The Law of Large Numbers in a Metric Space with a Convex Combination Operation

Pedro Terán[1] **and Ilya Molchanov**[2,3]

We consider a separable complete metric space equipped with a convex combination operation. For such spaces, we identify the corresponding convexification operator and show that the invariant elements for this operator appear naturally as limits in the strong law of large numbers. It is shown how to uplift the suggested construction to work with subsets of the basic space in order to develop a systematic way of proving laws of large numbers for such operations with random sets.

1. INTRODUCTION

Many laws of large numbers for random elements in linear spaces can be extended to the setting of random elements in general metric spaces.[10,25,30] These generalisations begin with a definition of an averaging operation on a metric space. At this stage some conditions on the metric space are imposed to ensure that the defined operation yields a non-empty result. The averaging operation is used to define expectation of random elements and eventually to prove the corresponding law of large numbers.

If convex combinations of points are well defined, it is possible to extend them to average sets, where the convex combination of sets is

---

[1] Facultad de Ciencias Económicas y Empresariales, Grupo Decisión Multicriterio Zaragoza, Universidad de Zaragoza, Gran Vía 2, E-50005 Zaragoza, Spain.
  E-mail: teran@unizar.es
[2] Department of Mathematical Statistics and Actuarial Sciences, University of Bern, Sidlerstr. 5, CH-3012 Bern, Switzerland. E-mail: ilya.molchanov@stat.unibe.ch
[3] To whom correspondence should be addressed.

defined elementwise as the set of all possible convex combinations of points from the sets in question. This issue is non-trivial even in linear spaces, where convex combinations of points are naturally defined. The corresponding convex combinations of sets are known under the name of Minkowski sums. It is well known that the Minkowski sums of sets convexify the summands, i.e. the sum is "more convex" than the summands (see Ref. 2). It is worth noticing that convex sets appear as limits in the law of large numbers, corresponding to the fact that convex sets are decomposable with respect to the Minkowski addition.

In this paper, we aim to single out the crucial properties of convex combination operations in metric spaces, derive from these properties the strong law of large numbers for normalised sums of random elements and characterise possible limits. A related issue concerns the properties of this operation uplifted to act on subsets of metric spaces.

Section 2 formalises the concept of the convex combination operation in a metric space $\mathbb{E}$, establishes its main algebraic properties and relationship to the metric on $\mathbb{E}$ that results in the requirement that $\mathbb{E}$ should be negatively curved with respect to the convex combination. Section 3 concerns the limits of the repeated convex combinations of $u \in \mathbb{E}$ with itself. The limits of these iterated convex combinations define a linear operator K on $\mathbb{E}$ that further appears in the definition of the expectation of $\mathbb{E}$-valued random elements in Section 4 and in the strong law of large numbers proved in Section 5. Section 6 describes the uplift of the convex combination operation extended pointwisely to subsets of $\mathbb{E}$.

In Section 7, we explore the connection between our framework and the concept of the Doss expectation in a metric space. As a by-product we show that the conventional (i.e. linear) convex combination operation in a Banach space is the unique operation such that the corresponding operator K is the identity. Section 8 shows that despite of some freedom of choice of the convex combination operation, the strong law of large numbers holds uniformly over all such operations. Section 9 lists several examples.

2. CONVEX COMBINATIONS

Let $(\mathbb{E}, d)$ be a metric space, endowed with a *convex combination operation* which for all $n \geqslant 2$, numbers $\lambda_1, \ldots, \lambda_n > 0$ satisfying $\sum_{i=1}^{n} \lambda_i = 1$, and all $u_1, \ldots, u_n \in \mathbb{E}$ this operation produces an element of $\mathbb{E}$, denoted by $[\lambda_i, u_i]_{i=1}^{n}$ or $[\lambda_1, u_1; \ldots; \lambda_n, u_n]$. Assume that $[1, u] = u$ for every $u \in \mathbb{E}$.

This operation makes it possible to consider situations where meaningful addition and product by scalars in $\mathbb{E}$ do not exist, but points

can nevertheless be averaged. This is the case of metric spaces where geometric considerations may allow to define averages.

It is assumed that the considered operation satisfies the following algebraic properties:

(i) $[\lambda_i, u_i]_{i=1}^n = [\lambda_{\sigma(i)}, u_{\sigma(i)}]_{i=1}^n$ for every permutation σ of $\{1, \ldots, n\}$;

(ii) $[\lambda_i, u_i]_{i=1}^{n+2} = [\lambda_1, u_1; \ldots; \lambda_n, u_n; \lambda_{n+1} + \lambda_{n+2}, [\{\frac{\lambda_{n+j}}{\lambda_{n+1}+\lambda_{n+2}}; u_{n+j}\}]_{j=1}^2]$.

