
SUBADDITIVE ERGODIC THEOREMS FOR RANDOM SETS IN INFINITE DIMENSIONS

JENNIE C. HANSEN¹ *and* PAUL HULSE²

Abstract

We prove pointwise and mean versions of the subadditive ergodic theorem for superstationary families of compact, convex random subsets of a real Banach space, extending previously known results that were obtained in finite dimensions or with additional hypotheses on the random sets. We also show how the techniques can be used to obtain the strong law of large numbers for pairwise independent random sets, as well as results in the weak topology.

1. Introduction

One of the most fundamental questions of probability theory concerns the asymptotic behaviour of sequences of random variables. The classic result is the strong law of large numbers for independent, identically distributed sequences. This was superseded by Birkhoff's ergodic theorem for stationary sequences. A substantial generalisation of Birkhoff's result was obtained by Kingman [10], who proved an ergodic theorem for subadditive stationary processes. Further generalisations involved relaxing the stationarity properties: Krengel [11] extended Birkhoff's theorem to superstationary sequences, then Abid [1] obtained Kingman's theorem for subadditive superstationary processes. In another direction, Liggett [13] obtained Kingman's result under weaker subadditivity conditions. There is also a corresponding relaxation of the stationarity conditions, although not to include superstationary processes. However, it is worth noting that, with no more than the obvious modifications to the proof, one can obtain a superstationary version of Liggett's result, which includes Abid's result as a special case. Further generalisations were obtained by Derriennic [6] and Schürger [16], who considered different notions of almost subadditivity.

Various analogues of these results have been obtained for random sets. Artstein and Vitale [3] proved a strong law for random compact subsets of Euclidean space. Puri and Ralescu [14] extended this to any Banach space, under the additional assumption that the sets are convex. Artstein and Hansen [2] deduced the non-convex case, reducing it to the convex case by a convexification lemma. The same lemma was obtained independently by Hiai [9]; it is quite general and can be applied, for example, to the ergodic theorem obtained by Hess [8] for the convex case.

In [15], Schürger obtained pointwise and mean versions of the subadditive ergodic theorem for superstationary families of random compact convex subsets in finite dimensions. The extra assumption of convexity is necessary in this case, as Schürger demonstrated. Indeed, the example given shows that the assumption is needed even when the sets are non-random. Whereas the idea of the proofs of the strong law in [3] and [14] was to embed the sets in a suitable Banach space and use the strong law for vector-valued processes, Schürger used support functions to reduce the problem to an application of Abid's version of the subadditive ergodic theorem for random variables. More

¹ Dept. of Actuarial Mathematics & Statistics, Heriot-Watt University, Edinburgh EH14 4AS, UK.

² 64, Curriehill Castle Drive, Balerno, Midlothian EH14 5TD, UK.

recently, Krupa [12] used a similar approach to extend Schürger's results to infinite dimensions, but with an additional compactness condition on the random sets.

In this paper, we extend these results to obtain pointwise and mean ergodic theorems valid for any subadditive superstationary collection of random compact convex subsets of a real Banach space; in particular, the Banach space can be infinite-dimensional and there is no need for any additional conditions on the random sets. The general approach is similar to that of Schürger and Krupa in that the problem is reduced to an application of the subadditive ergodic theorem for random variables. However, it differs in that this is done using a distance function instead of support functions. This gives a more elementary proof in the sense that it uses only basic properties of the distance function, whereas the support function approach relies ultimately on the Hahn-Banach Theorem, or some equivalent result. The key to obtaining a more general result, however, is a compactness lemma (Lemma 3.2), which says, in effect, that any sequence of random compact sets with the appropriate stationarity properties is contained (with probability 1) in a compact set. This and other preliminary results needed to carry out the reduction are in §2 and §3. The statement and proofs of the theorems are in §4. We note that our subadditivity and stationarity hypotheses are also weaker than those in [12] and [15]. This is accounted for by the fact that the subadditivity and stationarity hypotheses needed to prove an ergodic theorem for random sets are essentially determined by the corresponding theorem for random variables. Thus, we are able to relax these hypotheses by using the superstationary version of Liggett's result instead of Abid's result.

While the main purpose of this paper is to prove the aforementioned subadditive theorems, the methods used are fairly general and can be applied in other contexts to extend results for random variables to random sets. For example, the same approach can be used to obtain analogues of the results in [6] and [16] for almost subadditive processes. In the remainder of the paper, we consider two other such situations.

