J. Math. Anal. Appl. 411 (2014) 271–280

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Large deviations for heavy-tailed random elements in convex cones

Christoph Kopp [1], Ilya Molchanov *

Institute of Mathematical Statistics and Actuarial Science, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland

ARTICLE INFO

Article history:
Received 31 May 2013
Available online 25 September 2013
Submitted by U. Stadtmueller

Keywords:
Large deviation
Convex cone
Heavy-tailed distribution
Embedding
Random set

ABSTRACT

We prove large deviation results for sums of heavy-tailed random elements in rather general convex cones being semigroups equipped with a rescaling operation by positive real numbers. In difference to previous results for the cone of convex sets, our technique does not use the embedding of cones in linear spaces. Examples include the cone of convex sets with the Minkowski addition, positive half-line with maximum operation and the family of square integrable functions with arithmetic addition and argument rescaling.

1. Introduction

Most results concerning limiting behavior of sums of random elements in linear spaces can be extended for random closed sets in linear spaces, see [10, Ch. 3]. The sum of sets is defined in the Minkowski sense, i.e. the sum of two sets is the closure of the set of pairwise sums of elements from these sets. It is well known that this addition is not invertible. The most typical way to handle this setting is to consider first random convex compact sets and embed them into the Banach space of continuous functions on the unit sphere in the dual space using the support function. Then the Minkowski sum of sets corresponds to the arithmetic sum of their support functions and the Hausdorff distance between sets turns into the uniform distance in the space of support functions, which opens the possibility to use the results available for random elements in Banach spaces, see e.g. [4]. Finally, it is usually argued that the results for possibly non-convex random compact sets are identical to their convex case counterparts in view of the convexification property of the Minkowski addition, see [1].

The family of limit theorems for random sets has been recently extended with several large deviation results in the heavy-tail setting in [8] and [9]. The crucial assumption is the regular variation condition on the tail, which is similar to one that appears in limit theorems for unions of random closed sets, see [10, Ch. 4]. Let S_n denote the Minkowski sum of i.i.d. regularly varying random compact sets $\xi_1, \dots, \xi_n$ in $\mathbb{R}^m$ with tail index $\alpha > 0$ and tail measure μ. In particular, [8] show that

$$\gamma_n \mathbf{P}(S_n \in \lambda_n U) \to \mu(U) \tag{1}$$

* Corresponding author.
E-mail addresses: chris.kopp@stat.unibe.ch (C. Kopp), ilya.molchanov@stat.unibe.ch (I. Molchanov).
[1] Partially supported by Swiss National Foundation Grant Nr. 200021-126503.

0022-247X/$ – see front matter
http://dx.doi.org/10.1016/j.jmaa.2013.09.042

for λ_n that grows sufficiently fast and all μ-continuous measurable subsets U of the family of all compact sets bounded away from zero. The sequence of normalizing constants $\{\gamma_n\}$ is related to the tail behavior of the norm of a single random compact set defined as its Hausdorff distance to the origin. Especially, it is required that $\lambda_n/n \to \infty$ in case $\alpha \geqslant 1$.

This large deviation result has been refined in [9], where it is shown that, for regularly varying convex random compact sets with integrable norm (so $\alpha \geqslant 1$),

$$\gamma_n \mathbf{P}(S_n \in \lambda_n U + n\mathbf{E}\xi_1) \to \mu(U), \tag{2}$$

where λ_n grows slower than in (1) and $\mathbf{E}\xi_1$ is the expectation of ξ_1, see [10, Sec. 2.1]. The method of the proof is based on the embedding argument combined with a use of classical large deviation results from [11] and [6].

The setting of random compact sets can be considered as a special case of random elements in convex cones (also called conlinear spaces), being semigroups with a scaling operation by positive reals, see [3]. A simple example is the cone of positive numbers with the maximum operation. It should be noted that in that case the embedding argument is not applicable any longer, so one has to prove the corresponding results in the cone without using any centering or symmetrization arguments.

In this paper, we generalize the above mentioned results from [8] and [9] for heavy-tailed random elements in convex cones. While the general scheme of our proofs follows the lines of the proofs from [8] and [9], it requires extra care caused by the impossibility to use the embedding device. In particular, this concerns our generalization of (2), since there is no generally consistent definition of the expectation in convex cones.

2. Regularly varying random elements in cones

A Borel function $f : (c,\infty) \mapsto (0,\infty)$ for some $c > 0$ is said to be *regularly varying* (at infinity) with index ρ if

$$f(\lambda x)/f(x) \to \lambda^\rho \quad \text{as } x \to \infty$$

for all $\lambda > 0$, see e.g. [2]. If $\rho = 0$, then f is called *slowly varying* and usually denoted by the letter ℓ instead of f. Any regularly varying function f with index ρ has a representation $f(x) = x^\rho \ell(x)$ for a slowly varying function ℓ. We write $f \sim g$ as a shorthand for $f(x)/g(x) \to 1$ as $x \to \infty$.

Theorem 2.1 *(Karamata). (See Th. 1.5.11 [2].) If f is regularly varying with index ρ and locally bounded on $[a,\infty)$, then*

(i) *for any $\beta \geqslant -(\rho+1)$,*

$$\lim_{x\to\infty} \frac{x^{\beta+1} f(x)}{\int_a^x t^\beta f(t)\,\mathrm{d}t} = \beta + \rho + 1;$$

(ii) *for any $\beta < -(\rho+1)$ (and for $\beta = -(\rho+1)$ if $\int_a^\infty t^{-(\rho+1)} f(t)\,\mathrm{d}t < \infty$),*

$$\lim_{x\to\infty} \frac{x^{\beta+1} f(x)}{\int_x^\infty t^\beta f(t)\,\mathrm{d}t} = -(\beta + \rho + 1).$$

Below we summarize several concepts from [3] concerning general convex cones. A *convex cone* $\mathbb{K}$ is a topological semigroup with neutral element $\mathbf{e}$ and an extra operation $x \mapsto ax$ of scaling $x \in \mathbb{K}$ by a positive number a, so that $a(x+y) = ax + ay$ for all $a > 0$, $x, y \in \mathbb{K}$. It should be noted that we do *not* require the validity of the second distributivity law $(a+b)x = ax + bx$. The second distributivity law holds for the cone of compact sets in $\mathbb{R}^d$ with the Minkowski addition and enables using the embedding argument.

We assume that $\mathbb{K}$ is a pointed cone, i.e. ax converges to the cone element $\mathbf{0}$ called the origin as $a \downarrow 0$ for all $x \neq \mathbf{e}$. Assume that $\mathbb{K}$ is metrized by a homogeneous metric d, i.e. $d(ax, ay) = ad(x, y)$ for all $x, y \in \mathbb{K}$ and $a > 0$. The value $\|x\| = d(x, \mathbf{0})$ is called the *norm* of x which in general constitutes an abuse of language since $\|\cdot\|$ is not necessarily sub-linear. Nevertheless, the norm is sub-linear if the metric is *sub-invariant*, i.e. if $d(x+h, x) \leqslant d(h, \mathbf{0}) = \|h\|$ for all $x, h \in \mathbb{K}$. A stronger assumption is the translation-invariance of the metric meaning that $d(x+h, y+h) = d(x, y)$ for all $x, y, h \in \mathbb{K}$. In a cone with sub-invariant metric, $\mathbf{0} = \mathbf{e}$, see [3, Lemma 2.7].