Observe that (i) and (ii) imply that arbitrary subsets of indices can be reordered and regrouped. In the setting of a linear space $\mathbb{E}$ these properties correspond to commutativity, associativity, mixed associativity and the distributivity law $\lambda(u + v) = \lambda u + \lambda v$, $\lambda \geqslant 0$.

Lemma 2.1. For every $u_1, \ldots, u_{nm} \in \mathbb{E}$ and $\lambda_1, \ldots, \lambda_n, \alpha_1, \ldots, \alpha_m > 0$ with $\sum_{i=1}^n \lambda_i = \sum_{j=1}^m \alpha_j = 1$,

$$[\lambda_i, [\alpha_j, u_{ij}]_{j=1}^m]_{i=1}^n = [\lambda_i \alpha_j, u_{ij}]_{j=1, i=1}^{j=m, i=n}.$$

The following assumptions relate the introduced convex combination operation to the metric on $\mathbb{E}$:

(iii) If $u, v \in \mathbb{E}$ and $\lambda^{(k)} \to \lambda \in (0, 1)$ as $k \to \infty$, then

$$[\lambda^{(k)}, u; 1 - \lambda^{(k)}, v] \to [\lambda, u; 1 - \lambda, v];$$

(iv) $(\mathbb{E}, d)$ is *negatively curved* (with respect to the chosen convex combination) in the sense that

$$d([\lambda, u_1; 1 - \lambda, u_2], [\lambda, v_1; 1 - \lambda, v_2]) \leqslant \lambda d(u_1, v_1) + (1 - \lambda) d(u_2, v_2)$$

for all $u_1, u_2, v_1, v_2 \in \mathbb{E}$ and $\lambda \in (0, 1)$.

Since regroupings are possible, properties (iii) and (iv) extend by an easy recurrence scheme to convex combinations of $m \geqslant 3$ elements (cf Ref. 10 Prop. 1.1). The following result follows from (iii) and (iv).

Lemma 2.2. The convex combination operation is jointly continuous in its $2n$ arguments.

3. CONVEXIFICATION AND DECOMPOSABILITY

The following assumption concerns the limiting behaviour of the convex combination operation:

(v) For each $u \in \mathbb{E}$, there exists $\lim_{n \to \infty} [n^{-1}, u]_{i=1}^n$, which will be denoted by Ku (or $K_{\mathbb{E}} u$ if ambiguity arises).

For ease of reference, a metric space $\mathbb{E}$ with the convex combination operation satisfying properties (i)–(v) will be called a *convex combination space*. The result Ku of applying K to $u \in \mathbb{E}$ is called the *convexification* of u, and K is called the *convexification* operator.

It is easy to realise that (v) is a necessary condition to the strong law of large numbers applied for deterministic elements of $\mathbb{E}$. If $\mathbb{E}$ is a linear space, then K is the identity operator. In a general $\mathbb{E}$ it is well possible that $[n^{-1}, u]_{i=1}^n \neq u$ and, moreover, $Ku \neq u$. The key idea in our approach is to show that the operator K contains sufficient information on the geometry of $\mathbb{E}$ in order to characterise the limits in the law of large numbers for *random* elements in $\mathbb{E}$.

If (v) does not hold for all $u \in \mathbb{E}$, we can restrict ourselves to those u, for which the limit exists. It is easily checked (like in Proposition 3.1) that the domain of definition of K is closed under convex combinations.

Proposition 3.1.   The operator K is linear, that is

$$K([\lambda_j, u_j]_{j=1}^m) = [\lambda_j, Ku_j]_{j=1}^m .$$

Proof.   Lemma 2.1 yields that

$$\begin{aligned} K([\lambda_j, u_j]_{j=1}^m) &= \lim_{n \to \infty} [n^{-1}, [\lambda_j, u_j]_{j=1}^m]_{i=1}^n \\ &= \lim_{n \to \infty} [\lambda_j, [n^{-1}, u_j]_{i=1}^n]_{j=1}^m. \end{aligned}$$

It suffices to note that the limit equals $[\lambda_j, Ku_j]_{j=1}^m$ by continuity.   □

A point $u \in \mathbb{E}$ is called *convexely decomposable* if

$$u = [\lambda_i, u]_{i=1}^n$$

for all $n \geqslant 2$ and $\lambda_1, \ldots, \lambda_n > 0$ with $\sum \lambda_i = 1$.

Proposition 3.2.   The image $K(\mathbb{E})$ of $\mathbb{E}$ under K coincides with the family of convexely decomposable elements of $\mathbb{E}$.

