

# NEW FRONTIERS IN APPLIED PROBABILITY

A Festschrift for SØREN ASMUSSEN  
Edited by P. GLYNN, T. MIKOSCH and T. ROLSKI

Part 3. Heavy tail phenomena

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

THOMAS MIKOSCH, *University of Copenhagen*

Department of Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100  
Copenhagen, Denmark. Email address: mikosch@math.ku.dk

ZBYNĚK PAWLAS, *Charles University in Prague*

Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics,  
Charles University in Prague, Sokolovská 83, Praha 8, Czech Republic.  
Email address: pawlas@karlin.mff.cuni.cz

GENNADY SAMORODNITSKY, *Cornell University*

School of Operations Research and Information Engineering, Cornell University, 220 Rhodes Hall,  
Ithaca, NY 14853, USA. Email address: gennady@orie.cornell.edu



APPLIED PROBABILITY TRUST  
AUGUST 2011

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

BY THOMAS MIKOSCH, ZBYNĚK PAWLAS AND GENNADY SAMORODNITSKY

---

## Abstract

We prove large deviation results for Minkowski sums  $S_n$  of independent and identically distributed 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 Mikosch, Pawlas and Samorodnitsky (2011) for generally nonconvex 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

2010 Mathematics Subject Classification: Primary 60F10

Secondary 52A22

---

## 1. Introduction

### 1.1. 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 the reader to [9]. Let  $F$  be a separable Banach space with norm  $\|\cdot\|$ . For  $A_1, A_2 \subseteq F$  and a real number  $\lambda$ , the Minkowski addition and scalar multiplication are respectively 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  $\mathcal{K} = \mathcal{K}(F)$  the class of all nonempty compact subsets of  $F$ . Note that this is not a vector space. However, it is well known that  $\mathcal{K}$  equipped with the Hausdorff distance

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

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

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

For any subset  $\mathcal{U}$  of  $\mathcal{K}$ , a real number  $\lambda$ , and a set  $A \in \mathcal{K}$ , 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  $\mathcal{U}_1$  and  $\mathcal{U}_2$  of  $\mathcal{K}$ , we define  $d(\mathcal{U}_1, \mathcal{U}_2) = \inf_{A_1 \in \mathcal{U}_1, A_2 \in \mathcal{U}_2} d(A_1, A_2)$ .

A random compact set  $X$  in  $F$  is a Borel measurable function from an abstract probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  into  $\mathcal{K}$ . 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 of [9] for an overview of results obtained until 2005. A general Cramér-type large deviation result for Minkowski sums of independent and identically distributed (i.i.d.) 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 deviation 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 Nagaev [10], [11] for sums of i.i.d. random variables (cf. [6] and [12]). 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.

**1.2. Regularly varying random sets**

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

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

Furthermore, we define the norm  $\|A\| = d(A, \{0\}) = \sup\{\|a\| : a \in A\}$  for  $A \in \mathcal{K}$ , and define  $\mathcal{B}_r = \{A \in \mathcal{K} : \|A\| \leq r\}$ . Let  $M_0 = M_0(\mathcal{K}_0)$  be the collection of Borel measures on  $\mathcal{K}_0$  whose restriction to  $\mathcal{K} \setminus \mathcal{B}_r$  is finite for each  $r > 0$ . Let  $\mathcal{C}_0$  denote the class of real-valued, bounded, and continuous functions  $f$  on  $\mathcal{K}_0$  such that, for each  $f$ , there exists  $r > 0$  and  $f$  vanishes on  $\mathcal{B}_r$ . The convergence  $\mu_n \rightarrow \mu$  in  $M_0$  is defined to mean the convergence  $\int f d\mu_n \rightarrow \int f d\mu$  for all  $f \in \mathcal{C}_0$ . By the portmanteau theorem (see [5, Theorem 2.4]),  $\mu_n \rightarrow \mu$  in  $M_0$  if and only if  $\mu_n(\mathcal{U}) \rightarrow \mu(\mathcal{U})$  for all Borel sets  $\mathcal{U} \subseteq \mathcal{K}$  which are bounded away from  $\{0\}$  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 linear metric space, a random compact set  $X$  is *regularly varying* if there exist a nonnull measure  $\mu \in M_0$  and a sequence  $\{a_n\}_{n \geq 1}$  of positive numbers such that