Furthermore, $\mathbb{S} = \{x \in \mathbb{K}\colon\ \|x\| = 1\}$ denotes the unit sphere. For $\varepsilon > 0$,

$$A^\varepsilon = \left\{x \in \mathbb{K}\colon\ d(x, A) \leqslant \varepsilon\right\}$$

is the ε-envelope of $A \subset \mathbb{K}$, where $d(x, A) = \inf_{a\in A} d(x, a)$. The Borel σ-algebra on $\mathbb{K}$ is denoted by $\mathcal{B}$ and used to define random cone elements ξ as measurable maps from a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ to $(\mathbb{K}, \mathcal{B})$.

Furthermore, $\operatorname{int} A$, $\operatorname{cl} A$ and ∂A denote the interior, closure and boundary of $A \subset \mathbb{K}$. A set $A \subset \mathbb{K}$ is said to be bounded away from a point $x \in \mathbb{K}$ if $x \notin \operatorname{cl} A$. If μ is a measure on $\mathcal{B}$, then $A \in \mathcal{B}$ is called a μ-continuity set if $\mu(\partial A) = 0$.

Let $M_{\mathbf{0}}$ be the space of all Borel measures on $\mathbb{K}' = \mathbb{K} \setminus \{\mathbf{0}\}$ taking finite values on $\mathbb{K} \setminus \{x \in \mathbb{K}\colon \|x\| \leqslant r\}$ for each $r > 0$. By $\mathcal{C}_0$ we denote the class of all real-valued bounded continuous functions on $\mathbb{K}'$ with support bounded away from $\mathbf{0}$. A sequence $\{\mu_n,\ n \geqslant 1\}$ of measures from $M_{\mathbf{0}}$ is said to converge to $\mu \in M_{\mathbf{0}}$ if

$$\int f \, \mathrm{d}\mu_n \to \int f \, \mathrm{d}\mu \quad \text{as } n \to \infty$$

for all $f \in \mathcal{C}_0$, equivalently $\mu_n(U) \to \mu(U)$ for all μ-continuity sets $U \in \mathcal{B}$ bounded away from $\mathbf{0}$.

The following definition does not rely on the semigroup operation and is available for random elements in metric spaces, where a scaling by positive real numbers is defined.

Definition 2.1. (See [5].) A random cone element ξ is called *regularly varying* (at infinity) if there exist a non-null measure $\mu \in M_{\mathbf{0}}$ and a sequence $\{a_n,\ n \geqslant 1\}$ of positive numbers such that

$$n\mathbf{P}(\xi \in a_n \cdot) \to \mu(\cdot) \quad \text{in } M_{\mathbf{0}} \text{ as } n \to \infty.$$

The *tail measure* μ necessarily scales like a power function, i.e. $\mu(\lambda U) = \lambda^{-\alpha}\mu(U)$ for every $\lambda > 0$, all μ-continuous U bounded away from $\mathbf{0}$ and $\alpha > 0$ called the *index* of regular variation of ξ.

By [5, Th. 3.1], regular variation of ξ implies

$$\frac{\mathbf{P}(\xi \in t \cdot)}{\mathbf{P}(\|\xi\| > t)} \to c\mu(\cdot) \quad \text{in } M_{\mathbf{0}} \text{ as } t \to \infty \tag{3}$$

for some $c > 0$. It will subsequently be assumed that $c = 1$ in (3), which is possible by scaling $\{a_n\}$.

By [5, Th. 3.1], ξ is regularly varying with index $\alpha > 0$ if and only if there exist a finite measure σ (called the *spectral measure*) on the unit sphere $\mathbb{S}$ and a sequence $\{\tilde{a}_n\}$ such that

$$\lim_{n\to\infty} n\mathbf{P}\big(\|\xi\|^{-1}\xi \in B, \|\xi\| > r\tilde{a}_n\big) = \sigma(B)r^{-\alpha}$$

for all $r > 0$ and all Borel $B \subset \mathbb{S}$ with $\sigma(\partial B) = 0$. It holds that $\tilde{a}_n \sim a_n$.

Karamata's theorem implies the following result.

Corollary 2.2. *Let ξ be regularly varying with index $\alpha > 0$ and let $T > 0$ and $\gamma > \alpha$. Then*

$$\mathbf{E}\big(\|\xi\|\mathbf{1}_{\|\xi\|\leqslant T}\big)^{\gamma} = \gamma \int_0^T \mathbf{P}\big(\|\xi\| > t\big)t^{\gamma-1}\,\mathrm{d}t \sim cT^{\gamma}\mathbf{P}\big(\|\xi\| > T\big) \quad \text{as } T \to \infty,$$

where $c > 0$ denotes a finite constant.

The letter c (also with subscripts) denotes finite, strictly positive constants; its value may change at every occurrence.

3. Large deviations with strong scaling

Consider a sequence $\{\xi_n\}_{n\geqslant 1}$ of i.i.d. random elements in $\mathbb{K}'$ and their partial sums

$$S_n = \xi_1 + \cdots + \xi_n, \quad n \geqslant 1.$$

Theorem 3.1. *Let $\xi, \xi_1, \ldots$ be i.i.d. regularly varying random elements with index $\alpha > 0$, tail measure μ and normalizing sequence $\{a_n\}$ in a convex cone with sub-invariant homogeneous metric d. Let $\{\lambda_n, n \geqslant 1\}$ be a sequence such that*

(i) $\lambda_n/a_n \to \infty$ *if* $\alpha < 1$,
(ii) $\lambda_n/n \to \infty$, $\lambda_n/a_n \to \infty$ *and* $(n/\lambda_n)\mathbf{E}(\|\xi\|\mathbf{1}_{\|\xi\|\leqslant\lambda_n}) \to 0$ *if* $\alpha = 1$,
(iii) $\lambda_n/n \to \infty$ *if* $\alpha > 1$.

Then

$$\gamma_n\mathbf{P}(S_n \in \lambda_n \cdot) \to \mu(\cdot) \quad \text{in } M_{\mathbf{0}} \text{ as } n \to \infty,$$

where $\gamma_n = (n\mathbf{P}(\|\xi\| > \lambda_n))^{-1}$.

The proof closely follows the lines of the proof of [8, Th. 1]. Note that the sequence $\{\lambda_n\}$ grows faster than $n^{\max(1,1/\alpha)}$.

Lemma 3.2. *In the setting of Theorem 3.1, $\lambda_n^{-1}\|S_n\| \to 0$ in probability.*

Proof. The sub-invariance of the metric implies the sub-linearity of the norm, thus it suffices to show that

$$\frac{\|\xi_1\| + \cdots + \|\xi_n\|}{\lambda_n} \to 0 \quad \text{in probability.} \tag{4}$$

If $\alpha > 1$, then $\mathbf{E}\|\xi\| < \infty$ and the strong law of large numbers with the growth conditions on $\{\lambda_n\}$ provides the result.
Assume $0 < \alpha \leqslant 1$. By [11, Th. 4.13], (4) holds if (and only if) the following three conditions hold.