Proof. It is clear that every convexely decomposable element belongs to $K(\mathbb{E})$. Let $m_1, \ldots, m_l, m$ be natural numbers such that $\sum m_i = m$. By Lemma 2.1 and continuity,

$$Ku = \lim_{n \to \infty} [(mn)^{-1}, u]_{i=1}^{mn} = \lim_{n \to \infty} [m_j m^{-1}, [(m_j n)^{-1}, u]_{i=1}^{m_j n}]_{j=1}^{l}$$
$$= [m_j m^{-1}, Ku]_{j=1}^{l}.$$

Hence, Ku is decomposable by rational convex combinations. By continuity, this property holds for arbitrary convex combinations. $\qquad \square$

Corollary 3.3. For all $n \geqslant 2$, $u \in \mathbb{E}$, and $\lambda_1, \ldots, \lambda_n > 0$ with $\sum \lambda_i = 1$ one has

$$K([\lambda_i, u]_{i=1}^{n}) = Ku = [\lambda_i, Ku]_{i=1}^{n}.$$

Corollary 3.4. K is an idempotent operator on $\mathbb{E}$.

Note that $[\lambda_1, u; \lambda_2, u]$ is not necessarily equal to u for $\lambda_1 + \lambda_2 = 1$ for $u \notin K(\mathbb{E})$. The following Proposition 3.5 establishes a variant of the distributivity law for convexely decomposable elements of $\mathbb{E}$. It easily extends to arbitrary subsets of indices by (i) and (ii).

Proposition 3.5. For every $\lambda_1, \lambda_2, \lambda_3 > 0$ with $\lambda_1 + \lambda_2 + \lambda_3 = 1$, and $u, v \in \mathbb{E}$,

$$[\lambda_1, u; \lambda_2, Kv; \lambda_3, Kv] = [\lambda_1, u; (\lambda_2 + \lambda_3), Kv].$$

Proof. By (ii) and Corollary 3.3,

$$[\lambda_1, u; \lambda_2, Kv; \lambda_3, Kv] = \left[\lambda_1, u; (\lambda_2 + \lambda_3), \left[\frac{\lambda_2}{\lambda_2 + \lambda_3}, Kv; \frac{\lambda_3}{\lambda_2 + \lambda_3}, Kv\right]\right]$$
$$= [\lambda_1, u; (\lambda_2 + \lambda_3), Kv].$$

$\qquad \square$

Proposition 3.6. K is non-expansive with respect to the metric d.

Proof. Fix $\varepsilon > 0$ and $u_1, u_2 \in \mathbb{E}$. For all sufficiently large n,

$$d(Ku_j, [n^{-1}, u_j]_{i=1}^{n}) < \varepsilon/2, \qquad j = 1, 2.$$

From (iv) it follows that $d(Ku_1, Ku_2) < d(u_1, u_2) + \varepsilon$. The result follows from the arbitrariness of ε. $\qquad \square$

A set $A \subset \mathbb{E}$ is said to be *convex* if $[\lambda_i, u_i]_{i=1}^n \in A$ for any $n \geqslant 2$, $u_1, \ldots, u_n \in A$ and positive numbers $\lambda_1, \ldots, \lambda_n$ that sum up to 1. The *convex hull*, co$A$, of $A \subset \mathbb{E}$ is the set of all possible convex combinations of points from A; the *closed convex hull* $\overline{\text{co}}\, A$ is the closure of co$A$ in $\mathbb{E}$. A number of elementary properties of convex sets in linear spaces still hold with the same proofs, e.g. the intersection of a family of convex sets is convex, and the closure of a convex set is convex.

Proposition 3.7. If V is a closed convex subset of $\mathbb{E}$, then $K(V)$ is closed convex and $K(V) \subset V$.

Proof. If $u \in V$, then $Ku = \lim_n [n^{-1}, u]_{i=1}^n \in V$, since V is convex closed. The convexity of $K(V)$ follows from that of V and Proposition 3.1. As for closedness, let $Ku_n \to u \in \mathbb{E}$ where $u_n \in V$. Since $K(V) \subset V$ is closed, $u \in V$. By Proposition 3.6, $K(Ku_n) \to Ku$. Since $K(Ku_n) = Ku_n$ by Corollary 3.4, the uniqueness of the limit yields $u = Ku \in K(V)$. $\qquad\square$

Proposition 3.8. Let $\mathbb{E}$ be a metric space such that (i)–(iv) hold. Then, the following are equivalent:

(a) property (v) holds with K being the identity mapping;
(b) all singletons are convex;
(c) all balls are convex.

