} \mathbb{P}(\nu(t) > k)\,(1 + \epsilon)^k = 0.$$

These conditions are satisfied for a wide variety of processes ν, including the homogeneous Poisson process on $(0, \infty)$. Klüppelberg and Mikosch [16] proved the following large deviation result for the random sums $S(t)$ (allowing for the more general condition of extended regular variation).

Theorem 4.6. *Assume that ν satisfies* **N1** *and* **N2** *and is independent of the i.i.d. non-negative sequence (X_i) which is regularly varying with index $\alpha > 1$. Then, for any $\gamma > 0$,*

$$\sup_{x \geq \gamma\lambda(t)} \left| \frac{\mathbb{P}(S(t) - m(t) > x)}{\lambda(t)\mathbb{P}(X > x)} - 1 \right|, \quad t \to \infty.$$

The same method of proof as in the previous sections in combination with the large deviation result of Theorem 4.6 yields the following statement. As usual, we assume that $(a(t))$ is a function such that $t\,\mathbb{P}(X > a(t)) \to 1$ as $t \to \infty$.

Corollary 4.4. *Assume the condition of Theorem 4.6. Let $(p(t))$ be an integer-valued function such that that $p(t) \to \infty$ as $t \to \infty$ and a growth condition is satisfied for every fixed $\gamma > 0$ and sufficiently large $t \geq t_0$:*

$$a(\lambda(t)p(t)) \geq \gamma\,\lambda(t). \tag{4.25}$$

Then the following limit relation holds for i.i.d. copies S_i of the random sum process S:

$$N_{p(t)} = \sum_{i=1}^{p(t)} \varepsilon_{(S_i(t) - m(t))/(a(\lambda(t)p(t)))} \xrightarrow{d} N = \sum_{i=1}^{\infty} \varepsilon_{\Gamma_i^{-1/\alpha}}, \quad t \to \infty,$$

where (Γ_i) *is defined in Theorem* 4.1.

Proof. In view of Proposition 2.1 the result is proved if we can show that as $t \to \infty$,

$$p(t)\, \mathbb{P}((a(\lambda(t)p(t)))^{-1}(S(t) - m(t)) > x) \sim \lambda(t)\, p(t)\, \mathbb{P}(X > a(\lambda(t)p(t))\, x) \to x^{-\alpha},$$

$$p(t)\, \mathbb{P}((a(\lambda(t)p(t)))^{-1}(S(t) - m(t)) < -x) \to 0, \quad x > 0.$$

But this follows by an application of Theorem 4.6 in combination with (4.25) and the regular variation of X. □

Remark 4.5. Since $a(\lambda(t)p(t)) = (\lambda(t)p(t))^{1/\alpha}\ell(\lambda(t)p(t))$ for a slowly varying function ℓ and $\ell(x) \geq x^{-\epsilon/\alpha}$ for any small $\epsilon > 0$ and sufficiently large x, (4.25) holds if $p(t) \geq (\lambda(t))^{\alpha-1+\epsilon'}$ for any choice of $\epsilon' > 0$.

4.4. An extension: the index of the point process is random

Let $(P_n)_{n \geq 0}$ be a sequence of positive integer-valued random variables. We assume that there exists a sequence of positive numbers (p_n) such that $p_n \to \infty$ and

$$\frac{P_n}{p_n} \xrightarrow{\mathbb{P}} 1, \quad n \to \infty. \tag{4.26}$$

This condition is satisfied for wide classes of integer-valued sequences (P_n), including the renewal counting processes and (inhomogeneous) Poisson processes when calculated at the positive integers. In particular, for renewal processes $p_n \sim c\, n$ provided the inter-arrival times have finite expectation.

We have the following analog of Proposition 2.1.