(i) $n\mathbf{P}(\|\xi\| > \lambda_n) \to 0$. This is the case, since

$$n\mathbf{P}\big(\|\xi\| > \lambda_n\big) = n\mathbf{P}\big(\|\xi\| > a_n\big)\frac{\mathbf{P}(\|\xi\| > \lambda_n)}{\mathbf{P}(\|\xi\| > a_n)},$$

where the first factor converges to one and the fraction converges to zero.

(ii) $\lambda_n^{-1} n\mathbf{E}(\|\xi\|\mathbf{1}_{\|\xi\|<\lambda_n}) \to 0$, which follows from Corollary 2.2 in case $\alpha < 1$, while for $\alpha = 1$ the convergence is assumed.

(iii) $\lambda_n^{-2} n \operatorname{Var}(\|\xi\|\mathbf{1}_{\|\xi\|<\lambda_n}) \to 0$. To confirm this, bound the variance by the second moment and apply Corollary 2.2. □

Proof of Theorem 3.1. Let $U \in \mathcal{B}$ with $U \neq \emptyset$, $\mu(\partial U) = 0$ and $\mathbf{0} \notin \operatorname{cl} U$. We start by bounding $\gamma_n \mathbf{P}(S_n \in \lambda_n U)$ from above. For any $\varepsilon > 0$,

$$\begin{aligned}
\mathbf{P}(S_n \in \lambda_n U) &= \mathbf{P}\left(S_n \in \lambda_n U, \bigcup_{i=1}^n \{\xi_i \in \lambda_n U^\varepsilon\}\right) + \mathbf{P}\left(S_n \in \lambda_n U, \bigcap_{i=1}^n \{\xi_i \notin \lambda_n U^\varepsilon\}\right) \\
&\leqslant n\mathbf{P}\big(\xi_1 \in \lambda_n U^\varepsilon\big) + \mathbf{P}\left(\bigcap_{i=1}^n \{d(S_n, \xi_i) > \varepsilon\lambda_n\}\right) \\
&= I_1 + I_2,
\end{aligned}$$

since $S_n \in \lambda_n U$ and $\xi_i \notin \lambda_n U^\varepsilon$ imply that $d(\lambda_n^{-1}S_n, \lambda_n^{-1}\xi_i) > \varepsilon$ so that $d(S_n, \xi_i) > \varepsilon\lambda_n$ by the homogeneity of d.
By (3), $\gamma_n I_1 \to \mu(U^\varepsilon)$ as $n \to \infty$ for μ-continuity sets U^ε with $\mathbf{0} \notin \operatorname{cl} U^\varepsilon$. Note that $\mathbf{0} \notin \operatorname{cl} U^\varepsilon$ for sufficiently small ε and that all but countably many U^ε-sets are μ-continuity sets since μ is finite outside any neighborhood of $\mathbf{0}$. It follows that $\mu(U^\varepsilon) \to \mu(U)$ as $\varepsilon \downarrow 0$.
To show that $\gamma_n I_2 \to 0$ as $n \to \infty$ for every $\varepsilon > 0$, consider for $\delta > 0$ the following events partitioning the probability space:

$$\begin{aligned}
D_1 &= \bigcup_{1 \leqslant i < j \leqslant n} \big\{\|\xi_i\| > \delta\lambda_n,\ \|\xi_j\| > \delta\lambda_n\big\}, \\
D_2 &= \bigcup_{i=1}^n \big\{\|\xi_i\| > \delta\lambda_n,\ \|\xi_j\| \leqslant \delta\lambda_n,\ j \neq i,\ j = 1, \ldots, n\big\}, \\
D_3 &= \Big\{\max_{i=1,\ldots,n} \|\xi_i\| \leqslant \delta\lambda_n\Big\}.
\end{aligned}$$

By the Bonferroni inequality and the independence of the ξ_i,

$$\gamma_n \mathbf{P}(D_1) \leqslant \binom{n}{2} \frac{\mathbf{P}(\|\xi_1\| > \delta\lambda_n)^2}{n\mathbf{P}(\|\xi\| > \lambda_n)},$$

which converges to zero as $n \to \infty$ because of the regular variation property and the growth condition on $\{\lambda_n\}$.
By the sub-invariance of the metric, $d(S_n, \xi_n) \leqslant \|S_{n-1}\|$. Therefore,

$$\begin{aligned}
\mathbf{P}\left(\bigcap_{i=1}^n \{d(S_n, \xi_i) > \varepsilon\lambda_n\}, D_2\right) &\leqslant \sum_{i=1}^n \mathbf{P}\big(d(S_n, \xi_i) > \varepsilon\lambda_n, \|\xi_i\| > \delta\lambda_n\big) \\
&\leqslant n\mathbf{P}\big(\|S_{n-1}\| > \varepsilon\lambda_n\big)\mathbf{P}\big(\|\xi_n\| > \delta\lambda_n\big)
\end{aligned}$$

which, if multiplied by γ_n, converges to zero as $n \to \infty$ because of the regular variation of ξ and Lemma 3.2.
Regarding D_3, by sub-invariance

$$\mathbf{P}\left(\bigcap_{i=1}^n \{d(S_n, \xi_i) > \varepsilon\lambda_n\}, \max_{i=1,\ldots,n} \|\xi_i\| \leqslant \delta\lambda_n\right) \leqslant \mathbf{P}\left(\sum_{i=1}^n \|\xi_i\|\mathbf{1}_{\|\xi_i\| \leqslant \delta\lambda_n} > \varepsilon\lambda_n\right),$$

which, after centering, becomes

$$\mathbf{P}\Biggl(\sum_{i=1}^{n}\bigl(\|\xi_i\|\mathbf{1}_{\|\xi_i\|\leqslant\delta\lambda_n}-\mathbf{E}\bigl(\|\xi_i\|\mathbf{1}_{\|\xi_i\|\leqslant\delta\lambda_n}\bigr)\bigr)>\varepsilon\lambda_n-n\mathbf{E}\bigl(\|\xi_1\|\mathbf{1}_{\|\xi_1\|\leqslant\delta\lambda_n}\bigr)\Biggr).$$

Since $n\lambda_n^{-1}\mathbf{E}(\|\xi_1\|\mathbf{1}_{\|\xi_1\|\leqslant\delta\lambda_n})\to 0$ (see the proof of Lemma 3.2), the right-hand side of the above inequality may be replaced by $\varepsilon\lambda_n/2$.

It remains to show that

$$\gamma_n\mathbf{P}\Biggl(\sum_{i=1}^{n}\bigl(\|\xi_i\|\mathbf{1}_{\|\xi_i\|\leqslant\delta\lambda_n}-\mathbf{E}\bigl(\|\xi_i\|\mathbf{1}_{\|\xi_i\|\leqslant\delta\lambda_n}\bigr)\bigr)>\varepsilon\lambda_n/2\Biggr)\to 0\quad\text{as } n\to\infty.$$

Since each summand

$$\eta_i=\|\xi_i\|\mathbf{1}_{\|\xi_i\|\leqslant\delta\lambda_n}-\mathbf{E}\bigl(\|\xi_i\|\mathbf{1}_{\|\xi_i\|\leqslant\delta\lambda_n}\bigr)$$

is centered and $\mathbf{E}|\eta_i|^p<\infty$ for any $p\geqslant 2$, the Fuk–Nagaev inequality (see e.g. [11, p. 78]) yields that

$$\mathbf{P}\Biggl(\sum_{i=1}^{n}\eta_i>\frac{\varepsilon\lambda_n}{2}\Biggr)\leqslant c_1 n(\varepsilon\lambda_n)^{-p}\mathbf{E}|\eta_1|^p+\exp\biggl\{-\frac{c_2(\varepsilon\lambda_n)^2}{(n\operatorname{Var}\eta_1)}\biggr\}=I_{3,1}+I_{3,2}$$

for any $p\geqslant 2$ where $c_1,c_2>0$ are finite constants.

