Applied Probability Trust (16 August 2010)

A LARGE DEVIATION PRINCIPLE FOR MINKOWSKI SUMS OF HEAVY-TAILED RANDOM COMPACT CONVEX SETS WITH FINITE EXPECTATION

THOMAS MIKOSCH,* *University of Copenhagen*
ZBYNĚK PAWLAS,** *Charles University in Prague*
GENNADY SAMORODNITSKY,*** *Cornell University*

Abstract

We prove large deviation results for Minkowski sums S_n of iid random compact sets where we assume that the summands have a regularly varying distribution and finite expectation. The main focus is on random convex compact sets. The results confirm the heavy-tailed large deviation heuristics: "large" values of the sum are essentially due to the "largest" summand. These results extend those in [8] for generally non-convex sets, where we assumed that the normalization of S_n grows faster than n.

Keywords: Minkowski sum, random compact set, large deviation, regularly varying distribution

2000 Mathematics Subject Classification: Primary 60F10
Secondary 52A22

1. Introduction

Preliminaries on random sets and Minkowski addition. The theory of random sets is summarized in the recent monograph [9]. For all definitions introduced below we refer to [9]. Let F be a separable Banach space with norm $\|\cdot\|$. For $A_1, A_2 \subseteq F$ and a real number λ, the Minkowski addition and scalar multiplication, respectively, are defined by

$$A_1 + A_2 = \{a_1 + a_2 : a_1 \in A_1, a_2 \in A_2\}, \quad \lambda A_1 = \{\lambda a_1 : a_1 \in A_1\}.$$

We denote by $\mathcal{K}(F)$ the class of all non-empty compact subsets of F. Note that this is not a vector space. However, it is well known that $\mathcal{K}(F)$ equipped with the Hausdorff distance

$$d(A_1, A_2) = \max\left\{\sup_{a_1\in A_1}\inf_{a_2\in A_2}\|a_1 - a_2\|, \sup_{a_2\in A_2}\inf_{a_1\in A_1}\|a_1 - a_2\|\right\}, \quad A_1, A_2 \in \mathcal{K}(F),$$

forms a complete separable metric space. The Hausdorff metric is subinvariant, i.e.,

$$d(A_1 + A, A_2 + A) \le d(A_1, A_2) \quad \text{for any } A_1, A_2, A \in \mathcal{K}(F).$$

* Postal address: Department of Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark, mikosch@math.ku.dk
** Postal address: Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, Praha 8, Czech Republic, pawlas@karlin.mff.cuni.cz
*** Postal address: School of Operations Research and Information Engineering, Cornell University, 220 Rhodes Hall, Ithaca, NY 14853, U.S.A., gennady@orie.cornell.edu

For any subset $\mathcal{U}$ of $\mathcal{K}(F)$, a real number λ and a set $A \in \mathcal{K}(F)$ we use the notation $\lambda\mathcal{U} = \{\lambda C : C \in \mathcal{U}\}$ and $\mathcal{U} + A = \{C + A : C \in \mathcal{U}\}$. For subsets $\mathcal{U}_1$, $\mathcal{U}_2$ of $\mathcal{K}(F)$ we denote $d(\mathcal{U}_1, \mathcal{U}_2) = \inf_{A_1 \in \mathcal{U}_1, A_2 \in \mathcal{U}_2} d(A_1, A_2)$.

A random compact set X in F is a Borel measurable function from an abstract probability space $(\Omega, \mathcal{F}, \mathrm{P})$ into $\mathcal{K}(F)$. Since addition and scalar multiplication are defined for random compact sets it is natural to study the strong law of large numbers, the central limit theorem, large deviations, etc., for sequences of such random sets; see Chapter 3 in [9] for an overview of results obtained until 2005. A general Cramér-type large deviation result for Minkowski sums of iid random compact sets was proved in [2]. Cramér-type large deviations require exponential moments of the summands. If such moments do not exist, then we are dealing with heavy-tailed random elements. Large deviations results for sums of heavy-tailed random elements significantly differ from Cramér-type results. In this case it is typical that only the largest summand determines the large deviation behavior; see the classical results by A. Nagaev [10] for sums of iid random variables; cf. [11, 6]. It is the aim of this paper to prove large deviation results for sums of *heavy-tailed random compact sets.* In what follows, we make this notion precise by introducing regularly varying random sets.

Regularly varying random sets. A special element of $\mathcal{K}(F)$ is $A_0 = \{0\}$. In what follows, we say that $\mathcal{U} \subseteq \mathcal{K}(F)$ is bounded away from A_0 if $A_0 \notin \mathrm{cl}\,\mathcal{U}$, where $\mathrm{cl}\,\mathcal{U}$ stands for the closure of $\mathcal{U}$. We consider the subspace $\mathcal{K}_0(F) = \mathcal{K}(F) \setminus \{A_0\}$, which is a separable metric space in the relative topology. For any Borel set $\mathcal{U} \subseteq \mathcal{K}_0(F)$ and $\varepsilon > 0$, we write

$$\mathcal{U}^\varepsilon = \{A \in \mathcal{K}_0(F) : d(A, C) \le \varepsilon \text{ for some } C \in \mathcal{U}\}.$$

Furthermore, we define the norm $\|A\| = d(A, A_0) = \sup\{\|a\| : a \in A\}$ for $A \in \mathcal{K}(F)$, and denote $\mathcal{B}_r = \{A \in \mathcal{K}(F) : \|A\| \le r\}$. Let $M_0\big(\mathcal{K}_0(F)\big)$ be the collection of Borel measures on $\mathcal{K}_0(F)$ whose restriction to $\mathcal{K}(F) \setminus \mathcal{B}_r$ is finite for each $r > 0$. Let $\mathcal{C}_0$ denote the class of real-valued, bounded and continuous functions f on $\mathcal{K}_0(F)$ such that for each f there exists $r > 0$ and f vanishes on $\mathcal{B}_r$. The convergence $\mu_n \longrightarrow \mu$ in $M_0\big(\mathcal{K}_0(F)\big)$ is defined to mean the convergence $\int f \,\mathrm{d}\mu_n \longrightarrow \int f \,\mathrm{d}\mu$ for all $f \in \mathcal{C}_0$. By the portmanteau theorem ([5], Theorem 2.4), $\mu_n \longrightarrow \mu$ in $M_0\big(\mathcal{K}_0(F)\big)$ if and only if $\mu_n(\mathcal{U}) \longrightarrow \mu(\mathcal{U})$ for all Borel sets $\mathcal{U} \subseteq \mathcal{K}(F)$ which are bounded away from A_0 and satisfy $\mu(\partial\mathcal{U}) = 0$, where $\partial\mathcal{U}$ is the boundary of $\mathcal{U}$.