Proposition 4.1. *Let* $(X_{ni})_{n=1,2,\dots;i=1,2,\dots}$ *be a triangular array of i.i.d. random variables assuming values in some state space* $E \subset \mathbb{R}^d$ *equipped with the Borel* σ-*field* $\mathcal{E}$. *Let* μ *be a Radon measure on* $\mathcal{E}$. *If the relation*

$$p_n\, \mathbb{P}(X_{n1} \in \cdot) \xrightarrow{v} \mu(\cdot), \quad n \to \infty, \tag{4.27}$$

holds on E, then

$$\widetilde{N}_p = \sum_{i=1}^{P_n} \varepsilon_{X_{ni}} \xrightarrow{d} N, \quad n \to \infty,$$

where N is PRM(μ) *on E.*

Proof. We prove the result by showing convergence of the Laplace functionals. The arguments of a Laplace functional are elements of

$$C_K^+(E) = \{g\colon E \to \mathbb{R}_+ : g \text{ continuous with compact support}\}.$$

For $f \in C_K^+$ we have by independence of the (X_{ni}),

$$\mathbb{E}\left[\exp\left(-\int_E f\, d\widetilde{N}_p\right)\right] = \mathbb{E}\left[\exp\left(-\sum_{j=1}^{P_n} f(X_{nj})\right)\right] = \mathbb{E}[(\mathbb{E}[\exp(-f(X_{n1}))])^{P_n}].$$

In view of (4.26) there is a real sequence $\epsilon_n \downarrow 0$ such that

$$\lim_{n\to\infty} \mathbb{P}(|P_n/p_n - 1| > \epsilon_n) = \mathbb{P}(A_n^c) = 0. \tag{4.28}$$

Then

$$\mathbb{E}[(\mathbb{E}[\exp(-f(X_{n1})])^{P_n}]$$
$$= \mathbb{E}[(\mathbb{E}[\exp(-f(X_{n1})])^{P_n}(\mathbf{1}(A_n^c) + \mathbf{1}(A_n))]$$
$$= I_1 + I_2.$$

By (4.28) we have $I_1 \le \mathbb{P}(A_n^c) \to 0$ as $n \to \infty$, while

$$\mathbb{E}[(\mathbb{E}[\exp(-f(X_{n1})])^{(1+\epsilon_n)p_n}\mathbf{1}(A_n)] \le I_2 \le \mathbb{E}[(\mathbb{E}[\exp(-f(X_{n1})])^{(1-\epsilon_n)p_n}\mathbf{1}(A_n)]. \qquad (4.29)$$

In view of Proposition 2.1 and (4.27)

$$(\mathbb{E}[\exp(-f(X_{n1})])^{(1\pm\epsilon_n)p_n} \to \mathbb{E}\left(\exp\left(-\int_E (1 - e^{-f(\mathbf{x})})\,\mu(d\mathbf{x})\right)\right).$$

The right-hand side is the Laplace functional of a PRM(μ). Now an application of dominated convergence to I_2 in (4.29) yields the desired convergence result. $\qquad \square$

An immediate consequence of this result is that all point process convergences in Section 4 remain valid if the point processes N_p are replaced by their corresponding analogs $\widetilde{N}_p$ with a random index sequence (P_n) independent of (S_{ni}) and satisfying (4.26). Moreover, the growth rates for $p_n \to \infty$ remain the same.

4.5. Extension to the tail empirical process

We assume that (S_{ni}) are i.i.d. copies of a real-valued random walk (S_n). Instead of the point processes considered in the previous sections, one can also study the tail empirical process

$$N_p = \frac{1}{k}\sum_{i=1}^{p} \varepsilon_{c_{[p/k]}^{-1}(S_{ni}/\sqrt{n}-d_{[p/k]})},$$

where $k = k_n \to \infty$, $p = p_n \to \infty$ and $p_n/k_n \to \infty$, and (c_n) and (d_n) are suitable normalizing and centering constants. To illustrate the theory we consider two examples.