By Corollary 2.2,

$$\mathbf{E}|\eta_1|^p\leqslant\mathbf{E}\bigl(\|\xi_1\|\mathbf{1}_{\|\xi_1\|\leqslant\delta\lambda_n}\bigr)^p\sim c(\delta\lambda_n)^p\mathbf{P}\bigl(\|\xi_1\|>\delta\lambda_n\bigr)\quad\text{as } n\to\infty$$

for $p>\alpha$. For $p>\max\{2,\alpha\}$,

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\gamma_n I_{3,1}\leqslant c\lim_{\delta\downarrow 0}\delta^p\lim_{n\to\infty}\frac{\mathbf{P}(\|\xi_1\|>\delta\lambda_n)}{\mathbf{P}(\|\xi_1\|>\lambda_n)}=c\lim_{\delta\downarrow 0}\delta^{p-\alpha}=0.$$

To show that $\limsup_{n\to\infty}\gamma_n I_{3,2}=0$, consider these (disjoint) cases:

(i) If $\alpha\geqslant 2$ and $\operatorname{Var}\|\xi_1\|<\infty$, then $\operatorname{Var}\eta_1<\infty$ and the convergence follows.
(ii) If $0<\alpha<2$, then $\lambda_n^{-2}n\operatorname{Var}(\|\xi_1\|\mathbf{1}_{\|\xi_1\|\leqslant\delta\lambda_n})\sim cn\mathbf{P}(\|\xi\|>\delta\lambda_n)$ by Corollary 2.2, which implies the convergence.
(iii) If $\alpha=2$ and $\operatorname{Var}\|\xi\|=\infty$, then $\lambda_n^2\mathbf{P}(\|\xi\|>\lambda_n)$ and $\operatorname{Var}(\|\xi\|\mathbf{1}_{\|\xi\|\leqslant\delta\lambda_n})$ are both slowly varying functions of λ_n. Because $\lambda_n/n\to\infty$, the convergence follows.

Thus $\limsup_{n\to\infty}\gamma_n I_{3,2}=0$. Hence

$$\limsup_{n\to\infty}\gamma_n\mathbf{P}(S_n\in\lambda_n U)\leqslant\mu\bigl(U^\varepsilon\bigr)\to\mu(U)\quad\text{as }\varepsilon\downarrow 0$$

for any U bounded away from $\mathbf{0}$, establishing the upper bound.

For the lower bound, let $U\in\mathcal{B}$ now denote a μ-continuity set bounded away from $\mathbf{0}$ with nonempty interior. The set $U^{-\varepsilon}=((U^c)^\varepsilon)^c$ is bounded away from $\mathbf{0}$, is a μ-continuity set for all but countably many ε and $\operatorname{int}(U^{-\varepsilon})$ is nonempty for sufficiently small ε.

Writing $S_n^{\neq i}=\sum_{j=1,j\neq i}^n\xi_j$ for $i=1,\dots,n$, we see that

$$\begin{aligned}
\mathbf{P}(S_n\in\lambda_n U)&\geqslant\mathbf{P}\Biggl(S_n\in\lambda_n U,\bigcup_{i=1}^{n}\bigl\{\xi_i\in\lambda_n U^{-\varepsilon}\bigr\}\Biggr)\\
&\geqslant\mathbf{P}\Biggl(\bigcup_{i=1}^{n}\bigl\{d(S_n,\xi_i)<\varepsilon\lambda_n,\ \xi_i\in\lambda_n U^{-\varepsilon}\bigr\}\Biggr)\\
&\geqslant\mathbf{P}\Biggl(\bigcup_{i=1}^{n}\bigl\{\bigl\|S_n^{\neq i}\bigr\|<\varepsilon\lambda_n,\ \xi_i\in\lambda_n U^{-\varepsilon}\bigr\}\Biggr)\\
&\geqslant n\mathbf{P}\bigl(\bigl\|S_n^{\neq 1}\bigr\|<\varepsilon\lambda_n\bigr)\mathbf{P}\bigl(\xi_1\in\lambda_n U^{-\varepsilon}\bigr)-\binom{n}{2}\mathbf{P}\bigl(\xi_1\in\lambda_n U^{-\varepsilon}\bigr)^2\\
&=J_1-J_2,
\end{aligned}$$

where the second inequality holds because $d(S_n,\xi_i)<\varepsilon\lambda_n$ and $\xi_i\in\lambda_n U^{-\varepsilon}$ imply that $\lambda_n^{-1}S_n\in(U^{-\varepsilon})^\varepsilon\subset U$. The third inequality is implied by the sub-invariance of the metric.

By Lemma 3.2, $\mathbf{P}(\|S_n^{\neq 1}\| < \varepsilon\lambda_n) \to 1$ as $n \to \infty$ and

$$\liminf_{n\to\infty} \gamma_n J_1 \geqslant \lim_{n\to\infty} \frac{\mathbf{P}(\xi_1 \in \lambda_n U^{-\varepsilon})}{\mathbf{P}(\|\xi\| > \lambda_n)} = \mu\bigl(U^{-\varepsilon}\bigr)$$

which converges to $\mu(U)$ as $\varepsilon \downarrow 0$. Finally,

$$\limsup_{n\to\infty} \gamma_n J_2 \leqslant c \lim_{n\to\infty} n\mathbf{P}\bigl(\xi_1 \in \lambda_n U^{-\varepsilon}\bigr) = 0,$$

which establishes the lower bound and finishes the proof. □

4. Moderate scaling

Theorem 3.1 requires that the normalizing sequence $\{\lambda_n\}$ grows faster than n in case $\alpha \geqslant 1$. If λ_n grows slower than n, but faster than $n^{\max(1/\alpha,1/2)}$, then the large deviation result holds with an extra additive normalization. Care is required however, since the addition operation in general cones is not invertible and the expectation is not well defined.

Theorem 4.1. *Let $\xi, \xi_1, \ldots$ be i.i.d. regularly varying random elements with index $\alpha \geqslant 1$ and tail measure μ in a convex cone $\mathbb{K}'$ with a homogeneous sub-invariant metric d. Assume that $\mathbf{E}\|\xi\|$ is finite and there exists a sequence $\{A_n, n \geqslant 1\}$ of cone elements such that*

$$\lambda_n^{-1}\mathbf{E}d(S_n, A_n) \to 0 \quad \text{as } n \to \infty, \tag{5}$$

where $\lambda_n/n^{\max\{1/\alpha,1/2\}+\eta} \to \infty$ for some $\eta > 0$. If

(A) $A_n = \mathbf{e}$ *for all n,* *or* (B) *the metric d is invariant*

then, with $\gamma_n = (n\mathbf{P}(\|\xi\| > \lambda_n))^{-1}$,

$$\gamma_n \mathbf{P}(S_n \in \lambda_n \cdot + A_n) \to \mu(\cdot) \quad \text{in } M_{\mathbf{0}} \text{ as } n \to \infty.$$

The following result known in the setting of Banach spaces (see [7, Lemma 6.16]) extends to general semigroups and will be used to prove Theorem 4.1.