Following [5], for the general case of random elements with values in a separable metric space, a random compact set X is *regularly varying* if there exist a non-null measure $\mu \in M_0\big(\mathcal{K}_0(F)\big)$ and a sequence $\{a_n\}_{n \ge 1}$ of positive numbers such that

$$n\mathrm{P}(X \in a_n \cdot) \longrightarrow \mu(\cdot) \quad \text{in } M_0\big(\mathcal{K}_0(F)\big). \tag{1}$$

The *tail measure* μ necessarily has the property $\mu(\lambda\mathcal{U}) = \lambda^{-\alpha}\mu(\mathcal{U})$ for some $\alpha > 0$, all Borel sets $\mathcal{U}$ in $\mathcal{K}_0(F)$ and all $\lambda > 0$. We then also refer to *regular variation of X with index α and write for short* $X \in \mathrm{RV}(\alpha, \mu)$. From the definition of regular variation of X we get ([5], Theorem 3.1)

$$\left[\mathrm{P}(X \in t(\mathcal{K}(F) \setminus \mathcal{B}_1))\right]^{-1} \mathrm{P}(X \in t\cdot) \longrightarrow c\mu(\cdot) \quad \text{in } M_0\big(\mathcal{K}_0(F)\big) \text{ as } t \to \infty, \tag{2}$$

for some $c > 0$. The sequence $\{a_n\}_{n \ge 1}$ will always be chosen such that $n\mathrm{P}(X \in a_n(\mathcal{K}(F) \setminus \mathcal{B}_1)) \longrightarrow 1$. With this choice of $\{a_n\}_{n \ge 1}$, it follows that $c = 1$ in (2).

An important closed subset of $\mathcal{K}(F)$ is the family of non-empty *compact convex subsets* of F, denoted by $\operatorname{co}\mathcal{K}(F)$. Denote the topological dual of F by F^* and the unit ball of F^* by B^*, it is endowed with the weak-$*$ topology w^*. The *support function* h_A of a compact convex $A \in \operatorname{co}\mathcal{K}(F)$ is defined by (see [9])

$$h_A(u) = \sup\{u(x) : x \in A\}, \quad u \in B^*.$$

Since A is compact, $h_A(u) < \infty$ for all $u \in B^*$. The support function h_A is sublinear, i.e., it is subadditive ($h_A(u+v) \le h_A(u) + h_A(v)$ for all $u, v \in B^*$ with $u + v \in B^*$) and positively homogeneous ($h_A(cu) = ch_A(u)$ for all $c > 0$, $u \in B^*$ with $cu \in B^*$). Let $\mathcal{C}(B^*, w^*)$ be the set of continuous functions from B^* (endowed with the weak-$*$ topology) to $\mathbb{R}$ and consider the uniform norm $\|f\|_\infty = \sup_{u \in B^*} |f(u)|$, $f \in \mathcal{C}(B^*, w^*)$. The map $h : \operatorname{co}\mathcal{K}(F) \to \mathcal{C}(B^*, w^*)$ has the following properties

$$h_{A_1 + A_2} = h_{A_1} + h_{A_2}, \quad h_{\lambda A_1} = \lambda h_{A_1}, \quad A_1, A_2 \in \operatorname{co}\mathcal{K}(F), \quad \lambda \ge 0,$$

which make it possible to convert the Minkowski sums and scalar multiplication, respectively, of convex sets into the arithmetic sums and scalar multiplication of the corresponding support functions. Furthermore,

$$d(A_1, A_2) = \|h_{A_1} - h_{A_2}\|_\infty. \tag{3}$$

Hence, the support function provides an isometric embedding of $\operatorname{co}\mathcal{K}(F)$ into $\mathcal{C}(B^*, w^*)$ with the uniform norm. If $\mathcal{G} = h\big(\operatorname{co}\mathcal{K}(F)\big)$, then $\mathcal{G}$ is a closed convex cone in $\mathcal{C}(B^*, w^*)$, and h is an isometry between $\operatorname{co}\mathcal{K}(F)$ and $\mathcal{G}$.

A random compact convex set X is a Borel measurable function from a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ into $\operatorname{co}\mathcal{K}(F)$, which we endow with the relative topology inherited from $\mathcal{K}(F)$. The support function of a random compact convex set is, clearly, a $\mathcal{C}(B^*, w^*)$-valued random variable taking values in $\mathcal{G}$.

The definition of a regularly varying random compact convex set parallels that of a regularly varying random compact set above, and we are using the same notation: a random compact convex set X is *regularly varying* if there exist a non-zero measure $\mu \in M_0\big(\operatorname{co}\mathcal{K}_0(F)\big)$ and a sequence $\{a_n\}_{n \ge 1}$ of positive numbers such that

$$n\mathrm{P}(X \in a_n \cdot) \longrightarrow \mu(\cdot) \ \text{ in } M_0\big(\operatorname{co}\mathcal{K}_0(F)\big). \tag{4}$$

Once again, the tail measure μ necessarily scales, leading to the notion of the index of regular variation.

The following lemma is elementary.

Lemma 1. (i) *A random compact convex set X is regularly varying in $\operatorname{co}\mathcal{K}(F)$ if and only if its support function h_X is regularly varying in $\mathcal{C}(B^*, w^*)$. Specifically, if* (4) *holds for some sequence $\{a_n\}$, then for the same sequence we have*

$$n\mathrm{P}(h_X \in a_n \cdot) \longrightarrow \nu(\cdot) \ \text{ in } M_0\big(\mathcal{C}(B^*, w^*)\big), \tag{5}$$

where $\nu = \mu \circ h_X^{-1}$ (the "special element" of $\mathcal{C}(B^*, w^*)$ is, of course, the zero function). Conversely, if (5) *holds, then* (4) *holds as well with $\mu = \nu \circ h_X$. In particular, the indices of regular variation of X and h_X are the same.*

(ii) *If a random compact set X is regularly varying in $\mathcal{K}(F)$ then its convex hull* $\operatorname{co} X$ *is a random compact convex set, that is regularly varying in* $\operatorname{co}\mathcal{K}(F)$. *Specifically, if* (1) *holds, then so does* (4)*, with the tail measure replaced by the image of the tail measure from* (1) *under the map $A \mapsto \operatorname{co} A$ from $\mathcal{K}(F)$ to $\operatorname{co}\mathcal{K}(F)$. In particular, X and $\operatorname{co} X$ have the same indices of regular variation.*

Proof. Since isometry implies continuity, the support function is homogeneous of order 1, and assigns to the "special set" $\{0\}$ the "special element", the zero function, the statement of part (i) of the lemma follows from the mapping theorem (Theorem 2.5 in [5]). For part (ii) note that the map $A \mapsto \operatorname{co} A$ from $\mathcal{K}(F)$ to $\operatorname{co}\mathcal{K}(F)$ is a contraction in the Hausdorff distance, hence is continuous. It is also homogeneous of order 1. Since the "special set" $\{0\}$ is already convex, the statement follows once again from the mapping theorem.

Organization of the paper. In Section 2 we consider various examples of regularly varying compact random sets. In Section 3 we prove large deviation results for Minkowski sums S_n of iid regularly varying random compact sets. To the best of our knowledge, such results are not available in the literature; they parallel those proved by A. and S. Nagaev [10, 11] for sums of iid random variables. The case of general random compact sets is treated in [8]. The price one has to pay for this generality is that the normalizations λ_n of the sums S_n have to exceed the level n. The situation with milder normalizations considered in the present paper is much more delicate. Our main result here assumes that the random sums are convex, but we include partial results in the non-convex case as well.