Example 4.9. Assume the conditions and notation of Theorem 4.1. In this case, choose $c_n = 1/d_n$. Then

$$\mathbb{E}[N_p(x, \infty)] = \frac{p}{k}\,\mathbb{P}(S_n/\sqrt{n} > d_{[p/k]} + x/d_{[p/k]}) \to e^{-x},$$
$$\mathrm{var}(N_p(x, \infty)) \le \frac{p}{k^2}\,\mathbb{P}(S_n/\sqrt{n} > d_{[p/k]} + x/d_{[p/k]}) \to 0, \quad x \in \mathbb{R}, \quad n \to \infty,$$

provided $p/k < \exp(\gamma_n^2/2)$. It is not difficult to see that

$$N_p \overset{\mathbb{P}}{\to} -\log \Lambda.$$

Similarly, assume the conditions and the notation of Theorem 4.3 and consider

$$N_p = \frac{1}{k}\sum_{i=1}^{p} \varepsilon_{a_{[np/k]}^{-1}(S_{ni}-d_n)}.$$

Then, for $x > 0$ as $n \to \infty$,

$$\mathbb{E}[N_p(x, \infty)] = \frac{p}{k} \mathbb{P}(a_{[np/k]}^{-1}(S_n - d_n) > x) \sim \frac{np}{k} \mathbb{P}(X > a_{[np/k]} x) \to p_+ x^{-\alpha} = \mu_\alpha(x, \infty),$$

$$\text{var}(N_p(x, \infty)) \to 0,$$

$$\mathbb{E}[N_p(-\infty, -x]] = \frac{p}{k} \mathbb{P}(a_{[np/k]}^{-1}(S_n - d_n) \le -x) \to p_- x^{-\alpha} = \mu_\alpha(-\infty, -x],$$

$$\text{var}(N_p(-\infty, -x]) \to 0,$$

provided the modified sequence $p_n/k_n \to \infty$ satisfies the conditions imposed on (p_n) in Theorem 4.3. We notice that the values of μ_α on $(-\infty, -x]$ and (x, ∞) determine a Radon measure on $\mathbb{R}\backslash\{0\}$. From these relations we conclude that $N_p \xrightarrow{\mathbb{P}} \mu_\alpha$. Then, following the lines of Resnick and Stărică [30, Proposition 2.3], one can for example prove consistency of the Hill estimator based on the sample $(S_{ni})_{i=1,\dots,p}$. Assuming for simplicity $d_n = 0$, $p_+ > 0$, we write $S_{n,(1)} \ge \cdots \ge S_{n,(k)}$ for the k largest values. Then

$$\frac{1}{k} \sum_{i=1}^{k} \log \frac{S_{n,(i)}}{S_{n,(k)}} \xrightarrow{\mathbb{P}} \frac{1}{\alpha}.$$

4.6. Some related results

The largest values of sequences of i.i.d. normalized and centered partial sum processes play a role in the context of random matrix theory, which is also the main motivation for the present work. Consider a double array (X_{it}) of i.i.d. regularly varying random variables with index $\alpha \in (0, 4)$ (see (2.1)) and generic element X, and also assume that $\mathbb{E}[X] = 0$ if this expectation is finite. Consider the data matrix

$$\mathbf{X} := \mathbf{X}_n = (X_{it})_{i=1,\dots,p;t=1,\dots,n}$$

and the corresponding sample covariance matrix $\mathbf{XX}^\top = (S_{ij})$. Heiny and Mikosch [12] proved that

$$a_{np}^{-2}\|\mathbf{XX}^\top - \text{diag}(\mathbf{XX}^\top)\|_2 \xrightarrow{\mathbb{P}} 0, \quad n \to \infty,$$

where $\|A\|_2$ denotes the spectral norm of a $p \times p$ symmetric matrix A, $\text{diag}(A)$ consists of the diagonal of A, (a_k) is any sequence satisfying $k\mathbb{P}(|X| > a_k) \to 1$ as $k \to \infty$, and $p_n = n^\beta \ell(n)$ for some $\beta \in (0, 1]$ and a slowly varying function ℓ. Write $\lambda_{(1)}(A) \ge \cdots \ge \lambda_{(p)}(A)$ for the ordered eigenvalues of A. According to Weyl's inequality (see Bhatia [2]), the eigenvalues of $\mathbf{XX}^\top$ satisfy the relation