Let $\zeta_1, \ldots, \zeta_n$ be integrable random elements in a semigroup $\mathbb{K}$ with sub-invariant metric d. For $i \leqslant n$ write $\mathcal{A}_i = \sigma(\zeta_1, \ldots, \zeta_i)$ and let $\mathcal{A}_0$ denote the trivial sigma-algebra. Write $S_n^{\neq i} = \sum_{j=1,\, j\neq i}^n \zeta_j$ for $i \leqslant n$. Let z be any fixed cone element and define

$$d_i = \mathbf{E}\bigl(d(S_n, z) \bigm| \mathcal{A}_i\bigr) - \mathbf{E}\bigl(d(S_n, z) \mid \mathcal{A}_{i-1}\bigr), \quad i = 1, \ldots, n.$$

Then $d_1, \ldots, d_n$ is a real-valued martingale difference sequence and $\sum_{i=1}^n d_i = d(S_n, z) - \mathbf{E}d(S_n, z)$ almost surely.

Lemma 4.2. *Let $\zeta_1, \ldots, \zeta_n$ be independent. Then*

$$|d_i| \leqslant \|\zeta_i\| + \mathbf{E}\|\zeta_i\|$$

almost surely for every $i = 1, \ldots, n$.

Proof. Since $S_n^{\neq i}$ is independent of ζ_i,

$$\begin{aligned} d_i &= \mathbf{E}\bigl(d(S_n, z) \bigm| \mathcal{A}_i\bigr) - \mathbf{E}\bigl(d(S_n, z) \bigm| \mathcal{A}_{i-1}\bigr) + \mathbf{E}\bigl(d\bigl(S_n^{\neq i}, z\bigr) \bigm| \mathcal{A}_{i-1}\bigr) - \mathbf{E}\bigl(d\bigl(S_n^{\neq i}, z\bigr) \bigm| \mathcal{A}_i\bigr) \\ &= \mathbf{E}\bigl(d(S_n, z) - d\bigl(S_n^{\neq i}, z\bigr) \bigm| \mathcal{A}_i\bigr) - \mathbf{E}\bigl(d(S_n, z) - d\bigl(S_n^{\neq i}, z\bigr) \bigm| \mathcal{A}_{i-1}\bigr), \end{aligned}$$

where the equalities hold almost surely. The sub-invariance property yields that $d(S_n, z) - d(S_n^{\neq i}, z) \leqslant d(S_n, S_n^{\neq i}) \leqslant \|\zeta_i\|$, giving the result. □

Proof of Theorem 4.1. We start with an upper bound. Let $\varepsilon > 0$ and $U \in \mathcal{B}$ be a nonempty μ-continuity set bounded away from $\mathbf{0}$. Then

$$\begin{aligned} \mathbf{P}(S_n \in \lambda_n U + A_n) &= \mathbf{P}\Biggl(S_n \in \lambda_n U + A_n, \bigcup_{i=1}^n \bigl\{\xi_i \in \lambda_n U^{\varepsilon}\bigr\}\Biggr) + \mathbf{P}\Biggl(S_n \in \lambda_n U + A_n, \bigcap_{i=1}^n \bigl\{\xi_i \notin \lambda_n U^{\varepsilon}\bigr\}\Biggr) \\ &\leqslant n\mathbf{P}\bigl(\xi_1 \in \lambda_n U^{\varepsilon}\bigr) + \mathbf{P}\Biggl(S_n \in \lambda_n U + A_n, \bigcap_{i=1}^n \bigl\{\xi_i \notin \lambda_n U^{\varepsilon}\bigr\}\Biggr) \\ &= I_1 + I_2. \end{aligned}$$

As in the proof of Theorem 3.1, using (3) and the μ-continuity of U,

$$\mu(U) \leqslant \lim_{\varepsilon\downarrow 0}\liminf_{n\to\infty}\gamma_n I_1 \leqslant \lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\gamma_n I_1 = \mu(U).$$

Now fix $0<\delta\leqslant\varepsilon/3$ and partition Ω into

$$D_1 = \bigcup_{i=1}^n \{\|\xi_i\| > \delta\lambda_n\} \quad \text{and} \quad D_2 = \Big\{\max_{i=1,\dots,n}\|\xi_i\| \leqslant \delta\lambda_n\Big\}.$$

Starting with D_1, we see that

$$\begin{aligned}\mathbf{P}\left(S_n\in\lambda_n U + A_n, \bigcap_{k=1}^n \{\xi_k\notin\lambda_n U^\varepsilon\}, D_1\right) &\leqslant n\mathbf{P}\big(S_n\in\lambda_n U + A_n, \xi_1\notin\lambda_n U^\varepsilon, \|\xi_1\|>\delta\lambda_n\big)\\ &\leqslant n\mathbf{P}\big(d(\xi_2+\dots+\xi_n, A_n)>\varepsilon\lambda_n, \|\xi_1\|>\delta\lambda_n\big).\end{aligned}$$

To justify the last step, define the event

$$C = \big\{S_n\in\lambda_n U + A_n,\ \xi_1\notin\lambda_n U^\varepsilon\big\}.$$

If (A) holds, $C=\{\lambda_n^{-1}S_n\in U,\ \lambda_n^{-1}\xi_1\notin U^\varepsilon\}$ so that $d(S_n,\xi_1)>\varepsilon\lambda_n$ by the homogeneity of d. Now $d(S_n,\xi_1)\leqslant d(\xi_2+\dots+\xi_n, A_n)$ by the sub-invariance.

If (B) holds, $\xi_1\notin\lambda_n U^\varepsilon$ if and only if $\xi_1 + A_n\notin\lambda_n U^\varepsilon + A_n$ so that

$$C = \big\{S_n\in\lambda_n U + A_n,\ \xi_1 + A_n\notin\lambda_n U^\varepsilon + A_n\big\},$$

implying that $d(S_n,\xi_1+A_n)>\varepsilon\lambda_n$. Now apply the invariance of d again.

By independence of the ξ_i,

$$\gamma_n n\mathbf{P}\big(d(\xi_2+\dots+\xi_n, A_n)>\varepsilon\lambda_n, \|\xi_1\|>\delta\lambda_n\big) = \mathbf{P}\big(\lambda_n^{-1}d(\xi_2+\dots+\xi_n, A_n)>\varepsilon\big)\frac{\mathbf{P}(\|\xi_1\|>\delta\lambda_n)}{\mathbf{P}(\|\xi\|>\lambda_n)}.$$

The fraction in the right-hand side converges to $\delta^{-\alpha}$ as $n\to\infty$. The first factor converges to zero, since

$$\lambda_n^{-1}d\big(S_n^{\neq 1}, A_n\big)\leqslant\lambda_n^{-1}\|\xi_1\| + \lambda_n^{-1}d(S_n, A_n)$$

converges to zero in probability because $\lambda_n^{-1}\|\xi_1\|\to 0$ in probability and (5) implies that $\lambda_n^{-1}d(S_n,A_n)\to 0$ in probability.