2. Examples of regularly varying random sets

Simple examples of regularly varying random sets can be constructed from iid F-valued random elements $\xi_1, \dots, \xi_k$, $k \ge 2$, which are regularly varying with index $\alpha > 0$ and tail measure ν. The following three examples of random compact sets are distinct but the tail measures turn out to be the same. For the proofs we refer to [8].

Example 1. The *convex hull* $X = \operatorname{co}\{\xi_1, \dots, \xi_k\} \in \mathrm{RV}(\alpha, k\,\nu \circ T^{-1})$, where $T : F \to \operatorname{co}\mathcal{K}(F)$ is defined by the relation $T(x) = [0, x]$, and for $x, y \in F$, $[x, y]$ is the *random segment* with endpoints x, y. The *random zonotope* $X' = \sum_{i=1}^k [0, \xi_i] \in \mathrm{RV}(\alpha, k\,\nu \circ T^{-1})$. The random set $X = \bigcup_{1 \le i < j \le k} [\xi_i, \xi_j]$ is a compact, but generally non-convex, subset of F. The map $g : (z_1, \dots, z_k) \mapsto \cup_{1 \le i < j \le k} [z_i, z_j]$ from F^k to $\mathcal{K}(F)$ is continuous, homogeneous of order 1, and maps the zero point in F^k to A_0, which is now viewed as the "special element" of $\mathcal{K}(F)$. The continuous mapping argument used in the examples of [8] shows that $X \in \mathrm{RV}(\alpha, k\,\nu \circ T^{-1})$, where now we view $T(x) = [0, x]$ as a map from F to $\mathcal{K}(F)$. Note that the tail measure is supported by convex sets as in Theorem 2 below.

Another example of regularly varying random set is a sojourn set of multidimensional Brownian motion.

Example 2. (*Sojourn set.*) For $k \ge 3$, let $\{\mathbf{W}_t = (W_t^{(i)})_{i=1,\dots,k},\ t \in \mathbb{R}^+\}$ be a standard Brownian motion, i.e., the W_i's are independent standard Brownian motions in $\mathbb{R}$. Then $\{\|\mathbf{W}_t\|,\ t \in \mathbb{R}^+\}$ is a Bessel process of order k.

Consider the random set $X = \{t \in \mathbb{R}^+ : \|\mathbf{W}_t\| \le 1\}$. We claim that this set is regularly varying with index $\alpha = (k-2)/2$. To see this, let us define

$$M = \sup\{t : t \in X\} = \sup\{t \in \mathbb{R}^+ : \|\mathbf{W}_t\| \le 1\}.$$

It follows from the last part of Exercise 1.18, p. 450 in [12] that M^{-1} is χ^2_{k-2}-distributed. Therefore

$$\mathrm{P}(M > t) \sim \frac{1}{2^{(k-2)/2}\Gamma(k/2)}\, t^{-(k-2)/2} =: \nu(t,\infty), \quad t \to \infty.$$

The map $T : \mathbb{R}^+ \to \mathcal{K}(\mathbb{R})$ given by $T(x) = \{0, x\}$ is continuous, homogeneous of order 1 (and, hence, maps the zero point into A_0). Then the set $Y = \{0, M\} \subseteq X$ is regularly varying with index $\alpha = (k-2)/2$ and, with the measure ν on $\mathbb{R}^+$ defined above,

$$n\,\mathrm{P}\big(Y \in n^{2/(k-2)}\cdot\big) \longrightarrow \nu \circ T^{-1}(\cdot) \quad \text{in } M_0\big(\mathcal{K}_0(\mathbb{R})\big).$$

This relation remains valid with Y replaced by X, once one can show that for any $\varepsilon > 0$,

$$n\,\mathrm{P}\big(d(X,Y) > n^{2/(k-2)}\varepsilon\big) \longrightarrow 0. \tag{6}$$

Since $Y \subseteq X$, we have, with $T = \inf\{t > n^{2/(k-2)}\varepsilon : \|\mathbf{W}_t\| \le 1\} \in [n^{2/(k-2)}\varepsilon, \infty]$,

$$\begin{aligned}
&\mathrm{P}\big(d(X,Y) > n^{2/(k-2)}\varepsilon\big)\\
&\quad = \mathrm{P}\big(X \text{ contains a point separated by more that } n^{2/(k-2)}\varepsilon \text{ from both zero and } M\big)\\
&\quad \le \mathrm{P}\big(M - T > n^{2/(k-2)}\varepsilon\big).
\end{aligned}$$

Note that T is a stopping time, the process $\{\|\mathbf{W}_t\|,\, t \in \mathbb{R}^+\}$ is a Feller process, hence strongly Markov; see [12], p. 446. Therefore,

$$\mathrm{P}\big(M - T > n^{2/(k-2)}\varepsilon\big) = \mathbb{E}\big(\mathbf{1}_{\{T<\infty\}} \mathrm{P}_{\|\mathbf{W}_T\|}(M > n^{2/(k-2)}\varepsilon)\big).$$

At time $T < \infty$, the Brownian motion is inside the closed unit ball, hence returning to that closed unit ball at a later point means being within a distance of at most 2 of the initial point. Therefore, on the event $\{T < \infty\}$, with probability 1,

$$\begin{aligned}
\mathrm{P}_{\|\mathbf{W}_T\|}\big(M > n^{2/(k-2)}\varepsilon\big) &\le \mathrm{P}\big(\sup\{t \in \mathbb{R}^+ : \|\mathbf{W}_t\| \le 2\} > n^{2/(k-2)}\varepsilon\big)\\
&= \mathrm{P}\big(M > n^{2/(k-2)}\varepsilon/4\big) \le cn^{-1}
\end{aligned}$$

for n large enough, for some $c > 0$. We conclude that for large n,

$$\begin{aligned}
\mathrm{P}\big(d(X,Y) > n^{2/(k-2)}\varepsilon\big) &\le cn^{-1}\mathrm{P}(T < \infty)\\
&\le cn^{-1}\mathrm{P}(M \ge n^{2/(k-2)}\varepsilon) = O(n^{-2})
\end{aligned}$$

as $n \to \infty$, thus proving (6).

The random set of this example can be naturally embedded into the space $\mathbb{R}^k$ by defining

$$X_1 = \{t\mathbf{W}_t : t \in \mathbb{R}^+ : \|\mathbf{W}_t\| \le 1\} \subseteq \mathbb{R}^k.$$

It follows from what we already know about the set X, that X_1 is regularly varying, with the tail measure

$$\mu_1 = (\nu \times H) \circ T_1^{-1},$$

where H is the normalized Haar measure on the unit sphere $\mathbb{S}^{k-1}$, and $T_1 : \mathbb{R}^+ \times \mathbb{S}^{k-1} \to \mathcal{K}(\mathbb{R}^k)$ given by $T_1(x,s) = [0, sx]$.

3. Large deviations in the presence of expectation

In [8] we considered large deviations for the sums $S_n = X_1 + \cdots + X_n$ of iid regularly varying random compact sets X_i, $i = 1, 2, \ldots$, which were not necessarily convex. However, we had to assume that the scaling sequence $\{\lambda_n\}$ of $\{S_n\}$ had to grow faster than n. This is not a very natural condition if the index of regular variation $\alpha > 1$. In Theorem 1 below we will relax the conditions on $\{\lambda_n\}$ by assuming that we can define the expectation of a random set, but we will restrict ourselves to compact convex sets.