$$a_{np}^{-2} \sup_{i=1,\dots,p} |\lambda_{(i)}(\mathbf{XX}^\top) - \lambda_{(i)}(\text{diag}(\mathbf{XX}^\top))| \le a_{np}^{-2}\|\mathbf{XX}^\top - \text{diag}(\mathbf{XX}^\top)\|_2 \xrightarrow{\mathbb{P}} 0. \quad (4.30)$$

But of course, $\lambda_{(i)}(\text{diag}(\mathbf{XX}^\top))$ are the ordered values of the i.i.d. partial sums $S_{ii} = \sum_{t=1}^{n} X_{it}^2$, $i = 1, \dots, p$. In view of (4.30) the asymptotic theory for the largest eigenvalues of the normalized sample covariance matrix $a_{np}^{-2}\mathbf{XX}^\top$ (which also needs centering for $\alpha \in (2, 4)$) are determined through the Fréchet convergence of the processes with points $(a_{np}^{-2}S_{ii})_{i=1,\dots,p}$. Moreover, (4.30) implies the Fréchet convergence of the point processes of the normalized and centered eigenvalues of the sample covariance matrix.

The large deviation approach also works for proving limit theory for the point process of the off-diagonal elements of $\mathbf{X}\mathbf{X}^\top$ provided X has sufficiently high moments. Heiny, Mikosch, and Yslas [13] prove Gumbel convergence for the point process of the off-diagonal elements $(S_{ij})_{1 \le i < j \le p}$. The situation is more complicated because the points S_{ij} are typically dependent. Multivariate extensions of the normal large deviation approximation $0.5p^2\mathbb{P}(d_{p^2/2}(S_{12} - d_{p^2/2}) > x) \to \exp(-x)$ show that the point process of the standardized (S_{ij}) has the same limit Poisson process as if the S_{ij} were independent. Moreover, Heiny *et al.* [13] show that the point process of the diagonal elements (S_{ii}) (under suitable conditions on the rate of $p_n \to \infty$ and under $\mathbb{E}[|X|^s] < \infty$ for $s > 4$) converges to $\mathrm{PRM}(-\log \Lambda)$. This result indicates that the off-diagonal and diagonal entries of $\mathbf{X}\mathbf{X}^\top$ exhibit very similar extremal behavior. This is in stark contrast to the aforementioned results in [12], where the diagonal entries have Fréchet extremal behavior.

Related results can also be found in Gantert and Höfelsauer [11], who consider real-valued branching random walks and prove a large deviation principle for the position of the rightmost particle; see [11, Theorem 3.2]. The position of the rightmost particle is the maximum of a collection of a random number of dependent random walks. In this context, the authors also prove a related large deviation result under the assumption that the random walks considered are i.i.d. They show that the maximum of these i.i.d. random walks stochastically dominates the maximum of the branching random walks; see [11, Theorem 3.1 and Lemma 5.2]. An early comparison between maxima of branching and i.i.d. random walks was provided by Durrett [7].

Acknowledgements

References

[1] BASRAK, B. AND SEGERS, J. (2009). Regularly varying multivariate time series. *Stoch. Process. Appl.* **119**, 1055–1080.

[2] BHATIA, R. (1997). *Matrix Analysis* (Graduate Texts Math. **169**). Springer, New York.

[3] BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1987). *Regular Variation*. Cambridge University Press, Cambridge.

[4] CLINE, D. B. H. AND HSING, T. (1998). Large deviation probabilities for sums of random variables with heavy or subexponential tails. Technical report, Texas A&M University. Available at https://www.stat.tamu.edu/~dcline/Papers/large5.pdf.

[5] DAVIS, R. A. AND HSING, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Prob.* **23**, 879–917.