By the same reasoning as for D_1, for D_2 we get

$$\mathbf{P}\left(S_n\in\lambda_n U + A_n, \bigcap_{k=1}^n\{\xi_k\notin\lambda_n U^\varepsilon\}, D_2\right)\leqslant\mathbf{P}\big(d(\xi_2+\dots+\xi_n, A_n)>\varepsilon\lambda_n, D_2\big).$$

Since $d(S_n^{\neq 1}, A_n)\leqslant d(S_n^{\neq 1}, S_n) + d(S_n, A_n)$ and $d(S_n^{\neq 1}, S_n)\leqslant\|\xi_1\|$ the sub-invariance property yields that

$$\begin{aligned}\mathbf{P}\Big(d\big(S_n^{\neq 1}, A_n\big)>\varepsilon\lambda_n, \max_{i=1,\dots,n}\|\xi_i\|\leqslant\delta\lambda_n\Big) &\leqslant\mathbf{P}\Big(d(S_n, A_n)>\varepsilon\lambda_n/2, \max_{i=1,\dots,n}\|\xi_i\|\leqslant\delta\lambda_n\Big)\\ &\leqslant\mathbf{P}\Big(d(S_n, A_n)>\varepsilon\lambda_n/2, \max_{i=1,\dots,n}d(\xi_i, A_1)\leqslant 2\delta\lambda_n\Big)\end{aligned}$$

for sufficiently large n. The latter inequality holds since $d(\xi_i, A_1)\leqslant\|\xi_i\| + \|A_1\|$ and $\delta\lambda_n$ eventually exceeds $\|A_1\|$. Defining

$$\xi_i^\delta = \begin{cases}\xi_i, & d(\xi_i, A_1)\leqslant 2\delta\lambda_n,\\ \mathbf{e}, & d(\xi_i, A_1)>2\delta\lambda_n,\end{cases}\qquad i=1,\dots,n$$

and $S_n^\delta = \xi_1^\delta+\dots+\xi_n^\delta$, we see that it suffices to show that, for δ sufficiently small, $\gamma_n\mathbf{P}(d(S_n^\delta, A_n)>\varepsilon\lambda_n)\to 0$ as $n\to\infty$. For this, show that $\mathbf{E}d(S_n^\delta, A_n)\leqslant\varepsilon\lambda_n/2$ for sufficiently large n.

By the triangle inequality,

$$\lambda_n^{-1}\mathbf{E}d\big(S_n^\delta, A_n\big)\leqslant\lambda_n^{-1}\mathbf{E}d\big(S_n^\delta, S_n\big) + \lambda_n^{-1}\mathbf{E}d(S_n, A_n),$$

where the last term converges to zero as $n\to\infty$ by (5). To show that also $\lambda_n^{-1}\mathbf{E}d(S_n^\delta, S_n)\to 0$ as $n\to\infty$ for fixed δ, let $I(n)=\{1\leqslant i\leqslant n\colon\ d(\xi_i, A_1)>2\delta\lambda_n\}$, so that

$$d\big(S_n^\delta, S_n\big)\leqslant\Big\|\sum_{i\in I(n)}\xi_i\Big\|$$

by sub-invariance.

If $\alpha = 1$, $\lambda_n^{-1}\mathbf{E}\|\sum_{i\in I(n)}\xi_i\| \leqslant \lambda_n^{-1}n\mathbf{E}\|\xi_1\|$, which converges to zero. In case $\alpha > 1$, we have to show that

$$\lambda_n^{-1}n\mathbf{E}\big(\|\xi_1\|\mathbf{1}_{d(\xi_1,A_1)>2\delta\lambda_n}\big) \to 0 \quad \text{as } n\to\infty.$$

Since $d(\xi_1, A_1) \geqslant \|\xi_1\| - \|A_1\|$, it suffices that $\lambda_n^{-1}n\mathbf{E}(\|\xi_1\|\mathbf{1}_{\|\xi_1\|>3\delta\lambda_n})$ converges to 0, which follows from Corollary 2.2 and the growth rate of $\{\lambda_n\}$.

Now we assume that n is so large that $\mathbf{E}d(S_n^\delta, A_n) \leqslant \varepsilon\lambda_n/2$. Then

$$\mathbf{P}\big(d\big(S_n^\delta, A_n\big) > \varepsilon\lambda_n\big) \leqslant \mathbf{P}\big(d\big(S_n^\delta, A_n\big) - \mathbf{E}d\big(S_n^\delta, A_n\big) > \varepsilon\lambda_n/2\big).$$

Note that $d(S_n^\delta, A_n) - \mathbf{E}d(S_n^\delta, A_n)$ is almost surely equal to $\sum_{i=1}^n d_i$ with $d_i = \mathbf{E}(d(S_n^\delta, A_n) \mid \mathcal{A}_i) - \mathbf{E}(d(S_n^\delta, A_n) \mid \mathcal{A}_{i-1})$. Lemma 4.2 (applied for $\zeta_i = \xi_i^\delta$ and $z = A_n$) yields that $|d_i| \leqslant \|\xi_i^\delta\| + \mathbf{E}\|\xi_i^\delta\|$, which is almost surely smaller than $6\delta\lambda_n$ for sufficiently large n.

A martingale version of Bennett's inequality [7, Eq. (6.13)] for the martingale difference sequence $d_1, \dots, d_n$ yields that

$$\mathbf{P}\big(d\big(S_n^\delta, A_n\big) - \mathbf{E}d\big(S_n^\delta, A_n\big) > \varepsilon\lambda_n/2\big) \leqslant 2\exp\left\{\frac{\varepsilon}{12\delta} - \left(\frac{\varepsilon}{12\delta} + \frac{b}{144\delta^2\lambda_n^2}\right)\log\left(1 + \frac{6\delta\varepsilon\lambda_n^2}{b}\right)\right\}$$

for any $b \geqslant 4n\mathbf{E}\|\xi_1^\delta\|^2 \geqslant \sum_{i=1}^n \mathbf{E}d_i^2$.

In case $\alpha \geqslant 2$ and $\mathbf{E}\|\xi\|^2 < \infty$, we choose $b = 4n\mathbf{E}\|\xi\|^2$. In case $1 \leqslant \alpha \leqslant 2$ and $\mathbf{E}\|\xi\|^2$ is infinite, we choose $b = 8n(\|A_1\|^2 + \mathbf{E}\|\xi_1^\delta\|^2\mathbf{1}_{d(\xi_1,A_1)\leqslant 2\delta\lambda_n})$. Using the regular variation of $\|\xi\|$, we see that in both cases the growth conditions on $\{\lambda_n\}$ provide that for all $\delta > 0$ small enough,

$$\gamma_n\mathbf{P}\big(d\big(S_n^\delta, A_n\big) - \mathbf{E}d\big(S_n^\delta, A_n\big) > \varepsilon\lambda_n/2\big) \to 0 \quad \text{as } n\to\infty.$$

This establishes the upper bound.