Let X be a random compact set in F. Following [9], a random element $\xi \in F$ is a *selection of* X if $\xi \in X$ a.s. and if $\mathbb{E}\|\xi\| < \infty$, ξ is an integrable selection. The *selection expectation* of X is defined as $\mathbb{E}X = \mathrm{cl}\{\mathbb{E}\xi : \xi \text{ is an integrable selection of } X\}$. The selection expectation of a random compact convex set is defined in the same way. The selection expectation is necessarily a convex set (assuming sufficient richness of the underlying probability space), even if X itself is not convex. If X is a random compact convex set and $\mathbb{E}\|X\| < \infty$, then the selection expectation of X is the unique convex compact set $\mathbb{E}X$ satisfying $\mathbb{E}h_X(u) = h_{\mathbb{E}X}(u)$ for all $u \in B^*$; see [9], Theorem 2.1.22.

Theorem 1. *Let $\{X_n\}_{n\ge1}$ be an iid sequence of random compact convex sets, regularly varying with index $\alpha \ge 1$ and tail measure $\mu \in M_0\big(\mathrm{co}\,\mathcal{K}_0(F)\big)$. Assume that $\mathbb{E}\|X_1\| < \infty$. Consider a sequence $\{\lambda_n\}_{n\ge1}$ such that $\lambda_n \nearrow \infty$,*

$$\lambda_n^{-1} d(S_n, n\mathbb{E}X_1) \xrightarrow{\mathrm{P}} 0, \tag{7}$$

$$\lambda_n^{-1} \mathbb{E} d(S_n, n\mathbb{E}X_1) \longrightarrow 0, \tag{8}$$

and for some $\eta > 0$, (i) *$\lambda_n / n^{1/2+\eta} \longrightarrow \infty$ if $\alpha \ge 2$, and* (ii) *$\lambda_n / n^{1/\alpha+\eta} \longrightarrow \infty$ if $1 \le \alpha < 2$. Then, with $\gamma_n = [n\mathrm{P}(\|X_1\| > \lambda_n)]^{-1}$,*

$$\gamma_n \mathrm{P}\big(S_n \in \lambda_n \cdot + n\mathbb{E}X_1\big) \longrightarrow \mu(\cdot) \quad \textit{in } M_0\big(\mathrm{co}\,\mathcal{K}_0(F)\big).$$

Remark 1. Notice that the assumptions of the theorem imply that $\lambda_n / a_n \longrightarrow \infty$. Regarding the assumptions on the random set, we start by observing that the condition $\mathbb{E}\|X_1\| < \infty$ is automatic if $\alpha > 1$. Further, condition (7) can be easily verified if the random sets satisfy the central limit theorem. For example, if $d(S_n, n\mathbb{E}X_1)/\sqrt{n}$ converges in distribution (as it does when a Gaussian central limit theorem holds) and $\lambda_n/\sqrt{n} \longrightarrow \infty$ then (7), obviously, holds. This Gaussian central limit theorem requires $\alpha > 2$, and assumption (i) of the theorem already implies that $\lambda_n/\sqrt{n} \longrightarrow \infty$. Alternatively, if $d(S_n, n\mathbb{E}X_1)/a_n$ converges in distribution in the context of an α-stable central limit theorem, $1 < \alpha < 2$, then (7) also holds since $\lambda_n/a_n \longrightarrow \infty$. Sufficient conditions for the central limit theorem can be found in [4, 9]. If the Gaussian central limit theorem is satisfied, then condition (8) follows by the isometric embedding (3) and Corollary 10.2 in [7].

The usual choice of the scaling sequence is, of course, $\lambda_n = n$. Then condition (7) follows from the strong law of large numbers which is satisfied for any sequence $\{X_n\}$ of iid random compact convex sets in F by virtue of [4], Theorem 3.1 (the law of large numbers for random compact sets in $\mathbb{R}^d$ was established even earlier, by [1]). Since the law of large numbers in a separable Banach space implies the L^1 convergence, the isometric embedding (3) implies (8) as well. Conditions of the type (7), (8) and growth conditions on $\{\lambda_n\}$ similar to those used in Theorem 1 have been widely used in simple non-set-valued large deviation contexts; see e.g. [11, 3, 6].

Proof of Theorem 1. Let $\mathcal{U} \subseteq \operatorname{co}\mathcal{K}_0(F)$ be a μ-continuity set, bounded away from A_0. We will show that $\gamma_n \mathrm{P}(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1) \longrightarrow \mu(\mathcal{U})$. We start with an upper bound:

$$\begin{aligned}\mathrm{P}\big(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1\big) &= \mathrm{P}\big(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1, \cup_{i=1}^n\{X_i \in \lambda_n\mathcal{U}^\varepsilon\}\big)\\ &\quad + \mathrm{P}\big(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1, \cap_{i=1}^n\{X_i \notin \lambda_n\mathcal{U}^\varepsilon\}\big)\\ &\le n\mathrm{P}\big(X_1 \in \lambda_n\mathcal{U}^\varepsilon\big) + \mathrm{P}\big(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1, \cap_{i=1}^n\{X_i \notin \lambda_n\mathcal{U}^\varepsilon\}\big)\\ &:= I_1 + I_2.\end{aligned}$$

It follows from (2) and μ-continuity of $\mathcal{U}$ that

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

In order to show $\gamma_n I_2 \longrightarrow 0$ we use the isometric embedding $h : \operatorname{co}\mathcal{K}(F) \to \mathcal{C}(B^*, w^*)$ given by the support function. In the new language we have

$$\gamma_n I_2 = \gamma_n \mathrm{P}\big(\sum_{i=1}^n (h_{X_i} - \mathbb{E}h_{X_i}) \in \lambda_n\mathcal{V}, \cap_{i=1}^n\{h_{X_i} \notin \lambda_n\mathcal{V}^\varepsilon\}\big),$$

where $\mathcal{V} = h(\mathcal{U})$ is bounded away from the zero function. Note also that $\gamma_n = [n\mathrm{P}(\|h_{X_1}\|_\infty > \lambda_n)]^{-1}$. Let $Y_i = h_{X_i} - \mathbb{E}h_{X_i}$ and $\tilde{S}_n = \sum_{i=1}^n Y_i$. Then

$$I_2 \le \mathrm{P}\big(\cap_{i=1}^n\{\|\tilde{S}_n - h_{X_i}\|_\infty > \varepsilon\lambda_n\}\big).$$