Proof. The equivalence of (a) and (b) is trivial. It suffices to show that the convexity of all singletons implies the convexity of all balls. Consider a convex combination $[\lambda_i, u_i]_{i=1}^n$. Since the centre u of the ball is convex, one has $u = [\lambda_i, u]_{i=1}^n$. By (iv),

$$d([\lambda_i, u_i]_{i=1}^n, [\lambda_i, u]_{i=1}^n) \leqslant \sum_{i=1}^n \lambda_i d(u_i, u) \leqslant r,$$

where r is the radius of the ball. Therefore, the considered convex combination belongs to the ball. $\qquad\square$

4. CONSTRUCTION OF THE EXPECTATION

Here we define the expectation of a random element of $\mathbb{E}$ in a way which resembles the Lebesgue and Bochner integrals. From now on, we assume that $\mathbb{E}$ is separable and complete.

Let ξ be an $\mathbb{E}$-valued measurable simple function on a non-atomic probability space $(\Omega, \mathfrak{F}, \mathbf{P})$. That is, there exists a measurable partition $\{\Omega_1, \ldots, \Omega_m\}$ of Ω such that that ξ takes a constant value u_j on each non-null set Ω_j, $j = 1, \ldots, m$. Then the expectation of ξ is defined by

$$\mathbf{E}\xi = [\mathbf{P}(\Omega_i), Ku_i]_{i=1}^m. \tag{4.1}$$

By Proposition 3.5, the value of the expectation does not change if ξ is redefined as a step function on a finer partition. Note that the expectation is defined as the convex combination of Ku_i's and not u_i's themselves. Indeed, since the probability space is non-atomic, it is possible to partition any Ω_i further into any number $m \geqslant 2$ of subsets with equal probabilities. Then $[m^{-1}, u_i]_{j=1}^m$ converges to Ku_i as $m \to \infty$.

Fix any point $u_0 \in K(\mathbb{E})$ (by (v), $K(\mathbb{E})$ is non-empty). A random element $\xi \colon \Omega \to \mathbb{E}$ is said to be *integrable* if $d(u_0, \xi)$ is integrable real-valued random variable. Note that the definition does not depend on the chosen element u_0, since $|\mathbf{E}d(u_0, \xi) - \mathbf{E}d(v, \xi)| \leqslant d(u_0, v) < \infty$ for every $v \in K(\mathbb{E})$. The integrable functions form the space $L_{\mathbb{E}}^1$, where two functions ξ and η are considered to be equivalent if they coincide almost surely. The metric on $L_{\mathbb{E}}^1$ is defined by

$$\Delta(\xi, \eta) = \mathbf{E}d(\xi, \eta). \tag{4.2}$$

Since $\mathbb{E}$ is separable, every random element is almost surely approximable by a sequence of simple functions (see Ref. 29). Observe that simple functions form a subset of $L_{\mathbb{E}}^1$. As a consequence of property (iv) the operator $\mathbf{E}$ is non-expansive on simple functions and so it has a unique continuous (indeed non-expansive) extension onto the closure of simple functions in $L_{\mathbb{E}}^1$. This extension is called the expectation of ξ, and integrable ξ are exactly those given by limits of simple functions in the metric given by (4.2).

This extension makes it possible to calculate the expectation of every integrable random element of $\mathbb{E}$, as shown by the following proposition. Let $\{u_n, n \geqslant 1\}$ be a countable dense subset of $\mathbb{E}$ and let u_0 be the special element of $\mathbb{E}$ considered in the definition of $L_{\mathbb{E}}^1$. For each $k \geqslant 1$ define the mapping $\phi_k \colon \mathbb{E} \to \mathbb{E}$ such that $\phi_k(x) = u_{m_k(x)}$, where $m_k(x)$ is the smallest $i \in \{0, \ldots, k\}$ such that $d(u_i, x) = \min_{0 \leqslant j \leqslant k} d(u_j, x)$ and $d(u_0, u_i) \leqslant 2d(u_0, x)$.

Proposition 4.1. Let $\xi \in L_{\mathbb{E}}^1$. For each $k \geqslant 1$, the following statements hold:

(a) $\phi_k(\xi) \colon \Omega \to \mathbb{E}$ is a measurable simple function;
(b) $d(u_0, \phi_k(\xi)) \leqslant 2d(u_0, \xi)$;

(c) $d(\phi_k(\xi), \xi) \downarrow 0$ almost surely;
(d) $\Delta(\phi_k(\xi), \xi) \to 0$ as $k \to \infty$.

It is immediately seen from (4.1) that $\mathbf{E}\xi \in K(\mathbb{E})$ for a simple random element ξ. The following proposition extends this for any integrable ξ.

Proposition 4.2. If $\xi \in L^1_{\mathbb{E}}$, then $\mathbf{E}\xi$ is convexely decomposable.