$$nP(X \in a_n \cdot) \rightarrow \mu(\cdot) \quad \text{in } M_0. \tag{1}$$

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

$$[\text{P}(X \in t(\mathcal{K} \setminus \mathcal{B}_1))]^{-1}\text{P}(X \in t \cdot) \rightarrow c\mu(\cdot) \quad \text{in } M_0 \text{ as } t \rightarrow \infty, \tag{2}$$

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

An important closed subset of  $\mathcal{K}$  is the family of nonempty *compact convex subsets* of  $F$ , denoted by  $\text{co}\mathcal{K}$ . Denote the topological dual of  $F$  by  $F^*$  and the unit ball of  $F^*$  by  $B^*$ , endowed with the weak-\* topology  $w^*$ . The *support function*  $h_A$  of a compact convex  $A \in \text{co}\mathcal{K}$  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 ( $h_A(u + v) \leq h_A(u) + h_A(v)$  for all  $u, v \in B^*$  with  $u + v \in B^*$ ) and positively homogeneous ( $h_A(cu) = ch_A(u)$  for all  $c > 0$  and  $u \in B^*$  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  $h : \text{co}\mathcal{K} \rightarrow \mathcal{C}(B^*, w^*)$  has the properties that

$$h_{A_1+A_2} = h_{A_1} + h_{A_2}, \quad h_{\lambda A_1} = \lambda h_{A_1}, \quad A_1, A_2 \in \text{co}\mathcal{K}, \lambda \geq 0,$$

which make it possible to convert the Minkowski sums and scalar multiplication of convex sets respectively 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  $\text{co}\mathcal{K}$  into  $\mathcal{C}(B^*, w^*)$  with the uniform norm. If  $\mathcal{G} = h(\text{co}\mathcal{K})$  then  $\mathcal{G}$  is a closed convex cone in  $\mathcal{C}(B^*, w^*)$ , and  $h$  is an isometry between  $\text{co}\mathcal{K}$  and  $\mathcal{G}$ .

A random compact convex set  $X$  is a Borel measurable function from a probability space  $(\Omega, \mathcal{F}, P)$  into  $\text{co}\mathcal{K}$ , which we endow with the relative topology inherited from  $\mathcal{K}$ . 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 nonzero measure  $\mu \in M_0(\text{co}\mathcal{K}_0)$  and a sequence  $\{a_n\}_{n \geq 1}$  of positive numbers such that

$$nP(X \in a_n \cdot) \rightarrow \mu(\cdot) \quad \text{in } M_0(\text{co}\mathcal{K}_0). \tag{4}$$

Once again, the tail measure  $\mu$  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  $\text{co}\mathcal{K}$  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*

$$nP(h_X \in a_n \cdot) \rightarrow \nu(\cdot) \quad \text{in } M_0(\mathcal{C}(B^*, w^*)), \tag{5}$$

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  $\mathcal{K}$  then its convex hull  $\text{co}X$  is a random compact convex set, which is regularly varying in  $\text{co}\mathcal{K}$ . 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  $A \mapsto \text{co}A$  from  $\mathcal{K}$  to  $\text{co}\mathcal{K}$ . In particular,  $X$  and  $\text{co}X$  have the same indices of regular variation.*

*Proof.* Since isometry implies continuity, and 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 (see Theorem 2.5 of [5]). For part (ii), note that the map  $A \mapsto \text{co}A$  from  $\mathcal{K}$  to  $\text{co}\mathcal{K}$  is a contraction in the Hausdorff distance, and, 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.