For $0 < \delta \le \varepsilon/3$ consider the following disjoint partition of Ω: $B_1 = \cup_{i=1}^n\{\|h_{X_i}\|_\infty > \delta\lambda_n\}$ and $B_2 = \big\{\max_{i=1,\dots,n}\|h_{X_i}\|_\infty \le \delta\lambda_n\big\}$. Then

$$\begin{aligned}&\mathrm{P}\big(\cap_{i=1}^n\{\|\tilde{S}_n - h_{X_i}\|_\infty > \varepsilon\lambda_n\} \cap B_1\big)\\ &\quad\le \sum_{k=1}^n \mathrm{P}\big(\|\tilde{S}_n - h_{X_k}\|_\infty > \varepsilon\lambda_n, \|h_{X_k}\|_\infty > \delta\lambda_n\big)\\ &\quad\le \mathrm{P}\big(\|\tilde{S}_{n-1} - \mathbb{E}h_{X_1}\|_\infty > \varepsilon\lambda_n\big)\big[n\mathrm{P}(\|h_{X_1}\|_\infty > \delta\lambda_n)\big]\\ &\quad\le \mathrm{P}\big(\|\tilde{S}_{n-1}\|_\infty > \varepsilon\lambda_n - \|\mathbb{E}h_{X_1}\|_\infty\big)\big[n\mathrm{P}(\|h_{X_1}\|_\infty > \delta\lambda_n)\big].\end{aligned}$$

By (7), the first term in the right hand side above vanishes as $n \to \infty$, while by Lemma 1(i) the second term, when multiplied by γ_n, converges to a finite limit. As regards B_2, we denote $Y_i^\delta = Y_i\mathbf{1}_{\{\|Y_i\|_\infty \le 2\delta\lambda_n\}}$ and $\tilde{S}_n^\delta = \sum_{i=1}^n Y_i^\delta$. As $\delta \le \varepsilon/3$, we have for sufficiently large n,

$$\mathrm{P}\big(\cap_{i=1}^n\{\|\tilde{S}_n - h_{X_i}\|_\infty > \varepsilon\lambda_n\} \cap B_2\big) \le \mathrm{P}\big(\|\tilde{S}_{n-1}\|_\infty > \varepsilon\lambda_n/2, \max_{i=1,\dots,n-1}\|Y_i\|_\infty \le 2\delta\lambda_n\big).$$

The required upper bound in the theorem will follow once we can show that for $\delta > 0$ small enough, $\gamma_n\mathrm{P}\big(\|\tilde{S}_n^\delta\|_\infty > \varepsilon\lambda_n\big) \longrightarrow 0$. Observe that

$$\mathrm{P}\big(\|\tilde{S}_n^\delta\|_\infty > \varepsilon\lambda_n\big) \le \mathrm{P}\big(\|\tilde{S}_n^\delta\|_\infty - \mathbb{E}\|\tilde{S}_n^\delta\|_\infty > \varepsilon\lambda_n/2\big) + \mathbf{1}_{\{\mathbb{E}\|\tilde{S}_n^\delta\|_\infty > \varepsilon\lambda_n/2\}}. \tag{9}$$

Applying inequality (6.13) from [7], we see that for any $b \ge n\mathbb{E}\|Y_1^\delta\|_\infty^2$,

$$\mathrm{P}\left(\|\tilde{S}_n^\delta\|_\infty - \mathbb{E}\|\tilde{S}_n^\delta\|_\infty > \varepsilon\lambda_n/2\right) \le 2\exp\left[\frac{\varepsilon}{8\delta} - \left(\frac{\varepsilon}{8\delta} + \frac{b}{16\delta^2\lambda_n^2}\right)\log\left(1 + \frac{2\varepsilon\delta\lambda_n^2}{b}\right)\right].$$

If $\alpha \geq 2$ and $\mathbb{E}\|h_{X_1}\|_\infty^2 < \infty$, we take $b = n\mathbb{E}\|Y_1\|_\infty^2$. If $\alpha \leq 2$ and $\mathbb{E}\|h_{X_1}\|_\infty^2 = \infty$, then for large n we take $b = 2n\left(\|\mathbb{E}h_{X_1}\|_\infty^2 + \mathbb{E}\|h_{X_1}\|_\infty^2 \mathbf{1}_{\{\|h_{X_1}\|_\infty \leq 3\delta\lambda_n\}}\right)$. If $\alpha = 2$, by Karamata's theorem, we have $b = nl(n)$ for a a slowly varying (at infinity) function l. If $1 \leq \alpha < 2$, then by Karamata's theorem our choice results in $b \sim cn\lambda_n^2 \mathrm{P}\big(\|h_{X_1}\|_\infty > 3\delta\lambda_n\big)$ as $n \to \infty$ for some $c > 0$. In all three cases for all $\delta > 0$ small enough,

$$\gamma_n \mathrm{P}\left(\|\tilde{S}_n^\delta\|_\infty - \mathbb{E}\|\tilde{S}_n^\delta\|_\infty > \varepsilon\lambda_n/2\right) \longrightarrow 0.$$

The second term of (9) is zero for sufficiently large n. Indeed, if $\alpha = 1$ and $\mathbb{E}\|h_{X_1}\|_\infty < \infty$, then the choice of λ_n trivially shows that $\mathbb{E}\|\tilde{S}_n^\delta\|_\infty/\lambda_n \longrightarrow 0$. For $\alpha > 1$ we write

$$\mathbb{E}\|\tilde{S}_n^\delta\|_\infty/\lambda_n \leq \mathbb{E}\|\tilde{S}_n\|_\infty/\lambda_n + \mathbb{E}\|\tilde{S}_n - \tilde{S}_n^\delta\|_\infty/\lambda_n,$$

and observe that $\mathbb{E}\|\tilde{S}_n\|_\infty/\lambda_n \longrightarrow 0$ by the assumption (8). To see that

$$\mathbb{E}\|\tilde{S}_n - \tilde{S}_n^\delta\|_\infty/\lambda_n \longrightarrow 0$$

notice that, by Karamata's theorem and the choice of λ_n,

$$\lambda_n^{-1} n\mathbb{E}\|Y_1\|_\infty \mathbf{1}_{\{\|Y_1\|_\infty > 2\delta\lambda_n\}} \sim cn\mathrm{P}(\|h_{X_1}\|_\infty > 2\delta\lambda_n) \longrightarrow 0.$$

We conclude that, for any μ-continuity set $\mathcal{U}$ bounded away from A_0,

$$\limsup_{n\to\infty} \gamma_n \mathrm{P}(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1) \leq \mu(\mathcal{U}^\varepsilon) \underset{\varepsilon \searrow 0}{\longrightarrow} \mu(\mathcal{U}),$$

where the limit is taken along such $\varepsilon > 0$ that $\mathcal{U}^\varepsilon$ is a continuity set.