Proof. By Proposition 3.7, $K(\mathbb{E})$ is a closed convex subset of $\mathbb{E}$. Convexity and (4.1) yield $\mathbf{E}\xi \in K(\mathbb{E})$ for simple ξ, then closedness extends it to integrable ξ. Finally, by Proposition 3.2, convexely decomposable elements of $\mathbb{E}$ are exactly those in $K(\mathbb{E})$. □

Let B be a measurable subset of Ω of a positive probability. The conditional expectation $\mathbf{E}(\xi|B)$ of a simple random element ξ is defined as $[\mathbf{P}\{\xi = u_j | B\}, u_j]_{u_j \in B}$. The conditional expectation of a general ξ is defined by approximation.

If $\Omega = \{\omega_1, \omega_2, \dots, \}$ is purely discrete, then the expectation of ξ is defined as the limit (if it exists) of $[\mathbf{P}(\{\omega_i\})/\mathbf{P}(\{\omega_1, \dots, \omega_n\}), u_i]_{i=1}^n$. If Ω contains both the atomic part $\Omega' = \{\omega_1, \omega_2, \dots, \}$ and the non-atomic part Ω'', then the expectation is defined as

$$\mathbf{E}\xi = [\mathbf{P}(\Omega'), \mathbf{E}(\xi|\Omega'); \mathbf{P}(\Omega''), \mathbf{E}(\xi|\Omega'')].$$

5. STRONG LAW OF LARGE NUMBERS

The developed framework makes it possible to derive the strong law of large numbers for $\mathbb{E}$-valued random elements from the algebraic properties of the convex combination operation on $\mathbb{E}$ and the corresponding convexification operator K.

Theorem 5.1. Let $\xi \in L^1_{\mathbb{E}}$, and let $\xi_1, \xi_2, \dots$ be a sequence of pairwise independent random elements distributed as ξ. Then

$$[n^{-1}, \xi_i]_{i=1}^n \to \mathbf{E}\xi \qquad \text{a.s. as } n \to \infty.$$

Proof. First assume that ξ is simple with values $u_1, \dots, u_m$ on non-null sets $\Omega_1, \dots, \Omega_m$. Define

$$B_{n,j} = \text{card}\{i : 1 \leqslant i \leqslant n, \xi_i = u_j\}.$$

By Etemadi's strong law of large numbers,[7] $B_{n,j}/n \to \mathbf{P}(\Omega_j) > 0$ almost surely as $n \to \infty$ for $j = 1, \ldots, m$. Outside a null set and for all sufficiently large n, all $B_{n,j}$ are strictly positive. By Lemma 2.1, grouping together all equal ξ_i,

$$[n^{-1}, \xi_i]_{i=1}^n = [B_{n,j}n^{-1}, [B_{n,j}^{-1}, u_j]_{i=1}^{B_{n,j}}]_{j=1}^m \,.$$

By Lemma 2.2,

$$[B_{n,j}n^{-1}, [B_{n,j}^{-1}, u_j]_{i=1}^{B_{n,j}}]_{j=1}^m \to [\mathbf{P}(\Omega_j), Ku_j]_{j=1}^m = \mathbf{E}\xi \,.$$

Let us consider a general integrable random element ξ. For each $\varepsilon > 0$, there exists some $k \geqslant 1$ such that $\Delta(\phi_k(\xi), \xi) \leqslant \varepsilon$. Then,

$$\begin{aligned} d([n^{-1}, \xi_i]_{i=1}^n, \mathbf{E}\xi) &\leqslant d([n^{-1}, \xi_i]_{i=1}^n, [n^{-1}, \phi_k(\xi_i)]_{i=1}^n) \\ &\quad + d([n^{-1}, \phi_k(\xi_i)]_{i=1}^n, \mathbf{E}\phi_k(\xi)) + d(\mathbf{E}\phi_k(\xi), \mathbf{E}\xi) \\ &= \mathbf{(I)} + \mathbf{(II)} + \mathbf{(III)} \,. \end{aligned}$$

First, $\mathbf{(III)} \leqslant \varepsilon$ by the non-expansiveness of $\mathbf{E}$ with respect to d and Δ. By property (iv) and Etemadi's strong law of large numbers,[7]

$$\mathbf{(I)} \leqslant \sum_{i=1}^n n^{-1} d(\phi_k(\xi_i), \xi_i) \to \mathbf{E}d(\phi_k(\xi), \xi) \leqslant \varepsilon$$

outside a null set. Finally, $\mathbf{(II)} \to 0$ out of a null set since $\phi_k(\xi)$ is simple.

Therefore, for all sufficiently large n,

$$d([n^{-1}, \xi_i]_{i=1}^n, \mathbf{E}\xi) \leqslant 3\varepsilon \,.$$

The proof is completed by considering a sequence of values $\varepsilon \downarrow 0$ and taking a countable union of the corresponding exceptional null sets. $\square$

Note that our results (in particular the law of large numbers) are still valid if $\mathbb{E}$ is taken to be Hausdorff with the topology generated by a countable family of pseudometrics such that (iv) holds for each pseudometric. This is shown by applying the results to each pseudometric and then disposing, if necessary, of a countable union of null sets.