In general, the ergodic theorem and strong law can be viewed as corollaries of the subadditive ergodic theorem. However, this is not always the case; for example, Etemadi [7] has shown that the strong law holds for pairwise independent, identically distributed random variables, but no corresponding subadditive theorem is known. Our methods apply equally well in such a situation, and in §5 we show how Etemadi's result can be obtained for random sets. One difference with the subadditive case is that, in order to preserve additivity, we have to use linear functionals instead of the distance function, effectively equivalent to using support functions.

In all the results mentioned so far, pointwise convergence has been in the Hausdorff metric. If one is willing to accept weaker forms of convergence, then results can be obtained for a wider class of random sets. For example, the scalar convergence of subadditive families of weakly compact random sets was studied by Krupa in [12]. In §6, we outline how our methods can be applied in this situation.

2. The space $(\mathcal{C}(B), \rho)$

Throughout this paper, B denotes a real Banach space with norm $\|\cdot\|$, and $\mathcal{C}(B)$ denotes the set of non-empty compact subsets of B . Addition and scalar multiplication on $\mathcal{C}(B)$ are defined by

$$C + D = \{x + y : x \in C, y \in D\}, \quad \alpha C = \{\alpha x : x \in C\} \quad (C, D \in \mathcal{C}(B), \alpha \in \mathbb{R}).$$

The Hausdorff metric ρ on $\mathcal{C}(B)$ is defined as follows. For $C, D \in \mathcal{C}(B)$, let

$$d(C, D) = \inf\{\varepsilon : C \subseteq D + B(\varepsilon)\},$$

where $B(\varepsilon) = \{x \in B : \|x\| \leq \varepsilon\}$; then

$$\rho(C, D) = \sup\{d(C, D), d(D, C)\}.$$

The metric space $(\mathcal{C}(B), \rho)$ is complete.

The closure of $C \subseteq B$ is denoted by $cl C$, and the convex hull by $co C$. If $C \in \mathcal{C}(B)$, then $\|C\| = \sup\{\|x\| : x \in C\}$. The mapping $C \rightarrow \|C\|$ from $\mathcal{C}(B)$ into \mathbb{R} is continuous.

The following result was proved for convex sets in [5]. For completeness, we include a proof; note that it is valid for non-convex sets.

Lemma 2.1. *If $K \in \mathcal{C}(B)$ then $\{C \in \mathcal{C}(B) : C \subseteq K\}$ is compact in $\mathcal{C}(B)$.*

Proof. Let $\{C_n\}$ be a sequence in $\{C \in \mathcal{C}(B) : C \subseteq K\}$; we show that $\{C_n\}$ is sequentially compact. Given $\varepsilon > 0$, let $\{B_j\}_{j=1}^k$ be a finite cover of K by open balls of radius ε . We construct subsequences $\{n_i^{(j)}\}_{i=1}^\infty$, $1 \leq j \leq k$, as follows: given $\{n_i^{(j-1)}\}$, choose $\{n_i^{(j)}\} \subseteq \{n_i^{(j-1)}\}$ such that either $B_j \cap C_{n_i^{(j)}} \neq \emptyset$ for all $i \geq 1$, or $B_j \cap C_{n_i^{(j)}} = \emptyset$ for all $i \geq 1$. Then for each $1 \leq j \leq k$, either $B_j \cap C_{n_i^{(k)}} \neq \emptyset$ for all $i \geq 1$, or $B_j \cap C_{n_i^{(k)}} = \emptyset$ for all $i \geq 1$. Let D be the union of all the B_j such that $B_j \cap C_{n_i^{(k)}} \neq \emptyset$ for all $i \geq 1$ (note that D is non-empty, since the B_j 's cover K). Clearly, $\rho(C_{n_i^{(k)}}, D) \leq 2\varepsilon$ for all $i \geq 1$, and so $\rho(C_{n_i^{(k)}}, C_{n_j^{(k)}}) \leq 4\varepsilon$ for all $i, j \geq 1$. Note that this construction works for any sequence in $\{C \in \mathcal{C}(B) : C \subseteq K\}$ and any $\varepsilon > 0$; thus we can use a standard diagonalisation argument to obtain a Cauchy subsequence of $\{C_n\}$. Since $\mathcal{C}(B)$ is complete, the subsequence is convergent, and since K is closed, the limit is a subset of K . This proves the result.