**1.3. 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 i.i.d. regularly varying random compact sets. To the best of the authors’ knowledge, such results are not available in the literature; they parallel those proved by A. and S. Nagaev [10], [11], [12] for sums of i.i.d. 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  $\lambda_n$  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 nonconvex case as well. Large deviation results for Minkowski sums of random sets have their own interest; in [8] we considered first applications to intrinsic volumes.

**2. Examples of regularly varying random sets**

Simple examples of regularly varying random sets can be constructed from i.i.d.  $F$ -valued random elements  $\xi_1, \dots, \xi_k, k \geq 2$ , which are regularly varying with index  $\alpha > 0$  and tail measure  $\nu$ . The following three examples are distinct but their tail measures turn out to be the same; see [8] for proofs.

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

Next we consider an example of a nonconvex regularly varying random set whose tail measure is supported by convex sets.

**Example 2.** (*Sojourn set.*) For  $k \geq 3$ , let  $\{W_t = (W_t^{(i)})_{i=1, \dots, k}, t \in \mathbb{R}^+\}$  be a standard Brownian motion, i.e. the  $W_i$ s are independent standard Brownian motions in  $\mathbb{R}$ . Then  $\{\|W_t\|, t \in \mathbb{R}^+\}$  is a Bessel process of order  $k$ .

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

$$M = \sup\{t : t \in X\} = \sup\{t \in \mathbb{R}^+ : \|W_t\| \leq 1\}.$$

It follows from the last part of Exercise 1.18 of [13, p. 450] that  $M^{-1}$  is  $\chi_{k-2}^2$ -distributed. Therefore,

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

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

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

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

$$nP(d(X, Y) > n^{2/(k-2)}\varepsilon) \rightarrow 0. \tag{6}$$

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

$$\begin{aligned} P(d(X, Y) > n^{2/(k-2)}\varepsilon) &= P(X \text{ contains a point separated by more than } n^{2/(k-2)}\varepsilon \text{ from both } 0 \text{ and } M) \\ &\leq P(M - T > n^{2/(k-2)}\varepsilon). \end{aligned}$$

Note that  $T$  is a stopping time and that the process  $\{\|W_t\|, t \in \mathbb{R}^+\}$  is a Feller process and, hence, is strongly Markov; see [13, p. 446]. Therefore,

$$P(M - T > n^{2/(k-2)}\varepsilon) = E(\mathbf{1}_{\{T < \infty\}} P_{\|W_T\|}(M > n^{2/(k-2)}\varepsilon)).$$

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,

$$\begin{aligned} P_{\|W_T\|}(M > n^{2/(k-2)}\varepsilon) &\leq P(\sup\{t \in \mathbb{R}^+ : \|W_t\| \leq 2\} > n^{2/(k-2)}\varepsilon) \\ &= P\left(M > \frac{1}{4}n^{2/(k-2)}\varepsilon\right) \\ &\leq cn^{-1} \end{aligned}$$

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

$$P(d(X, Y) > n^{2/(k-2)}\varepsilon) \leq cn^{-1}P(T < \infty) \leq cn^{-1}P(M \geq n^{2/(k-2)}\varepsilon) = O(n^{-2})$$