6. AN UPLIFT TO SUBSETS OF $\mathbb{E}$

The convex combination operation defined on $\mathbb{E}$ can be naturally extended (*uplifted*) to act on subsets $A_1, \ldots, A_n$ of $\mathbb{E}$ by letting

$$[\lambda_i, A_i]_{i=1}^n = \mathrm{cl}\left\{[\lambda_i, u_i]_{i=1}^n : u_i \in A_i, \, i = 1, \ldots, n\right\}, \qquad (6.1)$$

where cl in the right-hand side denotes the closure in $\mathbb{E}$. Property (iv) implies that taking the closure in the right-hand side of (6.1) is not necessary if the sets $A_1, \ldots, A_n$ are compact. If $\mathbb{E}$ is a linear space with the linear convex combination operation, then the uplift (6.1) is the closure of the Minkowski sum of sets. Note that taking the closure is necessary, since the Minkowski sum of two non-compact closed sets is not necessarily closed.

The family $\mathcal{K}$ of non-empty compact subsets of $\mathbb{E}$ can be equipped with the Hausdorff metric d_{H} that is generated by the metric d on $\mathbb{E}$ as

$$d_{\mathrm{H}}(A_1, A_2) = \inf\{r \geqslant 0 : A_1 \subset A_2^r, \, A_2 \subset A_1^r\}, \qquad (6.2)$$

where the r-neighbourhood A^r is the set of all points $x \in \mathbb{E}$ such that the closed ball $B_r(x)$ of radius r intersects A. The Hausdorff metric can be also extended to a pseudometric for bounded subsets of $\mathbb{E}$.

The uplifted convex combinations and the corresponding convexification operator K can be defined for various families of sets, e.g. closed sets, compact sets, convex sets, etc. In this situation, it is natural to ask whether these spaces of sets, with the uplifted convex combination operation and the Hausdorff metric, are still convex combination spaces and how the operator K is uplifted. For instance, if $\mathbb{E}$ is a Banach space then it is known that K is the identity mapping but its uplifted version onto the family of compact sets is the convex hull operator. In this section, we address this question for the space $\mathcal{K}$ of non-empty compact subsets of a general convex combination space $\mathbb{E}$. The following inequality is well-known in the Banach space setting.

Lemma 6.1. For any two bounded sets $A, C \subset \mathbb{E}$, $d_{\mathrm{H}}(\overline{\mathrm{co}}\, A, \overline{\mathrm{co}}\, C) \leqslant d_{\mathrm{H}}(A, C)$.

Proof. Let $d_{\mathrm{H}}(A, C) = r$. By symmetry and the fact that closures do not affect the value of the Hausdorff pseudometric, it suffices to check that for every $x = [\lambda_j, u_j]_{j=1}^m \in \mathrm{co} A$, there exists $y \in \mathrm{co} C$ such that $d(x, y) \leqslant r$. For each u_j there exists $v_j \in C$ with $d(u_j, v_j) \leqslant r$. Set $y = [\lambda_j, v_j]_{j=1}^m \in \mathrm{co} C$. By property (iv), $d(x, y) \leqslant r$. $\qquad \square$

The following theorem shows that the uplifted $K_\mathcal{K}$ is the convex hull operator if $\mathcal{K}$ is a family of compact sets in a convex combination space $\mathbb{E}$. This provides an extra reason for calling K a convexification operator.

Theorem 6.2. Let $\mathcal{K}$ be the family of compact sets in $\mathbb{E}$ with the uplifted convex combination operation and the Hausdorff metric. Then, $\mathcal{K}$ is a convex combination space and its convexification operator $K_\mathcal{K}$ is $\mathrm{co}K_\mathbb{E}$.

Proof. Properties (i), (ii) and (iv) follow from their analogues in $\mathbb{E}$ and the definitions of the uplifted convex combination (6.1) and the Hausdorff metric.