Lemma 2.2. *Let \mathcal{K} be a closed subset of $\mathcal{C}(B)$, and let $\{K_n\}_{n=1}^\infty$ be a sequence in \mathcal{K} such that $cl(\bigcup_n K_n)$ is compact, and $\{d(K_n, C)\}$ is convergent for all $C \in \mathcal{K}$. Then $\{K_n\}$ is convergent in $(\mathcal{C}(B), \rho)$.*

Proof. It follows from Lemma 2.1 that $\{C \in \mathcal{C}(B) : C \subseteq cl(\bigcup_n K_n)\}$ is compact in $\mathcal{C}(B)$. Therefore, there exists a subsequence $\{n_i\}$ and $C \in \mathcal{K}$ such that $K_{n_i} \rightarrow C$ as $i \rightarrow \infty$. Since $\{d(K_n, C)\}$ is convergent, it follows that $\lim_{n \rightarrow \infty} d(K_n, C) = 0$. We claim that $\lim_{n \rightarrow \infty} d(C, K_n) = 0$ also, in which case, $K_n \rightarrow C$ as $n \rightarrow \infty$. Suppose not; then there exists $\varepsilon > 0$, $D \in \mathcal{K}$ and a subsequence $\{m_i\}$ such that $d(C, K_{m_i}) \geq \varepsilon$ for all i , and $K_{m_i} \rightarrow D$ as $i \rightarrow \infty$. It follows that $d(C, D) \geq \varepsilon$; moreover,

$$d(C, D) \leq d(C, K_{n_i}) + d(K_{n_i}, D),$$

and so $\liminf_{i \rightarrow \infty} d(K_{n_i}, D) \geq \varepsilon$, whereas $\lim_{i \rightarrow \infty} d(K_{m_i}, D) = 0$. This is a contradiction; hence $\lim_{n \rightarrow \infty} d(C, K_n) = 0$, as required.

3. Random sets

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space. A *random compact set* is a mapping $X : \Omega \rightarrow \mathcal{C}(B)$ which is measurable with respect to the Borel σ -algebra of $(\mathcal{C}(B), \rho)$. We note that if $E(\|X\|) < \infty$, there is a separable subspace of B which contains, with probability one, all the values taken by X .

It is implicitly assumed throughout that any given countable collection of random (compact) sets is defined on a common probability space $(\Omega, \mathcal{F}, \mu)$ and takes values in $\mathcal{C}(B)$. Furthermore, in view of the previous remark, we can assume without loss of generality that B is separable, in which case, so is $\mathcal{C}(B)$.

Recall that a sequence of random variables $\{X_n\}_{n=1}^\infty$ (on $(\Omega, \mathcal{F}, \mu)$) is *superstationary* if $\int f(X_1, X_2, \dots) d\mu \geq \int f(X_2, X_3, \dots) d\mu$ for all bounded Borel functions $f : \mathbb{R}^\mathbb{N} \rightarrow \mathbb{R}$ which are increasing with respect to the usual partial ordering on $\mathbb{R}^\mathbb{N}$ (that is, $x \leq y$ if $x_i \leq y_i$ for all $i \in \mathbb{N}$). Set inclusion induces a similar partial ordering on $\mathcal{C}(B)^\mathbb{N}$: if $C = (C_1, C_2, \dots)$, $D = (D_1, D_2, \dots) \in \mathcal{C}(B)^\mathbb{N}$

$\mathcal{C}(B)^\mathbb{N}$, then $C \leq D$ if $C_i \subseteq D_i$ for all $i \in \mathbb{N}$. A sequence of random sets $\{X_n\}_{n=1}^\infty$ is *superstationary* if $\int f(X_1, X_2, \dots) d\mu \geq \int f(X_2, X_3, \dots) d\mu$ for all increasing, bounded Borel functions $f: \mathcal{C}(B)^\mathbb{N} \rightarrow \mathbb{R}$.

We have the following corollary to Lemma 2.2.

Lemma 3.1. *Let \mathcal{K} be a closed subset of $\mathcal{C}(B)$, and let $\{Y_n\}$ be a sequence of \mathcal{K} -valued random sets such that $cl(\bigcup_n Y_n)$ is compact a.e. If $\{d(Y_n, C)\}$ is convergent a.e. whenever $C \in \mathcal{K}$, then $\{Y_n\}$ is convergent in $(\mathcal{C}(B), \rho)$ a.e.*

Proof. Let \mathcal{D} be a countable, dense subset of \mathcal{K} , and choose $\Omega' \subseteq \Omega$, $\mu(\Omega') = 1$, so that $cl(\bigcup_n Y_n)$ is compact and $\{d(Y_n, C)\}$ is convergent for all $\omega \in \Omega'$ and $C \in \mathcal{D}$. Since $|d(Y_n, C) - d(Y_n, C')| \leq \rho(C, C')$ for any $C, C' \in \mathcal{C}(B)$, it follows that $\{d(Y_n, C)\}$ is convergent for all $C \in \mathcal{K}$ whenever $\omega \in \Omega'$. Then by Lemma 2.2, $\{Y_n\}$ is convergent for all $\omega \in \Omega'$.