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

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

$$X_1 = \{tW_t, t \in \mathbb{R}^+ : \|W_t\| \leq 1\} \subseteq \mathbb{R}^k.$$

It follows from what we already know about the set  $X$  that  $X_1$  is regularly varying, with tail measure  $\mu_1 = (v \times H) \circ T_1^{-1}$ , where  $H$  is the normalized Haar measure (see, e.g. [7, p. 365]) on the unit sphere  $\mathbb{S}^{k-1}$ , and  $T_1 : \mathbb{R}^+ \times \mathbb{S}^{k-1} \rightarrow \mathcal{K}(\mathbb{R}^k)$  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  $S_n = X_1 + \dots + X_n$  of i.i.d. regularly varying random compact sets  $X_i, i = 1, 2, \dots$ , 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, p. 151], a random element  $\xi \in F$  is a *selection of  $X$*  if  $\xi \in X$  almost surely (a.s.), and if  $E\|\xi\| < \infty$ ,  $\xi$  is an integrable selection. The *selection expectation* of  $X$  is defined as  $EX = \text{cl}\{E\xi : \xi \text{ is an integrable selection of } X\}$ . 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  $E\|X\| < \infty$ , then the selection expectation of  $X$  is the unique convex compact set  $EX$  satisfying  $Eh_X(u) = h_{EX}(u)$  for all  $u \in B^*$ ; see [9, Theorem 2.1.22].

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

$$\lambda_n^{-1}d(S_n, nEX_1) \xrightarrow{P} 0, \tag{7}$$

$$\lambda_n^{-1}Ed(S_n, nEX_1) \rightarrow 0, \tag{8}$$

and, for some  $\eta > 0$ ,

(i)  $\lambda_n/n^{1/2+\eta} \rightarrow \infty$  if  $\alpha \geq 2$ , and

(ii)  $\lambda_n/n^{1/\alpha+\eta} \rightarrow \infty$  if  $1 \leq \alpha < 2$ .

Then, with  $\gamma_n = [nP(\|X_1\| > \lambda_n)]^{-1}$ ,

$$\gamma_n P(S_n \in \lambda_n \cdot + nEX_1) \rightarrow \mu(\cdot) \text{ in } M_0(\text{co}\mathcal{K}_0).$$

**Remark 1.** Note that the assumptions of Theorem 1 imply that  $\lambda_n/a_n \rightarrow \infty$ . Regarding the assumptions on the random set, we start by observing that the condition  $E\|X_1\| < \infty$  is automatic if  $\alpha > 1$ . Furthermore, condition (7) can be easily verified if the random sets satisfy the central limit theorem. For example, if  $d(S_n, nEX_1)/\sqrt{n}$  converges in distribution (as it does when a Gaussian central limit theorem holds) and  $\lambda_n/\sqrt{n} \rightarrow \infty$ , then (7), obviously, holds. This Gaussian central limit theorem requires  $\alpha > 2$ , and assumption (i) of Theorem 1 already implies that  $\lambda_n/\sqrt{n} \rightarrow \infty$ . Alternatively, if  $d(S_n, nEX_1)/a_n$  converges in distribution in the context of an  $\alpha$ -stable central limit theorem,  $1 < \alpha < 2$ , then (7) also holds since  $\lambda_n/a_n \rightarrow \infty$ . Sufficient conditions for the central limit theorem can be found in [4] and [9]. If the Gaussian central limit theorem is satisfied then condition (8) follows by the isometric embedding (3) and Corollary 10.2 of [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 i.i.d. 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 in [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 nonset-valued large deviation contexts; see, e.g. [3], [6], and [12].

*Proof of Theorem 1.* Let  $\mathcal{U} \subseteq \text{co}\mathcal{K}_0$  be a  $\mu$ -continuity set, bounded away from  $\{0\}$ . We will show that  $\gamma_n P(S_n \in \lambda_n \mathcal{U} + nEX_1) \rightarrow \mu(\mathcal{U})$ . We start with an upper bound:

$$\begin{aligned} P(S_n \in \lambda_n \mathcal{U} + nEX_1) &= P\left(S_n \in \lambda_n \mathcal{U} + nEX_1, \bigcup_{i=1}^n \{X_i \in \lambda_n \mathcal{U}^\varepsilon\}\right) \\ &\quad + P\left(S_n \in \lambda_n \mathcal{U} + nEX_1, \bigcap_{i=1}^n \{X_i \notin \lambda_n \mathcal{U}^\varepsilon\}\right) \\ &\leq nP(X_1 \in \lambda_n \mathcal{U}^\varepsilon) + P\left(S_n \in \lambda_n \mathcal{U} + nEX_1, \bigcap_{i=1}^n \{X_i \notin \lambda_n \mathcal{U}^\varepsilon\}\right) \\ &:= I_1 + I_2. \end{aligned}$$