Let us prove (iii) for finite sets A_1, A_2. Assume that

$$d_\mathrm{H}([\lambda_i^{(k)}, A_i]_{i=1}^2, [\lambda_i, A_i]_{i=1}^2) > \varepsilon$$

for all k, whence $d(v, [\lambda_i^{(k)}, A_i]_{i=1}^2) > \varepsilon$ for some $v = [\lambda_i, u_i]_{i=1}^n \in [\lambda_i, A_i]_{i=1}^2$ or $d(v_k, [\lambda_i, A_i]_{i=1}^2) > \varepsilon$ for $v_k = [\lambda_i^{(k)}, u_i^{(k)}]_{i=1}^2 \in [\lambda_i^{(k)}, A_i]_{i=1}^2$. In the first case,

$$d(v, [\lambda_i^{(k)}, A_i]_{i=1}^2) \leqslant d([\lambda_i, u_i]_{i=1}^2, [\lambda_i^{(k)}, u_i]_{i=1}^2) \to 0$$

by property (iii) of $\mathbb{E}$ leading to a contradiction. In the second case, using the finiteness of $A_1 \times A_2$, and by taking a further subsequence, it is possible to assume that $u_i^{(k)} = u_i$ for some u_i and all i, k. Then the above arguments using (iii) yield a contradiction.

For the proof of the general case, fix $\varepsilon > 0$. By the compactness of A_i, we can take finite sets $A_i^{(\varepsilon)} \subset A_i$ such that $d_\mathrm{H}(A_i, A_i^{(\varepsilon)}) \leqslant \varepsilon$. Taking into account property (iv) of $\mathcal{K}$,

$$\begin{aligned}
d_\mathrm{H}([\lambda_i^{(k)}, A_i]_{i=1}^2, [\lambda_i, A_i]_{i=1}^2) &\leqslant d_\mathrm{H}([\lambda_i^{(k)}, A_i]_{i=1}^2, [\lambda_i^{(k)}, A_i^{(\varepsilon)}]_{i=1}^2 \\
&\quad + d_\mathrm{H}([\lambda_i^{(k)}, A_i^{(\varepsilon)}]_{i=1}^2, [\lambda_i, A_i^{(\varepsilon)}]_{i=1}^2) \\
&\quad + d_\mathrm{H}([\lambda_i, A_i^{(\varepsilon)}]_{i=1}^2, [\lambda_i, A_i]_{i=1}^2)) \\
&\leqslant \varepsilon + \varepsilon_k + \varepsilon,
\end{aligned}$$

where $\varepsilon_k \to 0$. We conclude routinely that (iii) holds for $\mathcal{K}$.

Let us finally prove (v) with $K_\mathcal{K} = \overline{\mathrm{co}}\, K_\mathbb{E}$. Assume first that A is finite. If $A_{(n)} = [n^{-1}, A]_{i=1}^n$ does not d_H-converge to $\overline{\mathrm{co}}\, K_\mathbb{E}(A)$ as $n \to \infty$, by taking a subsequence we can assume that $d(v, A_{(n)}) > \varepsilon$ for some $v = [\lambda_j, K_\mathbb{E} u_j]_{j=1}^l \in \overline{\mathrm{co}}\, K_\mathbb{E}(A)$ or $d(v_n, \overline{\mathrm{co}}\, K_\mathbb{E}(A)) > \varepsilon$ for some $v_n = [n^{-1}, u_i^{(n)}]_{i=1}^n \in A_{(n)}$.

In the first case, a contradiction will be reached if we construct a sequence $v_n \in A_{(n)}$ such that $v_n \to v$. For each $n \geqslant 1$ define natural numbers $m_j(n)$ such that $\frac{m_j(n)-1}{n} < \lambda_i \leqslant \frac{m_j(n)}{n}$ for all $j = 1, \dots, l-1$ and $m_l(n) = n - \sum_{j=1}^{l-1} m_j(n)$. Then $m_j(n) \to \infty$ and $\frac{m_j(n)}{n} \to \lambda_j$ as $n \to \infty$ and $m_l(n) > 0$ for all sufficiently large n. Let

$$v_n = [\underbrace{n^{-1}, u_1; \dots; n^{-1}u_1}_{m_1(n)}; \underbrace{n^{-1}u_2; \dots; n^{-1}u_2}_{m_2(n)}; \dots; \underbrace{n^{-1}u_l; \dots; n^{-1}u_l}_{m_l(n)}],$$

where each u_j appears exactly $m_j(n)$ times. By Lemma 2.1 and joint continuity,

$$v_n = [m_j(n)n^{-1}, [m_j(n)^{-1}, u_j]_{i=1}^{m_j(n)}]_{j=1}^l \to [\lambda_j, K_{\mathbb{E}} u_j]_{j=1}^l = v.$$

In the second case, define $m_j(n) = \mathrm{card}\{1 \leqslant i \leqslant n : u_i^{(n)} = u_j\}$ for $j = 1, \dots, l$. Since each sequence $\{m_j(n)n^{-1}, n \geqslant 1\}$ is contained in $[0, 1]$, we can assume by taking subsequences that it converges to $\lambda_j \in [0, 1]$.