To prove the corresponding lower bound, write for $\mathcal{U}$ as above

$$\begin{aligned} \mathrm{P}\big(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1\big) &\geq \mathrm{P}\big(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1, \cup_{i=1}^n \{X_i \in \lambda_n\mathcal{U}^{-\varepsilon}\}\big) \\ &\geq \mathrm{P}\big(\cup_{i=1}^n \{X_i \in \lambda_n\mathcal{U}^{-\varepsilon}\}\big) \\ &\quad - \mathrm{P}\big(S_n \notin \lambda_n\mathcal{U} + n\mathbb{E}X_1, \cup_{i=1}^n \{X_i \in \lambda_n\mathcal{U}^{-\varepsilon}\}\big) := I_1 - I_2. \end{aligned}$$

The same argument as in the proof of the upper bound shows that $\gamma_n I_2 \longrightarrow 0$ as $n \to \infty$. Furthermore, a Bonferroni argument shows that

$$\gamma_n\, I_1 \geq n\gamma_n \mathrm{P}\big(X_1 \in \lambda_n\mathcal{U}^{-\varepsilon}\big) - \gamma_n\, 0.5\, n(n-1)[\mathrm{P}(X_1 \in \lambda_n\mathcal{U}^{-\varepsilon})]^2.$$

By the choice of λ_n, for $\varepsilon > 0$ so small that $\mathcal{U}^{-\varepsilon}$ is bounded away from A_0 and a μ-continuity set, $\liminf_{n\to\infty} \gamma_n I_1 \geq \mu\big(\mathcal{U}^{-\varepsilon}\big)$. Letting $\varepsilon \to 0$ establishes the required lower bound, completing the proof.

The statement of Theorem 1 is a bit unusual in the context of large deviation results: while $\mathrm{P}(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1)$, $\mathcal{U}$ a measurable subset of $\mathrm{co}\,\mathcal{K}_0(F)$, is, in fact, a probability measure on $\mathrm{co}\,\mathcal{K}_0(F)$, the sets $\lambda_n\mathcal{U} + n\mathbb{E}X_1$ do not cover all measurable subsets of $\mathrm{co}\,\mathcal{K}_0(F)$, except in the trivial case $X_1 = A_0$ a.s. This is especially inconvenient in the case of linear scaling $\lambda_n = an$ for some $a > 0$, when the statement of Theorem 1 can be written in the form $\gamma_n \mathrm{P}((an)^{-1}S_n \in \cdot + a^{-1}\mathbb{E}X_1) \longrightarrow \mu(\cdot)$ in $M_0\big(\mathrm{co}\,\mathcal{K}_0(F)\big)$ which leaves unanswered the obvious question of how the law of $(an)^{-1}S_n$ behaves on sets that are not in $\mathrm{co}\,\mathcal{K}_0(F) + a^{-1}\mathbb{E}X_1$. The following proposition yields the expected answer: at the usual large deviation scaling the mass outside of $\mathrm{co}\,\mathcal{K}_0(F) + a^{-1}\mathbb{E}X_1$ asymptotically vanishes.

Proposition 1. *Under the assumptions of Theorem 1,* $\gamma_n \mathrm{P}((an)^{-1}S_n \in \mathcal{U}) \longrightarrow 0$ *as* $n \to \infty$ *for every* $a > 0$ *and measurable subset* $\mathcal{U}$ *with* $\tau := d\big(\mathcal{U}, \operatorname{co}\mathcal{K}_0(F) + a^{-1}\mathbb{E}X_1\big) > 0$.

Proof. We again switch to the isometric embedding $h : \operatorname{co}\mathcal{K}(F) \to \mathcal{C}(B^*, w^*)$ given by the support function. Let $\mathcal{V} = h(\mathcal{U})$ and $\mathcal{W} = h\big(\operatorname{co}\mathcal{K}_0(F) + a^{-1}\mathbb{E}X_1\big)$. By isometry,

$$\inf_{f \in \mathcal{V}, g \in \mathcal{W}} \|f - g\|_\infty = \tau. \tag{10}$$

For $\delta > 0$ we write in the notation of the proof of Theorem 1,

$$\begin{aligned}
\mathrm{P}((an)^{-1}S_n \in \mathcal{U}) &= \mathrm{P}((an)^{-1}\tilde{S}_n + a^{-1}\mathbb{E}h_{X_1} \in \mathcal{V}) \\
&\le \mathrm{P}\big(\|h_{X_j}\|_\infty > \delta n \text{ for at least two different } j = 1, \dots, n\big) \\
&\quad + \sum_{j=1}^{n} \mathrm{P}((an)^{-1}\tilde{S}_n + a^{-1}\mathbb{E}h_{X_1} \in \mathcal{V},\, \|h_{X_i}\|_\infty \le \delta n, i \ne j, i = 1, \dots, n) \\
&:= I_1 + I_2.
\end{aligned}$$

We already know that $\gamma_n I_1 \longrightarrow 0$. Furthermore,

$$\begin{aligned}
I_2 &= n\mathrm{P}((an)^{-1}\tilde{S}_n + a^{-1}\mathbb{E}h_{X_1} \in \mathcal{V},\, \|h_{X_i}\|_\infty \le \delta n \text{ for } i = 1, \dots, n-1) \\
&\le n\mathrm{P}\big((an)^{-1}\tilde{S}_n + a^{-1}\mathbb{E}h_{X_1} \in \mathcal{V},\, \|\tilde{S}_{n-1}\|_\infty \le \tau an/2\big) \\
&\quad + n\mathrm{P}\big(\|\tilde{S}_{n-1}\|_\infty > \tau an/2,\, \|h_{X_i}\|_\infty \le \delta n \text{ for } i = 1, \dots, n-1\big) \\
&:= I_{21} + I_{22}.
\end{aligned}$$

Note that

$$(an)^{-1}\tilde{S}_n + a^{-1}\mathbb{E}h_{X_1} = (an)^{-1}h_{X_n} + a^{-1}\mathbb{E}h_{X_1} + (an)^{-1}\tilde{S}_{n-1} - (an)^{-1}\mathbb{E}h_{X_1}.$$

Clearly, $h_{X_n}/(an) + \mathbb{E}h_{X_1}/a \in \mathcal{W}$, while on the event described in I_{21},

$$\|(an)^{-1}\tilde{S}_{n-1} - (an)^{-1}\mathbb{E}h_{X_1}\|_\infty \le 0.5\tau + (an)^{-1}\|\mathbb{E}h_{X_1}\|_\infty < \tau$$

for large n. Therefore, (10) says that $I_{21} = 0$ for large n. Furthermore, we have already established in the proof of Theorem 1 that $\gamma_n I_{22} \longrightarrow 0$ as $n \to \infty$ if δ is small enough, relative to τ. The statement of the proposition follows.

An interesting question is whether Theorem 1 extends to generally non-convex random compact sets. A first observation is the following: while the set function $\mathrm{P}(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1)$ is a measure on measurable subsets $\mathcal{U}$ of $\operatorname{co}\mathcal{K}_0(F)$, it is generally NOT a measure on all measurable subsets $\mathcal{U}$ of $\mathcal{K}_0(F)$. For example, for disjoint collections of compact sets, $\mathcal{U}_1$ and $\mathcal{U}_2$, the collections $\mathcal{U}_1 + n\mathbb{E}X_1$ and $\mathcal{U}_2 + n\mathbb{E}X_1$ may not be disjoint. Therefore we cannot hope for a result stated as convergence of measures but one can hope for a convergence result of set functions evaluated on certain sets; see below. We only consider regularly varying random compact sets in $\mathbb{R}^d$ for some $d \ge 1$ for which the tail measure is supported by $\operatorname{co}\mathcal{K}(\mathbb{R}^d)$. Informally, those are random compact sets whose tails are the heaviest "in the convex directions". A good comparison is with real-valued regularly varying random variables whose tail measures are supported by the positive half-line, e.g. α-stable variables with $1 \le \alpha < 2$ and $\beta = 1$; see [13], Chapter 1. We only consider linear scaling sequences $\{\lambda_n\}$.