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

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

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

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

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

$$I_2 \leq \mathbb{P} \left( \bigcap_{i=1}^n \{\|\tilde{S}_n - h_{X_i}\|_\infty > \varepsilon \lambda_n\} \right).$$

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

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

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

$$\mathbb{P} \left( \bigcap_{i=1}^n \{\|\tilde{S}_n - h_{X_i}\|_\infty > \varepsilon \lambda_n\} \cap B_2 \right) \leq \mathbb{P} \left( \|\tilde{S}_{n-1}\|_\infty > \frac{\varepsilon \lambda_n}{2}, \max_{i=1, \dots, n-1} \|Y_i\|_\infty \leq 2\delta \lambda_n \right).$$

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

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

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

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

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  $b = 2n(\|\mathbb{E}h_{X_1}\|_\infty^2 + \mathbb{E}\|h_{X_1}\|_\infty^2 \mathbf{1}_{\{\|h_{X_1}\|_\infty \leq 3\delta \lambda_n\}})$ . If  $\alpha = 2$ , by Karamata's theorem, we have  $b = nl(n)$  for a slowly varying (at infinity) function  $l$ . If  $1 \leq \alpha < 2$  then, by



Karamata’s theorem, our choice results in  $b \sim cn\lambda_n^2\mathbb{P}(\|h_{X_1}\|_\infty > 3\delta\lambda_n)$  as  $n \rightarrow \infty$  for some  $c > 0$ . In all three cases, for all small enough  $\delta > 0$ ,

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

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

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

and observe that  $\mathbb{E}\|\tilde{S}_n\|_\infty/\lambda_n \rightarrow 0$  by assumption (8). To see that  $\mathbb{E}\|\tilde{S}_n - \tilde{S}_n^\delta\|_\infty/\lambda_n \rightarrow 0$ , note that, by Karamata’s theorem and the choice of  $\lambda_n$ ,

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

We conclude that, for any  $\mu$ -continuity set  $\mathcal{U}$  bounded away from  $\{0\}$ ,

$$\limsup_{n \rightarrow \infty} \gamma_n\mathbb{P}(S_n \in \lambda_n\mathcal{U} + nEX_1) \leq \mu(\mathcal{U}^\varepsilon) \rightarrow \mu(\mathcal{U}) \quad \text{as } \varepsilon \searrow 0,$$

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

$$\begin{aligned} \mathbb{P}(S_n \in \lambda_n\mathcal{U} + nEX_1) &\geq \mathbb{P}\left(S_n \in \lambda_n\mathcal{U} + nEX_1, \bigcup_{i=1}^n \{X_i \in \lambda_n\mathcal{U}^{-\varepsilon}\}\right) \\ &\geq \mathbb{P}\left(\bigcup_{i=1}^n \{X_i \in \lambda_n\mathcal{U}^{-\varepsilon}\}\right) \\ &\quad - \mathbb{P}\left(S_n \notin \lambda_n\mathcal{U} + nEX_1, \bigcup_{i=1}^n \{X_i \in \lambda_n\mathcal{U}^{-\varepsilon}\}\right) \\ &:= I_1 - I_2. \end{aligned}$$