For the lower bound, we let $U \in \mathcal{B}$ again be a μ-continuity set bounded away from $\mathbf{0}$ with nonempty interior and write

$$\begin{aligned}\mathbf{P}(S_n \in \lambda_n U + A_n) &\geqslant \mathbf{P}\left(S_n \in \lambda_n U + A_n, \bigcup_{i=1}^n\{\xi_i \in \lambda_n U^{-\varepsilon}\}\right)\\ &= \mathbf{P}\big(\cup_{i=1}^n\{\xi_i \in \lambda_n U^{-\varepsilon}\}\big) - \mathbf{P}\left(S_n \notin \lambda_n U + A_n, \bigcup_{i=1}^n\{\xi_i \in \lambda_n U^{-\varepsilon}\}\right)\\ &= I_1 - I_2.\end{aligned}$$

To show that $\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\gamma_n I_2 = 0$, first note that

$$\mathbf{P}\left(S_n \notin \lambda_n U + A_n, \bigcup_{i=1}^n\{\xi_i \in \lambda_n U^{-\varepsilon}\}\right) \leqslant n\mathbf{P}\big(S_n \notin \lambda_n U + A_n, \xi_1 \in \lambda_n U^{-\varepsilon}\big).$$

Now $S_n \notin \lambda_n U + A_n, \xi_1 \in \lambda_n U^{-\varepsilon}$ implies that $d(S_n^{\neq 1}, A_n) > \varepsilon\lambda_n$, by separately considering cases (A) and (B) as above. Then the same arguments as in the upper bound part of the proof apply.

It remains to show that $\limsup_{n\to\infty}\gamma_n I_1 \geqslant \mu(U^{-\varepsilon})$. By a Bonferroni argument,

$$\gamma_n I_1 \geqslant \frac{\mathbf{P}(\xi_1 \in \lambda_n U^{-\varepsilon})}{\mathbf{P}(\|\xi_1\| > \lambda_n)} - \frac{n-1}{2}\frac{\mathbf{P}(\xi_1 \in \lambda_n U^{-\varepsilon})^2}{\mathbf{P}(\|\xi_1\| > \lambda_n)}.$$

By choosing ε sufficiently small (such that $U^{-\varepsilon}$ is a μ-continuity set bounded away from $\mathbf{0}$), we may apply (3) and conclude that the lim sup of the positive summand converges to $\mu(U^{-\varepsilon})$. The upper limit of the second term is zero because $\lambda_n/a_n \to \infty$. Letting $\varepsilon \to 0$ establishes the lower bound. $\square$

A convex cone with metric d is said to be *isometrically embeddable* in a Banach space $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ if there exists a measurable map $I : \mathbb{K} \to \mathbb{B}$ such that $I(x + y) = I(x) + I(y)$ and $d(x, y) = \|I(x) - I(y)\|_{\mathbb{B}}$ for all $x, y \in \mathbb{K}$. If the second distributivity law in $\mathbb{K}$ holds, this is always possible and then I becomes a linear map, i.e. $I(ax) = aI(x)$ for all $a > 0$.

By [7, Prop. 9.11], a Banach space $\mathbb{B}$ of (Rademacher) type $p \in [1, 2]$ has the property that for every finite sequence $X_1, \dots, X_n$ of independent mean zero p-integrable Radon random variables in $\mathbb{B}$,

$$\mathbf{E}\left\|\sum_{i=1}^n X_i\right\|_{\mathbb{B}}^p \leqslant (2c)^p\sum_{i=1}^n \mathbf{E}\|X_i\|_{\mathbb{B}}^p. \tag{6}$$

If ξ is a random element in $\mathbb{K}$ with $\mathbf{E}\|\xi\| < \infty$, then $I(\xi)$ is a random element in $\mathbb{B}$ which is strongly integrable with expectation $\mathbf{E}I(\xi)$.

Corollary 4.3. *Let $\xi, \xi_1, \dots$ be i.i.d. regularly varying random elements with index $\alpha \geqslant 1$ and tail measure μ in a convex cone $\mathbb{K}$ which is isometrically embeddable by $I : \mathbb{K} \to \mathbb{B}$ in a separable Banach space $\mathbb{B}$ of type $\min(\alpha, 2)$. Assume that $\mathbf{E}\|\xi\|^{\min(\alpha,2)}$ is finite and $I(A_n) = n\mathbf{E}I(\xi)$, $n \geqslant 1$, for some cone elements $\{A_n, n \geqslant 1\}$. Then (5) holds, and so the statement of Theorem 4.1 follows.*

Proof. Note that the existence of the embedding implies that the metric d is invariant. If $\tilde{S}_n = I(S_n) - n\mathbf{E}I(\xi)$, then

$$\mathbf{E}d(S_n, A_n) = \mathbf{E}\|\tilde{S}_n\|_{\mathbb{B}}.$$

Since $\mathbb{B}$ is a Banach space of type $\min(\alpha, 2)$ and each summand $X_i = I(\xi_i) - \mathbf{E}I(\xi)$, $i = 1, \dots, n$, is centered and Radon by the separability of $\mathbb{B}$, it holds that

$$\mathbf{E}\|\tilde{S}_n\|_{\mathbb{B}}^{\min(\alpha,2)} \leqslant cn\mathbf{E}\|X_1\|_{\mathbb{B}}^{\min(\alpha,2)}$$

for some finite c by (6). Further

$$\sup_n n^{-1}\mathbf{E}\|\tilde{S}_n\|_{\mathbb{B}}^{\min(\alpha,2)} < \infty.$$

By Jensen's inequality,

$$\sup_n n^{-1/\min(\alpha,2)+\eta}\mathbf{E}\|\tilde{S}_n\|_{\mathbb{B}} \to 0$$

for any $\eta > 0$, so that (5) holds. □

5. Examples

Example 5.1 *(Half-line with maximum).* Let $\mathbb{K} = [0, \infty)$ with the semigroup operation $x + y = \max\{x, y\}$ and the usual multiplication, so that $S_n = \max\{\xi_1, \dots, \xi_n\}$. The metric $d(x, y) = |x - y|$ is homogeneous and sub-invariant, and $\mathbf{e} = \mathbf{0} = 0$. Regularly varying random elements are precisely the non-negative random variables with regularly varying right tail of index $\alpha > 0$, so their distributions are in the maximum domain of attraction of the Fréchet distribution and $a_n = n^{1/\alpha}\ell(n)$ for some slowly varying function ℓ. Thus, Theorem 3.1 applies.

Since the metric is not invariant, condition (A) in Theorem 4.1 is imposed, i.e. $A_n = 0$ for all n. Condition (5) requires that $\lambda_n^{-1}\mathbf{E}\max\{\xi_1, \dots, \xi_n\} \to 0$. If $\alpha = 1$, it follows from the law of large numbers, since $\max\{\xi_1, \dots, \xi_n\}/\lambda_n \leqslant (\xi_1 + \dots + \xi_n)/n^{1+\eta}$ and $\mathbf{E}\xi_1 < \infty$. In case $\alpha > 1$, we can use the fact that $\xi_1^{\alpha-\varepsilon}$ is integrable for any $\varepsilon > 0$. Since $\alpha - \varepsilon \geqslant 1$ for sufficiently small $\varepsilon > 0$, Jensen's inequality yields that

$$n^{-1-\eta'}\big(\mathbf{E}\max(\xi_1, \dots, \xi_n)\big)^{\alpha-\varepsilon} \to 0$$

for each $\eta' > 0$ and it remains to choose η' and $\varepsilon > 0$ such that $(1 + \eta')/(\alpha - \varepsilon) \leqslant 1/\alpha + \eta$.