The next result is the key to removing the need for additional compactness conditions on the random sets.

Lemma 3.2. *Let $\{X_n\}_{n=1}^\infty$ be a superstationary sequence of $\mathcal{C}(B)$ -valued random sets such that $E(\|X_1\|) < \infty$. Then $cl(\bigcup_{n=1}^\infty n^{-1} \sum_{i=1}^n X_i)$ is compact a.e.*

Proof. Let $\{x_k\}_{k=0}^\infty$ be a countable dense subset of B , with $x_0 = 0$. Let $C_k = co\{x_0, \dots, x_k\}$ ($k \geq 1$) and $d_k(C) = d(C, C_k)$ ($C \in \mathcal{C}(B)$). Then for each k , $\{d_k(X_n)\}_{n=1}^\infty$ is a superstationary sequence of random variables, and $d_k(X_1) \leq \|X_1\|$ (since $0 \in C_k$). Moreover, since X_1 is compact a.e., $\lim_{k \rightarrow \infty} d_k(X_1) = 0$ a.e.; hence, $\lim_{k \rightarrow \infty} E(d_k(X_1)) = 0$. Given $\varepsilon > 0$, choose k such that $E(d_k(X_1)) \leq \varepsilon^2/4$. By the superstationary ergodic theorem [11], there exists $f \in L^1(\Omega, \mathcal{F}, \mu)$ such that $n^{-1} \sum_{i=1}^n d_k(X_i) \rightarrow f$ a.e. (μ) . By Ergoroff's theorem, there exists N such that

$$\mu\left\{\omega : \sup_{n \geq N} \left| n^{-1} \sum_{i=1}^n d_k(X_i(\omega)) - f \right| > \varepsilon/2 \right\} \leq \varepsilon/2.$$

Moreover, $E(f) \leq \varepsilon^2/4$, and so $\mu\{f > \varepsilon/2\} \leq \varepsilon/2$. Therefore,

$$\mu\left\{\omega : \sup_{n \geq N} n^{-1} \sum_{i=1}^n d_k(X_i(\omega)) > \varepsilon \right\} \leq \varepsilon,$$

and so, since C_k is convex,

$$\mu\left\{\omega : \sup_{n \geq N} d_k(n^{-1} \sum_{i=1}^n X_i(\omega)) > \varepsilon \right\} \leq \varepsilon.$$

Hence, if

$$A_\varepsilon = \bigcup_{k, N} \{\omega : \bigcup_{n=N}^\infty n^{-1} \sum_{i=1}^n X_i(\omega) \subseteq C_k + B(\varepsilon)\},$$

then $\mu(A_\varepsilon) \geq 1 - \varepsilon$, and $\varepsilon \leq \varepsilon' \Rightarrow A_\varepsilon \subseteq A_{\varepsilon'}$. Thus, $\mu(\bigcap_{\varepsilon > 0} A_\varepsilon) = 1$. Now, given $\varepsilon > 0$ and $\omega \in \bigcap_{\varepsilon > 0} A_\varepsilon$, choose k, N such that

$$\bigcup_{n=N}^\infty n^{-1} \sum_{i=1}^n X_i(\omega) \subseteq C_k + B(\varepsilon/2).$$

Since C_k is totally bounded, it follows that $C_k + B(\varepsilon/2)$ is contained in a finite number of balls of radius ε . Clearly, $\bigcup_{n=1}^{N-1} n^{-1} \sum_{i=1}^n X_i$ is contained in a finite number of balls of radius ε , and so, since ε is arbitrary, $\bigcup_{n=1}^\infty n^{-1} \sum_{i=1}^n X_i$ is totally bounded, and the result follows.

Remark. The above proof could be simplified slightly by using the maximal ergodic theorem instead of the ergodic theorem and Ergoroff's theorem. However, the latter approach illustrates the point that Lemma 3.2 holds whenever $n^{-1} \sum_{i=1}^n d_k(X_i)$ converges a.e. ($k \geq 1$).