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

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

By the choice of  $\lambda_n$ , for  $\varepsilon > 0$  so small that  $\mathcal{U}^{-\varepsilon}$  is bounded away from  $\{0\}$  and a  $\mu$ -continuity set,  $\liminf_{n \rightarrow \infty} \gamma_n I_1 \geq \mu(\mathcal{U}^{-\varepsilon})$ . Letting  $\varepsilon \rightarrow 0$  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  $\mathbb{P}(S_n \in \lambda_n\mathcal{U} + nEX_1)$ ,  $\mathcal{U}$ , a measurable subset of  $\text{co}\mathcal{K}_0$ , is, in fact, a probability measure on  $\text{co}\mathcal{K}_0$ , and the sets  $\lambda_n\mathcal{U} + nEX_1$  do not cover all measurable subsets of  $\text{co}\mathcal{K}_0$ , except in the trivial case  $X_1 = \{0\}$  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  $\gamma_n\mathbb{P}((an)^{-1}S_n \in \cdot + a^{-1}EX_1) \rightarrow \mu(\cdot)$  in  $M_0(\text{co}\mathcal{K}_0)$ , which leaves unanswered the obvious question of how the law of  $(an)^{-1}S_n$  behaves on sets that are not in  $\text{co}\mathcal{K}_0 + a^{-1}EX_1$ . The following proposition yields the expected answer: at the usual large deviation scaling the mass outside of  $\text{co}\mathcal{K}_0 + a^{-1}EX_1$  asymptotically vanishes.

**Proposition 1.** *Under the assumptions of Theorem 1,  $\gamma_n P((an)^{-1}S_n \in \mathcal{U}) \rightarrow 0$  as  $n \rightarrow \infty$  for every  $a > 0$  and measurable subset  $\mathcal{U}$  with  $\tau := d(\mathcal{U}, \text{co}\mathcal{K}_0 + a^{-1}EX_1) > 0$ .*

*Proof.* We again switch to the isometric embedding  $h: \text{co}\mathcal{K} \rightarrow \mathcal{C}(B^*, w^*)$  given by the support function. Let  $\mathcal{V} = h(\mathcal{U})$  and  $\mathcal{W} = h(\text{co}\mathcal{K}_0 + a^{-1}EX_1)$ . 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,

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

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

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

Note that

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

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

$$\|(an)^{-1}\tilde{S}_{n-1} - (an)^{-1}Eh_{X_1}\|_\infty \leq 0.5\tau + (an)^{-1}\|Eh_{X_1}\|_\infty < \tau$$

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  $\gamma_n I_{22} \rightarrow 0$  as  $n \rightarrow \infty$  if  $\delta$  is small enough, relative to  $\tau$ . The statement of the proposition follows.

An interesting question is whether Theorem 1 extends to generally nonconvex random compact sets. A first observation is the following: while the set function  $P(S_n \in \lambda_n \mathcal{U} + nEX_1)$  is a measure on measurable subsets  $\mathcal{U}$  of  $\text{co}\mathcal{K}_0$ , it is generally *not* a measure on all measurable subsets  $\mathcal{U}$  of  $\mathcal{K}_0$ . For example, for disjoint collections of compact sets,  $\mathcal{U}_1$  and  $\mathcal{U}_2$ , the collections  $\mathcal{U}_1 + nEX_1$  and  $\mathcal{U}_2 + nEX_1$  may not be disjoint. Therefore, we cannot hope for a result stated as the convergence of measures, but we can hope for a convergence result of set functions evaluated on certain sets; see below. We consider only regularly varying random compact sets in  $\mathbb{R}^d$  for some  $d \geq 1$  for which the tail measure is supported by  $\text{co}\mathcal{K}(\mathbb{R}^d)$ . 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.  $\alpha$ -stable variables with  $1 \leq \alpha < 2$  and  $\beta = 1$ ; see [14, Chapter 1]. We consider only linear scaling sequences  $\{\lambda_n\}$ .