[6] DENISOV, D., DIEKER, A. B. AND SHNEER, V. (2008). Large deviations for random walks under subexponentiality: the big-jump domain. *Ann. Prob.* **36**, 1946–1991.

[7] DURRETT, R. (1979). Maxima of branching random walks vs. independent random walks. *Stoch. Process. Appl.* **9**, 117–135

[8] EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.

[9] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, vol. II, 2nd edn. Wiley, New York.

[10] FOSS, S., KORSHUNOV, D. AND ZACHARY, S. (2013). *An Introduction to Heavy-Tailed and Subexponential Distributions*, 2nd edn. Springer, New York.

[11] GANTERT, N. AND HÖFELSAUER, T. (2019). Large deviations for the maximum of a branching random walk. *Electron. J. Prob.* **23**, 1–12.

[12] HEINY, J. AND MIKOSCH, T. (2017). Eigenvalues and eigenvectors of heavy-tailed sample covariance matrices with general growth rates: the iid case. *Stoch. Process. Appl.* **127**, 2179–2242.

[13] HEINY, J., MIKOSCH, T. AND YSLAS, J. (2019). Gumbel convergence of the maximum entry in a sample covariance matrix. Technical report.

[14] HULT, H., LINDSKOG, F., MIKOSCH, T. AND SAMORODNITSKY G. (2005). Functional large deviations for multivariate regularly varying random walks. *Ann. Appl. Prob.* **15**, 2651–2680.

[15] IBRAGIMOV, I. A. AND LINNIK, Y. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen.

[16] KLÜPPELBERG, C. AND MIKOSCH, T. (1997). Large deviation of heavy-tailed random sums with applications in insurance and finance. *J. Appl. Prob.* **34**, 293–308.

[17] LINNIK, Y. V. (1961). Limit theorems allowing large deviations for sums of independent variables I, II. *Theory Prob. Appl.* **6**, 145–161, 377–391.

[18] MICHEL, R. (1974). Results on probabilities of moderate deviations. *Ann. Prob.* **2**, 349–353.

[19] MIKOSCH, T. AND NAGAEV, A. V. (1998). Large deviations of heavy-tailed sums with applications in insurance. *Extremes* **1**, 81–110.

[20] MIKOSCH, T. AND WINTENBERGER, O. (2016). A large deviations approach to limit theory for heavy-tailed time series. *Prob. Theory Relat. Fields* **166** 233–269.

[21] NAGAEV, S. V. (1965). Limit theorems on large deviations. *Theory Prob. Appl.* **10** 231–254.

[22] NAGAEV, A. V. (1969). Limit theorems for large deviations where Cramér's conditions are violated (in Russian). *Izv. Akad. Nauk UzSSR Ser. Fiz.–Mat. Nauk* **6**, 17–22.

[23] NAGAEV, A. V. (1969). Integral limit theorems for large deviations when Cramér's condition is not fulfilled I, II. *Theory Prob. Appl.* **14**, 51–64, 193–208.

[24] NAGAEV, A. V. (1977). A property of sums of independent random variables. *Theory Prob. Appl.* **22**, 335–346.

[25] NAGAEV, S. V. (1979). Large deviations of sums of independent random variables. *Ann. Prob.* **7**, 745–789.

[26] PETROV, V. V. (1972). *Sums of Independent Random Variables* (in Russian). Nauka, Moscow.

[27] PETROV, V. V. (1995). *Limit Theorems of Probability Theory*. Oxford University Press, Oxford.

[28] RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*, 2008 reprint. Springer, New York.

[29] RESNICK, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.

[30] RESNICK, S. I. AND STĂRICĂ, C. (1995). Consistency of Hill's estimator for dependent data. *J. Appl. Prob.* **32**, 139–167.

[31] ROZOVSKI, L. V. (1993). Probabilities of large deviations on the whole axis. *Theory Prob. Appl.* **38**, 53–79.