Theorem 2. *For $d \geq 1$ let $\{X_n\}_{n\geq 1}$ be an iid sequence of random compact sets in $\mathbb{R}^d$, $X_1 \in \mathrm{RV}(\alpha,\mu)$ with $\alpha \geq 1$, $\mathbb{E}\|X_1\| < \infty$ and $\mu \in M_0\big(\mathcal{K}_0(\mathbb{R}^d)\big)$ supported by $\mathrm{co}\,\mathcal{K}(\mathbb{R}^d)$. For $a > 0$ and $\mathcal{U} \subseteq \mathcal{K}(\mathbb{R}^d)$ let*

$$\mathcal{V}^* = \big\{V \in \mathrm{co}\,\mathcal{K}(\mathbb{R}^d) : V + a^{-1}\mathbb{E}X_1 \in \mathrm{cl}(\mathcal{U} + a^{-1}\mathbb{E}X_1)\big\},$$
$$\mathcal{V}_* = \big\{V \in \mathrm{co}\,\mathcal{K}(\mathbb{R}^d) : V + a^{-1}\mathbb{E}X_1 \in \mathrm{int}(\mathcal{U} + a^{-1}\mathbb{E}X_1)\big\}.$$

Then for $\mathcal{U}$ bounded away from the "special element" A_0, with $\gamma_n = (n\mathrm{P}(\|X_1\| > an))^{-1}$,

$$\mu\big(\mathcal{V}_*\big) \leq \liminf_{n\to\infty} \gamma_n \mathrm{P}(S_n \in an\,\mathcal{U} + n\mathbb{E}X_1) \leq \limsup_{n\to\infty} \gamma_n \mathrm{P}(S_n \in an\,\mathcal{U} + n\mathbb{E}X_1) \leq \mu\big(\mathcal{V}^*\big).$$

Consider the complete separable metric space $\mathcal{K}(F) \times \mathrm{co}\,\mathcal{K}(F)$ equipped with the topology of coordinate convergence. With "special element" $\big(\{0\},\{0\}\big)$, we define $M_0\big(\mathcal{K}(F) \times \mathrm{co}\,\mathcal{K}(F)\big)$ as the space of Borel measures on the metric space that are finite outside of a neighborhood of the "special element". Regular variation of a random pair $(X,Y) \in \mathcal{K}(F) \times \mathrm{co}\,\mathcal{K}(F)$ can be defined straightforwardly.

The proof of the following lemma is the same as that of the second part of Lemma 1.

Lemma 2. *If a random compact set X is regularly varying in $\mathcal{K}(F)$ then the pair $\big(X, \mathrm{co}\,X\big)$ is regularly varying in $\mathcal{K}(F) \times \mathrm{co}\,\mathcal{K}(F)$. Specifically, if (1) holds, then*

$$n\mathrm{P}\Big(\big(X, \mathrm{co}\,X\big) \in a_n\cdot\Big) \longrightarrow \nu(\cdot) \quad \text{in } M_0\big(\mathcal{K}(F) \times \mathrm{co}\,\mathcal{K}(F)\big),$$

where $\nu = \mu \circ (I,c)^{-1}$, with I the identity map, and $c : \mathcal{K}(F) \to \mathrm{co}\,\mathcal{K}(F)$ is the continuous map assigning to a compact set its convex hull.

Proof of Theorem 2. Let us start with the following consequence of the regular variation assumptions imposed in the theorem: for every $\varepsilon > 0$,

$$\mathrm{P}(t^{-1}d\big(X, \mathrm{co}\,X\big) > \delta \mid \mid \|X\| > \varepsilon t) \longrightarrow 0 \ \text{ as } t \to \infty \text{ for every } \delta > 0. \tag{11}$$

To prove (11) we may and will assume that $\varepsilon = 1$. Notice that by Lemma 2,

$$\mathrm{P}\left(t^{-1}\big(X, \mathrm{co}\,X\big) \in \cdot \mid \|X\| > t\right) \longrightarrow \frac{\nu\big(\{(A,B) \in \cdot,\ \|B\| > 1\}\big)}{\nu\big(\{(A,B) :\ \|B\| > 1\}\big)} \tag{12}$$

weakly in $\mathcal{K}(\mathbb{R}^d) \times \mathrm{co}\,\mathcal{K}(\mathbb{R}^d)$, and the limit measure is concentrated on pairs (B,B) where B is convex. Since the map $(A,B) \mapsto d(A,B)$, $\mathcal{K}(\mathbb{R}^d) \times \mathrm{co}\,\mathcal{K}(\mathbb{R}^d) \to [0,\infty)$ is continuous, we conclude that the conditional law of $d\big(X, \mathrm{co}\,X\big)/t$ given $\|X\| > t$ converges weakly to the law of $d(A,B)$, where the pair (A,B) is distributed according to the law in the right hand side of (12). However, $d(A,B) = 0$ a.s. according to the latter law, and so (11) follows.

Denote $S_n^0 = \mathrm{co}\,X_1 + \cdots + \mathrm{co}\,X_n$, $n \geq 1$. Let $\mathcal{U} \subseteq \mathcal{K}_0(F)$ be bounded away from A_0. For $\varepsilon > 0$ we write, with $\lambda_n = an$,

$$\begin{aligned}\gamma_n \mathrm{P}(S_n \in \lambda_n\,\mathcal{U} + n\mathbb{E}X_1) &= \gamma_n \mathrm{P}\big(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1,\ d(S_n, S_n^0) > \varepsilon\lambda_n\big) \\ &\quad + \gamma_n \mathrm{P}\big(S_n \in \lambda_n\mathcal{U} + n\mathbb{E}X_1,\ d(S_n, S_n^0) \leq \varepsilon\lambda_n\big) := I_1 + I_2.\end{aligned} \tag{13}$$

To estimate I_1, we will use the following estimate on the Hausdorff distance between sums of compact sets and their respective convex hulls (e.g. p. 195 in [9] or p. 881 in [1]): for any $n \geq 1$ and compact subsets $A_1, \dots, A_n$ of $\mathbb{R}^d$,

$$d(A_1 + \dots + A_n, \operatorname{co} A_1 + \dots + \operatorname{co} A_n) \leq d^{1/2} \max_{j=1,\dots,n} \|A_j\|. \tag{14}$$

Then

$$\begin{aligned}
I_1 &\leq \gamma_n \mathrm{P}\big(d(S_n, S_n^0) > \varepsilon\lambda_n\big) \\
&= \gamma_n \mathrm{P}\big(d(S_n, S_n^0) > \varepsilon\lambda_n,\, \|X_j\| \leq [\varepsilon/(2d^{1/2})]\lambda_n \text{ for each } j = 1, \dots, n\big) \\
&\quad + \gamma_n \mathrm{P}\big(d(S_n, S_n^0) > \varepsilon\lambda_n,\, \|X_j\| > [\varepsilon/(2d^{1/2})]\lambda_n \text{ for exactly one } j = 1, \dots, n\big) \\
&\quad + \gamma_n \mathrm{P}\big(d(S_n, S_n^0) > \varepsilon\lambda_n,\, \|X_j\| > [\varepsilon/(2d^{1/2})]\lambda_n \text{ for at least 2 different } j = 1, \dots, n\big) \\
&:= I_{11} + I_{12} + I_{13}.
\end{aligned}$$