4. The subadditive ergodic theorem

We start by defining what we mean by a subadditive superstationary process. As before, all countable collections of random variables and random sets are assumed to be defined on a common probability space $(\Omega, \mathcal{F}, \mu)$. A collection of random variables $\{X_{m,n}\}_{0 \leq m < n}$ ($m, n \in \mathbb{Z}$) is *subadditive* if $X_{0,n} \leq X_{0,m} + X_{m,n}$ a.e. for all $0 < m < n$; and $\{X_{m,n}\}_{0 \leq m < n}$ is *superstationary* if

- (i) $\{X_{(m-1)k, mk}\}_{m=1}^\infty$ is a superstationary sequence for each $k \geq 1$,
- (ii) for each $m \geq 0$, the joint distributions of $\{X_{m, m+n}\}_{n=1}^\infty$ dominate those of $\{X_{m+1, m+1+n}\}_{n=1}^\infty$; that is, $\int f(X_{m, m+1}, X_{m, m+2}, \dots) d\mu \geq \int f(X_{m+1, m+2}, X_{m+1, m+3}, \dots) d\mu$ for all increasing, bounded Borel functions $f: \mathbb{R}^\mathbb{N} \rightarrow \mathbb{R}$.

The following subadditive ergodic theorem was proved by Liggett in [13] under more restrictive stationarity hypotheses; however, we note that the same proof with no more than the obvious modifications works for this version.

Theorem 4.1. *If $\{X_{m,n}\}_{0 \leq m < n}$ is a subadditive, superstationary collection of random variables such that $E|X_{0,n}| < \infty$ for all n , and there exists a constant c such that $E(X_{0,n}) \geq -cn$ for all n , then $\lim_{n \rightarrow \infty} n^{-1}X_{0,n}$ exists a.e. If the sequences $\{X_{mk, (m+1)k}\}_{m=1}^\infty$ ($k \geq 1$) are ergodic, then the limit is constant a.e.*

Now let $\{X_{m,n}\}_{0 \leq m < n}$ be a collection of random sets. Then $\{X_{m,n}\}_{0 \leq m < n}$ is *subadditive* if $X_{0,n} \subseteq X_{0,m} + X_{m,n}$ a.e. for all $0 < m < n$, and *superstationary* if

- (i) $\{X_{(m-1)k, mk}\}_{m=1}^\infty$ is a superstationary sequence for each $k \geq 1$,
- (ii) for each $m \geq 0$, the joint distributions of $\{X_{m, m+n}\}_{n=1}^\infty$ dominate those of $\{X_{m+1, m+1+n}\}_{n=1}^\infty$; that is, $\int f(X_{m, m+1}, X_{m, m+2}, \dots) d\mu \geq \int f(X_{m+1, m+2}, X_{m+1, m+3}, \dots) d\mu$ for all increasing, bounded Borel functions $f: \mathcal{C}(B)^\mathbb{N} \rightarrow \mathbb{R}$.

Let $\mathcal{K}(B) = \{C \in \mathcal{C}(B) : C \text{ is convex}\}$; note that $\mathcal{K}(B)$ is a closed subset of $\mathcal{C}(B)$.

Theorem 4.2. *Let $\{X_{m,n}\}_{0 \leq m < n}$ be a subadditive, superstationary collection of $\mathcal{K}(B)$ -valued random sets such that $E\|X_{0,1}\| < \infty$. Then there exists a $\mathcal{K}(B)$ -valued random set X_∞ such that $n^{-1}X_{0,n} \rightarrow X_\infty$ in $(\mathcal{C}(B), \rho)$ a.e. as $n \rightarrow \infty$.*

Proof. Note that $X_{0,n} \subseteq \sum_{i=1}^n X_{i-1,i}$ and $\{X_{n-1,n}\}_{n=1}^\infty$ is a superstationary sequence. Hence, by Lemma 3.2, $cl(\bigcup_{n=1}^\infty n^{-1}X_{0,n})$ is compact. Therefore, by Lemma 3.1, it is enough to show that $\{d(n^{-1}X_{0,n}, C)\}$ is convergent a.e., for all $C \in \mathcal{K}(B)$. To this end, note that the random variables $\{d(X_{m,n}, (n-m)C)\}_{0 \leq m < n}$ are subadditive and superstationary. The superstationarity follows immediately from that of $\{X_{m,n}\}_{0 \leq m < n}$, since $d(\cdot, C')$ is an increasing function for any $C' \in \mathcal{C}(B)$. For subadditivity, note that if $0 < m < n$ and $x \in X_{0,n}$, there exist $x_1 \in X_{0,m}$, $x_2 \in X_{m,n}$ such that $x = x_1 + x_2$. Then for any $c_1, c_2 \in C$, we have $c = n^{-1}(mc_1 + (n-m)c_2) \in C$ since C is convex, and $\|x - nc\| \leq \|x_1 - mc_1\| + \|x_2 - (n-m)c_2\|$. It follows that