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

$$\begin{aligned} \mathcal{V}^* &= \{V \in \text{co}\mathcal{K}(\mathbb{R}^d) : V + a^{-1}EX_1 \in \text{cl}(\mathcal{U} + a^{-1}EX_1)\}, \\ \mathcal{V}_* &= \{V \in \text{co}\mathcal{K}(\mathbb{R}^d) : V + a^{-1}EX_1 \in \text{int}(\mathcal{U} + a^{-1}EX_1)\}. \end{aligned}$$

Then, for  $\mathcal{U}$  bounded away from the ‘special element’  $\{0\}$ , with  $\gamma_n = (n\text{P}(\|X_1\| > an))^{-1}$ ,

$$\mu(\mathcal{V}_*) \leq \liminf_{n \rightarrow \infty} \gamma_n \text{P}(S_n \in an\mathcal{U} + nEX_1) \leq \limsup_{n \rightarrow \infty} \gamma_n \text{P}(S_n \in an\mathcal{U} + nEX_1) \leq \mu(\mathcal{V}^*).$$

Consider the complete separable metric space  $\mathcal{K} \times \text{co}\mathcal{K}$  equipped with the topology of coordinate convergence. With ‘special element’  $(\{0\}, \{0\})$ , we define  $M_0(\mathcal{K} \times \text{co}\mathcal{K})$  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  $(X, Y) \in \mathcal{K} \times \text{co}\mathcal{K}$  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  $\mathcal{K}$  then the pair  $(X, \text{co}X)$  is regularly varying in  $\mathcal{K} \times \text{co}\mathcal{K}$ . Specifically, if (1) holds then  $n\text{P}((X, \text{co}X) \in a_n \cdot) \rightarrow \nu(\cdot)$  in  $M_0(\mathcal{K} \times \text{co}\mathcal{K})$ , where  $\nu = \mu \circ (I, c)^{-1}$ , with  $I$  the identity map, and  $c : \mathcal{K} \rightarrow \text{co}\mathcal{K}$  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  $\varepsilon > 0$ ,

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

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

$$\text{P}(t^{-1}(X, \text{co}X) \in \cdot \mid \|X\| > t) \rightarrow \frac{\nu(\{(A, B) \in \cdot, \|B\| > 1\})}{\nu(\{(A, B) : \|B\| > 1\})} \tag{12}$$

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

Define  $S_n^0 = \text{co}X_1 + \dots + \text{co}X_n$ ,  $n \geq 1$ . Let  $\mathcal{U} \subseteq \mathcal{K}_0$  be bounded away from  $\{0\}$ . For  $\varepsilon > 0$ , we write, with  $\lambda_n = an$ ,

$$\begin{aligned} \gamma_n \text{P}(S_n \in \lambda_n \mathcal{U} + nEX_1) &= \gamma_n \text{P}(S_n \in \lambda_n \mathcal{U} + nEX_1, d(S_n, S_n^0) > \varepsilon \lambda_n) \\ &\quad + \gamma_n \text{P}(S_n \in \lambda_n \mathcal{U} + nEX_1, d(S_n, S_n^0) \leq \varepsilon \lambda_n) \\ &:= I_1 + I_2. \end{aligned} \tag{13}$$

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

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

Then

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

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,

$$I_{13} \leq \gamma_n \mathbb{P}\left(\|X_j\| > \frac{\varepsilon}{2d^{1/2}} \lambda_n \text{ for at least two different } j = 1, \dots, n\right) \rightarrow 0$$

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

$$\begin{aligned}
 I_{12} &= \gamma_n \sum_{j=1}^n \mathbb{P}\left(d(S_n, S_n^0) > \varepsilon \lambda_n, \|X_j\| > \frac{\varepsilon}{2d^{1/2}} \lambda_n, \right. \\
 &\quad \left. \|X_i\| \leq \frac{\varepsilon}{2d^{1/2}} \lambda_n \text{ for } i = 1, \dots, n, i \neq j\right) \\
 &\leq \gamma_n \sum_{j=1}^n \mathbb{P}\left(d\left(\sum_{i \neq j} X_i, \sum_{i \neq j} \text{co}X_i\right) + d(X_j, \text{co}X_j) > \varepsilon \lambda_n, \right. \\
 &\quad \left. \|X_j\| > \frac{\varepsilon}{2d^{1/2}} \lambda_n, \|X_i\| \leq \frac{\varepsilon}{2d^{1/2}} \lambda_n \text{ for } i = 1, \dots, n, i \neq j\right) \\
 &\leq \gamma_n \sum_{j=1}^n \mathbb{P}\left(\frac{\varepsilon \lambda_n}{2} + d(X_j, \text{co}X_j) > \varepsilon \lambda_n, \|X_j\| > \frac{\varepsilon}{2d^{1/2}} \lambda_n\right) \\
 &= [\mathbb{P}(\|X_1\| > \lambda_n)]^{-1} \mathbb{P}\left(d(X_1, \text{co}X_1) > \frac{\varepsilon \lambda_n}{2}, \|X_1\| > \frac{\varepsilon}{2d^{1/2}} \lambda_n\right) \\
 &\rightarrow 0
 \end{aligned}$$