It follows from (14) that, if $\|X_j\| \leq [\varepsilon/(2d^{1/2})]\lambda_n$ for $j = 1, \dots, n$, then $d(S_n, S_n^0) \leq \varepsilon\lambda_n/2$, so that $I_{11} = 0$. Furthermore,

$$I_{13} \leq \gamma_n \mathrm{P}\big(\|X_j\| > [\varepsilon/(2d^{1/2})]\lambda_n \text{ for at least 2 different } j = 1, \dots, n\big) \longrightarrow 0$$

by the choice of the sequence $\{\lambda_n\}$. Finally, we use (14) once again to see that

$$\begin{aligned}
I_{12} &= \gamma_n \sum_{j=1}^{n} \mathrm{P}\big(d(S_n, S_n^0) > \varepsilon\lambda_n,\, \|X_j\| > [\varepsilon/(2d^{1/2})]\lambda_n, \\
&\qquad \|X_i\| \leq [\varepsilon/(2d^{1/2})]\lambda_n \text{ for } i = 1, \dots, n,\, i \neq j\big) \\
&\leq \gamma_n \sum_{j=1}^{n} \mathrm{P}\Big(d\big(\sum_{i\neq j} X_i, \sum_{i\neq j} \operatorname{co} X_i\big) + d\big(X_j, \operatorname{co} X_j\big) > \varepsilon\lambda_n, \\
&\qquad \|X_j\| > [\varepsilon/(2d^{1/2})]\lambda_n,\, \|X_i\| \leq [\varepsilon/(2d^{1/2})]\lambda_n \text{ for } i = 1, \dots, n,\, i \neq j\Big) \\
&\leq \gamma_n \sum_{j=1}^{n} \mathrm{P}\big(\varepsilon\lambda_n/2 + d\big(X_j, \operatorname{co} X_j\big) > \varepsilon\lambda_n,\, \|X_j\| > [\varepsilon/(2d^{1/2})]\lambda_n\big) \\
&= [\mathrm{P}(\|X_1\| > \lambda_n)]^{-1} \mathrm{P}\Big(d\big(X_1, \operatorname{co} X_1\big) > \varepsilon\lambda_n/2,\, \|X_1\| > [\varepsilon/(2d^{1/2})]\lambda_n\Big) \longrightarrow 0
\end{aligned}$$

by (11) and regular variation. Therefore, $I_1 \longrightarrow 0$ in the right hand side of (13).

For the second term in the right hand side of (13) we have, since $\lambda_n = an$,

$$I_2 \leq \gamma_n \mathrm{P}\left((an)^{-1} S_n^0 \in \big(\mathcal{U} + a^{-1}\mathbb{E}X_1\big)^{\varepsilon}\right).$$

Denote $\mathcal{V}^{*\varepsilon} = \big\{V \in \operatorname{co}\mathcal{K}(\mathbb{R}^d) : V + a^{-1}\mathbb{E}X_1 \in (\mathcal{U} + a^{-1}\mathbb{E}X_1)^{\varepsilon}\big\}$, so that

$$\begin{aligned}
I_2 &\leq \gamma_n \mathrm{P}\big((an)^{-1} S_n^0 \in \mathcal{V}^{*2\varepsilon} + a^{-1}\mathbb{E}X_1\big) + \gamma_n \mathrm{P}\big(d\big((an)^{-1} S_n^0, \operatorname{co}\mathcal{K}(\mathbb{R}^d) + a^{-1}\mathbb{E}X_1\big) > \varepsilon\big) \\
&:= I_{21} + I_{22}.
\end{aligned}$$

By Proposition 1, $I_{22} \longrightarrow 0$. Furthermore, the set $\mathcal{V}^{*2\varepsilon}$ is a closed subset of $\operatorname{co}\mathcal{K}(\mathbb{R}^d)$ that is bounded away from A_0. Therefore, by Theorem 1, $\limsup_{n\to\infty} I_{21} \leq \mu\big(\mathcal{V}^{*2\varepsilon}\big)$. Since $\mathcal{V}^{*2\varepsilon} \downarrow \mathcal{V}^*$ as $\varepsilon \to 0$, we conclude that

$$\limsup_{n\to\infty} \gamma_n \mathrm{P}(S_n \in an\,\mathcal{U} + n\mathbb{E}X_1) \leq \mu\big(\mathcal{V}^*\big),$$

and the proof of the corresponding lower bound on I_2 in the right hand side of (13) is similar.

Acknowledgment. Thomas Mikosch's research is partly supported by the Danish Natural Science Research Council (FNU) Grant 09-072331, "Point process modelling and statistical inference". Zbyněk Pawlas is partly supported by the Czech Ministry of Education, research project MSM 0021620839 and by the Grant Agency of the Czech Republic, grant P201/10/0472. Gennady Samorodnitsky's research is partially supported by a US Army Research Office (ARO) grant W911NF-10-1-0289 and a National Science Foundation (NSF) grant DMS-1005903 at Cornell University.

References

[1] Artstein, Z. and Vitale, R. (1975). A strong law of large numbers for random compact sets. *The Annals of Probability* **3**, 879–882.

[2] Cerf, R. (1999). Large deviaition results for sums of iid random compact sets. *Proceedings of the American Mathematical Society* **127**, 2431–2436.

[3] 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.

[4] Giné, E., Hahn, M. and Zinn, J. (1983). Limit theorems for random sets: application of probability in Banach space results. In: Beck, A. and Jacobs, K. (Eds.). *Probability in Banach Spaces IV (Oberwolfach, 1982). Lect. Notes Math.* **990**, pp. 112–135. Springer, Berlin.

[5] Hult, H. and Lindskog, F. (2006). Regular variation for measures on metric spaces. *Publications de l'Institut Mathématique, Nouvelle Série* **80**, 121–140.

[6] Hult, H., Lindskog, F., Mikosch, T. and Samorodnitsky, G. (2005). Functional large deviations for multivariate regularly varying random walks. *The Annals of Applied Probability* **15**, 2651–2680.

[7] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces.* Springer-Verlag, Berlin.

[8] Mikosch, T., Pawlas, Z. and Samorodnitsky, G. (2011). Large deviations for Minkowski sums of heavy-tailed generally non-convex random compact sets. *Vestnik of the University of St. Petersburg*, to appear.

[9] Molchanov, I. (2005). *Theory of Random Sets.* Springer, London.

[10] Nagaev, A. V. (1969). Integral limit theorems taking large deviations into account when Cramér's condition does not hold I, II. *Theory of Probability and its Applications* **14**, 51–64, 193–208.

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

[12] Revuz, D. and Yor, M. (2001). *Continuous Martingales and Brownian Motion.* Corrected second printing of the third edition. Springer, Berlin.

[13] Samorodnitsky, G. and Taqqu, M. (1994). *Stable Non-Gaussian Random Processes.* Chapman & Hall, New York.