$$d(X_{0,n}, nC) \leq d(X_{0,m}, mC) + d(X_{m,n}, (n-m)C).$$

Moreover, $d(X_{0,n}, nC) \leq \|X_{0,n}\| + \|nC\|$ and $\|X_{0,n}\| \leq \sum_{i=1}^n \|X_{i-1,i}\|$ a.e. Hence,

$$E(d(X_{0,n}, nC)) \leq n(E(\|X_{0,1}\|) + \|C\|) < \infty.$$

Therefore, by Theorem 4.1, $d(n^{-1}X_{0,n}, C)$ converges, and the result follows.

The question arises as to when the limit in Theorem 4.2 is constant. It follows from the more or less obvious observation that $C = C'$ ($C, C' \in \mathcal{K}(B)$) if and only if $d(C, D) = d(C', D)$ for all $D \in \mathcal{K}(B)$, that X_∞ is constant a.e. provided there is a set of measure 1 on which $d(n^{-1}X_{0,n}, C)$

converges to a constant for each $C \in \mathcal{K}(B)$. This will be the case if it is true for all C in some countable, dense subset of $\mathcal{K}(B)$. Hence, it follows from Theorem 4.1 that X_∞ is constant a.e. if the sequences $\{d(X_{mk, (m+1)k}, kC)\}_{m=1}^\infty$ ($k \geq 1, C \in \mathcal{K}(B)$) are ergodic.

As a corollary to Theorem 4.2, we have the following mean ergodic theorem.

Theorem 4.3. *Let $\{X_{m,n}\}_{0 \leq m < n}$ and X_∞ be as in Theorem 4.2, and let $1 \leq p < \infty$. If $E(\|X_{0,1}\|^p) < \infty$, then*

$$\lim_{n \rightarrow \infty} \int \rho(n^{-1}X_{0,n}, X_\infty)^p d\mu = 0.$$

Proof. Given $\varepsilon > 0$, there exists $A \in \mathcal{F}$ and $N \in \mathbb{N}$ such that $\mu(A) \geq 1 - \varepsilon$ and $\rho(n^{-1}X_{0,n}, X_\infty) \leq \varepsilon$ for all $n \geq N$ and $\omega \in A$. Then $\int_A \rho(n^{-1}X_{0,n}, X_\infty)^p d\mu \leq \varepsilon^p(1 - \varepsilon)$. Now,

$$\left(\int_{A^c} \rho(n^{-1}X_{0,n}, X_\infty)^p d\mu \right)^{1/p} \leq \left(\int_{A^c} \|n^{-1}X_{0,n}\|^p d\mu \right)^{1/p} + \left(\int_{A^c} \|X_\infty\|^p d\mu \right)^{1/p}.$$

Note that $\|n^{-1}X_{0,n}\| \leq n^{-1} \sum_{i=1}^n \|X_{i-1,i}\|$ a.e., and $\{\|X_{n-1,n}\|\}_{n=1}^\infty$ is a superstationary sequence of random variables. Hence, by the mean ergodic theorem for superstationary processes (which can easily be deduced from Krengel's pointwise result [11]), there exists $f \in L^p(\Omega, \mathcal{F}, \mu)$ such that $\int |f|^p d\mu \leq E(\|X_{0,1}\|^p)$ and $n^{-1} \sum_{i=1}^n \|X_{i-1,i}\|$ converges to f in $L^p(\Omega, \mathcal{F}, \mu)$. Therefore,

$$\limsup_{n \rightarrow \infty} \int_{A^c} \|n^{-1}X_{0,n}\|^p d\mu \leq \int_{A^c} f^p d\mu.$$

Moreover, by Fatou's Lemma,

$$\int_{A^c} \|X_\infty\|^p d\mu \leq \liminf_{n \rightarrow \infty} \int_{A^c} \|n^{-1}X_{0,n}\|^p d\mu,$$

and so

$$\limsup_{n \rightarrow \infty} \left(\int_{A^c} \rho(n^{-1}X_{0,n}, X_\infty)^p d\mu \right)^{1/p} \leq 2 \left(\int_{A^c} f^p d\mu \right)^{1/p}.$$

Letting $\varepsilon \rightarrow 0$ gives the result.