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

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

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

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

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

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

$\mathcal{V}^{*2\varepsilon} \downarrow \mathcal{V}^*$  as  $\varepsilon \rightarrow 0$ , we conclude that

$$\limsup_{n \rightarrow \infty} \gamma_n P(S_n \in an\mathcal{U} + nEX_1) \leq \mu(\mathcal{V}^*),$$

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

### Acknowledgements

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

### References

- [1] ARTSTEIN, Z. AND VITALE, R. A. (1975). A strong law of large numbers for random compact sets. *Ann. Prob.* **3**, 879–882.
- [2] CERF, R. (1999). Large deviations for sums of i.i.d. random compact sets. *Proc. Amer. Math. Soc.* **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. Tech. Rep., Texas A&M University.
- [4] GINÉ, E., HAHN, M. G. AND ZINN, J. (1983). Limit theorems for random sets: an application of probability in Banach space results. In *Probability in Banach Spaces IV* (Oberwolfach, 1982; Lecture Notes Math. **990**), eds A. Beck and K. Jacobs, Springer, Berlin, pp. 112–135.
- [5] HULT, H. AND LINDSKOG, F. (2006). Regular variation for measures on metric spaces. *Publ. Inst. Math. 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. *Ann. Appl. Prob.* **15**, 2651–2680.
- [7] LEDOUX, M. AND TALAGRAND, M. (1991). *Probability in Banach Spaces*. Springer, Berlin.
- [8] MIKOSCH, T., PAWLAS, Z. AND SAMORODNITSKY, G. (2011). Large deviations for Minkowski sums of heavy-tailed generally non-convex random compact sets. To appear in *Vestnik St. Petersburg University: Mathematics*.
- [9] MOLCHANOV, I. (2005). *Theory of Random Sets*. Springer, London.
- [10] NAGAEV, A. V. (1969). Integral limit theorems taking large deviations into account when Cramer's condition does not hold. I. *Theory Prob. Appl.* **14**, 51–64.
- [11] NAGAEV, A. V. (1969). Integral limit theorems taking large deviations into account when Cramer's condition does not hold. II. *Theory Prob. Appl.* **14**, 193–208.
- [12] NAGAEV, S. V. (1979). Large deviations of sums of independent random variables. *Ann. Prob.* **7**, 745–789.
- [13] REVUZ, D. AND YOR, M. (2001). *Continuous Martingales and Brownian Motion*. Corrected second printing of the third edition. Springer, Berlin.
- [14] SAMORODNITSKY, G. AND TAQQU, M. S. (1994). *Stable Non-Gaussian Random Processes*. Chapman and Hall, New York.

THOMAS MIKOSCH, *University of Copenhagen*

Department of Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark.

Email address: mikosch@math.ku.dk

ZBYNĚK PAWLAS, *Charles University in Prague*

Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University in Prague, Sokolovská 83, Praha 8, Czech Republic. Email address: pawlas@karlin.mff.cuni.cz

GENNADY SAMORODNITSKY, *Cornell University*

School of Operations Research and Information Engineering, Cornell University, 220 Rhodes Hall, Ithaca, NY 14853, USA. Email address: gennady@orie.cornell.edu