Example 5.2 *(Compact sets in $\mathbb{R}^m$).* The cone of compact sets in the Euclidean space with the Minkowski addition metrized by the Hausdorff metric falls into the scheme of Theorem 3.1. This case is considered in [8].

Example 5.3 *(Compact convex sets in $\mathbb{R}^m$).* Let $\mathbb{K}$ consist of nonempty compact convex sets in $\mathbb{R}^m$, equipped with Minkowski addition and the usual scaling. The support function of the set X is denoted by $h_X(u)$ for u from the unit sphere $\mathbb{S}^{m-1}$ in $\mathbb{R}^m$. The Hausdorff distance between compact convex sets equals the uniform distance between their support functions, so it is possible to embed $\mathbb{K}$ into the Banach space of continuous functions on the unit sphere. This argument has been used in [9] to derive large deviation results for random convex compact sets with integrable norm, which is also a special case of Theorem 4.1.

However, it is possible to get rid of condition (5) by considering another metric for convex compact sets. For $p \in [1, \infty)$, define the distance between convex compact sets X and Y using the L_p-distance between their support functions as

$$d_p(X, Y) = \bigg(\int_{\mathbb{S}^{m-1}} \big|h_X(u) - h_Y(u)\big|^p \,du\bigg)^{1/p},$$

see also [12]. Note that d_p is homogeneous and invariant and the support function provides an isometric embedding of $\mathbb{K}$ into the space $\mathbb{B} = L_p(\mathbb{S}^{m-1})$. Since $\mathbb{B}$ is a separable Banach space of type $\min(p, 2)$, Corollary 4.3 applies and condition (5) is not needed.

Example 5.4 *(Functions with argument rescaling).* Let $\mathbb{K}$ consist of continuous functions $f : \mathbb{R}_+ \to \mathbb{R}$ such that $\int_0^\infty xf(x)^2\,dx < \infty$, i.e. $f \in L_2(R_+, \mu)$, where $\mu(dx) = x\,dx$. The addition is defined pointwisely (so $\mathbf{e}$ is the zero function) and the cone multiplication $\cdot$ is defined as

$$(a \cdot f)(x) = f\left(\frac{x}{a}\right)$$

for $a > 0$. The metric

$$d(f,g) = \left(\int_0^\infty x\big(f(x) - g(x)\big)^2 \, dx\right)^{1/2}$$

is invariant and homogeneous, and $\mathbb{K}$ is isometrically embeddable into the space $\mathbb{B} = L_2(\mathbb{R}_+, \mu)$ of type 2. Note however that the scaling in $\mathbb{K}$ differs from the scaling in $\mathbb{B}$. Thus, Theorem 3.1 applies and also (5) in Theorem 4.1 holds, see Corollary 4.3.

In order to construct an example of a regularly varying function in $\mathbb{K}'$, take any random function η from $\mathbb{K}$ with $\|\eta\| = 1$ a.s. and define

$$\xi(x) = (\zeta \cdot \eta)(x) = \eta\big(\zeta^{-1}x\big), \quad x \geqslant 0,$$

for a non-negative and independent of η random variable ζ with regularly varying tail with index $\alpha \geqslant 1$. Then $\|\xi\| = \zeta$ and $\|\xi\|^{-1} \cdot \xi = \eta$ a.s. and

$$n\mathbf{P}\big(\|\xi\|^{-1} \cdot \xi \in B, \|\xi\| > a_n r\big) = n\mathbf{P}(\eta \in B)\mathbf{P}(\zeta > a_n r) \to \sigma(B) r^{-\alpha}$$

as $n \to \infty$, where a_n is the normalizing sequence associated to ζ, so that ξ is indeed regularly varying in $\mathbb{K}$.

The condition $\mathbf{E}\|\xi\| < \infty$ in Theorem 4.1 means that $\mathbf{E}\zeta < \infty$. We define $A_n(x) = n\mathbf{E}\xi(x)$ for all x. Condition (5) holds in case $\mathbf{E}\zeta^{\min(\alpha,2)} < \infty$ by Corollary 4.3.

A number of further examples fall into the scope of the proved large deviation theorems. They include the cone of compact sets with the union operation and the Hausdorff metric, and the cone of integrable probability measures with the convolution operation and rescaling of the argument and the Wasserstein metric, see [3, Sec. 8] for these examples in view of the stability properties.

Acknowledgments

The authors are grateful to Zbyněk Pawlas for explaining some details of the proofs of large deviation results for Minkowski sums of random sets. The authors benefited from the advice of Sergei Foss concerning asymptotic behavior of random walks with heavy-tailed distributions.

References

[1] Z. Artstein, J.C. Hansen, Convexification in limit laws of random sets in Banach spaces, Ann. Probab. 13 (1985) 307–309.
[2] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular Variation, Cambridge University Press, Cambridge, 1989.
[3] Y. Davydov, I. Molchanov, S. Zuyev, Strictly stable distributions on convex cones, Electron. J. Probab. 13 (2008) 259–321.
[4] E. Giné, M.G. Hahn, J. Zinn, Limit theorems for random sets: application of probability in Banach space results, in: A. Beck, K. Jacobs (Eds.), Probability in Banach Spaces, IV, Oberwolfach, 1982, in: Lecture Notes in Math., vol. 990, Springer, Berlin, 1983, pp. 112–135.
[5] H. Hult, F. Lindskog, Regular variation for measures on metric spaces, Publ. Inst. Math. (Beograd) (N.S.) 80 (94) (2006) 121–140.
[6] H. Hult, F. Lindskog, T. Mikosch, G. Samorodnitsky, Functional large deviations for multivariate regularly varying random walks, Ann. Appl. Probab. 15 (2005) 2651–2680.
[7] M. Ledoux, M. Talagrand, Probability in Banach Spaces, Springer-Verlag, Berlin, 1991.
[8] T. Mikosch, Z. Pawlas, G. Samorodnitsky, A large deviation principle for Minkowski sums of heavy-tailed random compact convex sets with finite expectation, J. Appl. Probab. 48A (2011) 133–144.
[9] T. Mikosch, Z. Pawlas, G. Samorodnitsky, Large deviations for Minkowski sums of heavy-tailed generally non-convex random compact sets, Vestnik S.-Peterburg. Univ. Ser. 1. Mat. Mekh. Astronom. (2011) 70–78.
[10] I. Molchanov, Theory of Random Sets, Springer, London, 2005.
[11] V.V. Petrov, Limit Theorems of Probability Theory, Oxford University Press, Oxford, 1995.
[12] R.A. Vitale, L_p metrics for compact, convex sets, J. Approx. Theory 45 (1985) 280–287.