As mentioned in the introduction, weaker notions of subadditivity have been studied: Derriennic [6] and Schürger [16] have both obtained ergodic theorems for random variables which are almost subadditive, that is, subadditive with respect to another family of random variables. The above methods can be used to obtain versions of these theorems for random sets. For example, in [16], $\{X_{m,n}\}_{0 \leq m < n}$ is said to be subadditive with respect to $\{U_{m,n}\}_{0 \leq m < n}$ if

$$X_{m_0, m_k} \leq \sum_{i=1}^k (X_{m_{i-1}, m_i} + U_{m_{i-1}, m_i})$$

for all $m_0 < \dots < m_k$. If almost subadditivity of random sets is defined analogously and the $\mathcal{K}(B)$ -valued random sets $\{X_{m,n}\}_{0 \leq m < n}$ are subadditive with respect to the $\mathcal{C}(B)$ -valued $\{U_{m,n}\}_{0 \leq m < n}$, where $\lim_{n \rightarrow \infty} n^{-1}U_{0,n} = \{0\}$ a.e., then, for any $C \in \mathcal{K}(B)$, $\{d(X_{m,n}, (n-m)C)\}_{0 \leq m < n}$ is subadditive with respect to $\{d(U_{m,n}, \{0\})\}_{0 \leq m < n}$ and hence (under suitable conditions) $\{d(n^{-1}X_{0,n}, C)\}$ is convergent a.e. Moreover,

$$X_{m,n} \subseteq \sum_{i=m+1}^n (X_{i-1,i} + U_{i-1,i})$$

for all $0 \leq m < n$. Thus, under suitable stationarity conditions, one can apply Lemma 3.2 to the sequence $\{X_{i-1,i} + U_{i-1,i}\}$ and, as in the proof of Theorem 4.2, deduce the convergence of $n^{-1}X_{0,n}$.

5. The strong law

If $\{X_n\}$ is a superstationary sequence of random compact sets, it follows immediately from Theorem 4.2 (and the convexification result in [2] and [9]) that $n^{-1} \sum_{i=1}^n X_i$ converges a.e. However, there are situations where these averages converge but there is no corresponding version of the subadditive ergodic theorem from which this can be deduced. One such situation is when the X_n are identically distributed and pairwise independent (i.e., X_i and X_j are independent for all $i, j \in \mathbb{N}$, $i \neq j$). In the corresponding case for random variables, Etemadi [7] has shown that the strong law holds. The techniques of §3 can be used to extend this result to random sets.

Theorem 5.1. *Let $\{X_n\}_{n=1}^\infty$ be a sequence of pairwise independent, identically distributed random compact sets, such that $E(\|X_1\|) < \infty$. Then there exists $C \in \mathcal{K}(B)$ such that $n^{-1} \sum_{i=1}^n X_i \rightarrow C$ a.e. as $n \rightarrow \infty$.*

Proof. It follows from the convexification result in [2] and [9] that it is enough to prove the result when the X_n are convex, and that any limit takes values in $\mathcal{K}(B)$. Note that if $f \in B^*$, then $\sup f(C + D) = \sup f(C) + \sup f(D)$ and $\sup f(\alpha C) = \alpha \sup f(C)$ for all $C, D \in \mathcal{C}(B)$, $\alpha > 0$. Moreover, the random variables $\{\sup f(X_n)\}$ are pairwise independent and identically distributed, as are the $\{d_k(X_n)\}_{n=1}^\infty$ ($k \geq 1$), defined in the proof of Lemma 3.2. Let Λ be a countable, dense subset of B^* . Then it follows from Etemadi's result [7] and the remark after Lemma 3.2 that there exists $\Omega' \subseteq \Omega$, $\mu(\Omega') = 1$, such that, for all $f \in \Lambda$ and $\omega \in \Omega'$, $n^{-1} \sum_{i=1}^n \sup f(X_i)$ converges to a constant and $cl(\bigcup_n n^{-1} \sum_{i=1}^n X_i)$ is compact. Fix $\omega \in \Omega'$. By Lemma 2.1, $\{C \in \mathcal{C}(B) : C \subseteq cl(\bigcup_n n^{-1} \sum_{i=1}^n X_i)\}$ is compact in $\mathcal{C}(B)$. Therefore, there exists a convergent subsequence of $n^{-1} \sum_{i=1}^n X_i$. Suppose $C, D \in \mathcal{K}(B)$ are limit points for the subsequences $\{m_k\}, \{n_k\}$ respectively. Then for all $f \in \Lambda$, $m_k^{-1} \sum_{i=1}^{m_k} \sup f(X_i) \rightarrow \sup f(C)$ and $n_k^{-1} \sum_{i=1}^{n_k} \sup f(X_i) \rightarrow \sup f(D)$ as $k \rightarrow \infty$, and so, $\sup f(C) = \sup f(D)$. It follows from the Hahn-Banach theorem that $C = D$, and so $n^{-1} \sum_{i=1}^n X_i$ is convergent. It also follows that if the limit is X_∞ , then X_∞ is constant a.e., because $\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \sup f(X_i) = \sup f(X_\infty)$ is constant on Ω' , for all $f \in \Lambda$.

Remark. The limit in Theorem 5.1 is $E(coX_1)$; for the definition of the expectation of a random set, see [4].

6. Scalar convergence

A sequence $\{C_n\}_{n=1}^\infty$ of subsets of B converges to C in the *scalar topology* if $\lim_{n \rightarrow \infty} s(f, C_n) = s(f, C)$ for all $f \in B^*$, where $s(\cdot, \cdot)$ is the support function; that is, $s(f, C) = \sup_{x \in C} f(x)$ ($f \in B^*$, $C \subseteq B$). We can deduce a scalar convergence version of Theorem 4.2 for subadditive families of weakly compact random sets. If $\{X_{m,n}\}_{0 \leq m < n}$ is a subadditive, superstationary collection of weakly compact, convex random sets such that $E(\|X_{0,1}\|) < \infty$, and $f \in B^*$, then it follows immediately from Theorem 4.1 that $\sup n^{-1} f(X_{0,n})$ converges a.e., as does $n^{-1} \|X_{0,n}\|$. By considering a countable, dense subset of B^* and using a standard approximation argument, we deduce that there is a subset Ω' of measure 1 on which $\sup n^{-1} f(X_{0,n})$ is convergent for all $f \in B^*$. The convergence of $\{n^{-1} X_{0,n}\}$, on a set of measure 1, follows from weak compactness versions of Lemma 2.1 (see [5, Lemma 3.2]) and Lemma 3.2. Since B^* is separable, the set of weakly compact subsets of B with the scalar topology is metrisable, and so the required version of Lemma 3.2 can be obtained by replacing the function $d(\cdot, \cdot)$ by a suitable metric in the proof of Lemma 3.2.

References

1. M. Abid, Un théorème ergodique pour des processus sous-additifs et sur-stationnaires, *C.R. Acad. Sci. Paris, Série A*, 287 (1978), 149-152.
2. Z. Artstein & J. C. Hansen, Convexification in limit laws of random sets in Banach space, *Ann. Probab.* 13 (1985), 307-309.
3. Z. Artstein & R. Vitale, A strong law of large numbers for random sets, *Ann. Probab.* 3 (1975), 879-882.
4. R. J. Aumann, Integrals of set-valued functions, *J. Math. Analysis and Appl.* 12 (1965), 1-12.
5. E. J. Balder & C. Hess, Two generalizations of Komlós theorem with lower closure-type applications, *J. Convex Anal.* 3 (1996), 25-44.
6. Y. Derriennic, Un théorème ergodique presque sous-additif, *Ann. Probab.* 11 (1983), 669-677.
7. N. Etemadi, An elementary proof of the strong law of large numbers, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 55 (1981), 119-122.
8. C. Hess, Théorème ergodique et loi forte des grands nombres pour des ensembles aléatoires, *C.R. Acad. Sci. Paris, Série A*, 288 (1979), 519-522.
9. F. Hiai, Strong laws of large numbers for multivalued random variables, in *Multifunctions and Integrands*, Lecture Notes in Mathematics 1091, Springer, Berlin, 1984.
10. J. F. C. Kingman, The ergodic theory of subadditive processes, *J. Roy. Statist. Soc., Ser. B*, 30 (1968), 499-510.
11. U. Krengel, Un théorème ergodique pour les processus surstationnaires, *C.R. Acad. Sci. Paris, Série A*, 282 (1976), 1019-1021.
12. G. Krupa, Ergodic theorems for subadditive superstationary families of random sets with values in Banach spaces, *Studia Math.* 131 (1998), 289-302.
13. T. M. Liggett, An improved subadditive ergodic theorem, *Ann. Probab.* 13 (1985), 1279-1285.
14. M. L. Puri & D. A. Ralescu, Strong law of large numbers for Banach space valued random sets, *Ann. Probability* 11 (1983), 222-224.
15. K. Schürger, Ergodic theorems for subadditive superstationary families of convex compact random sets, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 62 (1983), 125-135.
16. K. Schürger, Almost subadditive extensions of Kingman's ergodic theorem, *Ann. Probab.* 19 (1991), 1575-1586.