Observe that at least one λ_{j_0} is strictly positive. Define the triangular array $\{w_i^{(n)}\}_{i,n} \subset A$ to be $w_i^{(n)} = u_j$ if $u_i^{(n)} = u_j$ and $m_j(n)n^{-1} \not\to 0$, and $w_i^{(n)} = u_{j_0}$ otherwise. Define also $y_n = [n^{-1}, w_i^{(n)}]_{i=1}^n$. By (iv),

$$d(v_n, y_n) \leqslant n^{-1} \mathrm{card}\{1 \leqslant i \leqslant n : u_i^{(n)} \neq w_i^{(n)}\} \cdot \mathrm{diam}(A) \to 0.$$

Since

$$d(v_n, \overline{\mathrm{co}}\, K_{\mathbb{E}}(A)) \leqslant d(v_n, y_n) + d(y_n, \overline{\mathrm{co}}\, K_{\mathbb{E}}(A))$$

eventually also $d(y_n, \overline{\mathrm{co}}\, K_{\mathbb{E}}(A)) > \varepsilon$. Then we can write, relabelling the points u_j if necessary,

$$y_n = [m_j(n(k))n(k)^{-1}, [m_j(n(k))^{-1}, u_j]_{i=1}^{n(k)}]_{j=1}^s$$

for some $1 \leqslant s \leqslant l$ and a sequence $n(k) \to \infty$ as $k \to \infty$. By joint continuity, $y_n \to [\lambda_j, K_{\mathbb{E}} u_j]_{j=1}^s \in \overline{\mathrm{co}}\, K_{\mathbb{E}}(A)$. That is a contradiction.

For the general case of a compact A one proceeds like for (iii) using approximation of A by finite sets $A^{(\varepsilon)}$. One just has to take into account that, by Lemma 6.1 and Proposition 3.6,

$$d_{\mathrm{H}}(\overline{\mathrm{co}}\, K_{\mathbb{E}}(A^{(\varepsilon)}), \overline{\mathrm{co}}\, K_{\mathbb{E}}(A)) \leqslant d_{\mathrm{H}}(K_{\mathbb{E}}(A^{(\varepsilon)}), K_{\mathbb{E}}(A)) \leqslant \varepsilon.$$

Finally, it is possible to get rid of the closure in the right-hand side of $K_{\mathcal{K}} = \overline{\mathrm{co}}\, K_{\mathbb{E}}$. Indeed, the continuity of K implies that $K_{\mathbb{E}} A = \{K_{\mathbb{E}} u : u \in A\}$ is a compact set for every compact set $A \in \mathcal{K}$. Then it suffices to note that the convex hull of a compact set is closed. $\square$

Let X be an integrably bounded random compact set in a Banach space $\mathbb{E}$, i.e. $d_H(\{0\}, X)$ is integrable. The uplifted K is the convex hull operator, whence the expectation of the simple random compact set is the weighted Minkowski sums of the convex hulls of its possible values. For a general integrably bounded random compact set X, its expectation $\mathbf{E}X$ in the space $\mathcal{K}$ with the uplifted convex combination operation is a compact set whose support function equals the expected support function of X. It is known (see, e.g. Ref. 14, Th. II.1.21) that if the probability space is non-atomic, then $\mathbf{E}X$ coincides with the *selection expectation* (also called the Aumann expectation) of X. The selection expectation is defined as the closure of the set of expectations $\mathbf{E}\xi$ for all integrable $\mathbb{E}$-valued random elements ξ such that $\xi \in X$ a.s. Note that such ξ is called a measurable selection of X. This original definition goes back to Ref. 3, while its relationship to the law of large numbers for random compact sets was discovered in Ref. 2. By Theorem 5.1, the normalised Minkowski sum $[n^{-1}, X_i]_{i=1}^n = n^{-1}(X_1 + \cdots + X_n)$ almost surely converges to $\mathbf{E}X$, i.e. the uplift construction immediately yields the strong law of large numbers for Minkowski sums of identically distributed random compact sets in a Banach space $\mathbb{E}$ (cf. Ref. 1). It should be noted that our setup does not cover the case of non-compact summands (cf. Refs. 11, 12 and 28).

It is possible to uplift the convex combination operation even further by applying it to $\mathcal{K}$-valued functions. Let $\mathcal{U}$ be the space of $\mathcal{K}$-valued non-decreasing left-continuous functions $F(t)$, $t \in (0, 1]$. The convex combination of $F_1, \ldots, F_n \in \mathcal{U}$ is defined as the function $F(t)$ whose value at t is the uplifted convex combination (6.1) of the sets $F_1(t), \ldots, F_n(t)$. Endow $\mathcal{U}$ with the uniform metric

$$d_H^\infty(F_1, F_2) = \sup_{0 < t \leqslant 1} d_H(F_1(t), F_2(t)). \tag{6.3}$$

Note that the space $\mathcal{U}$ with the uniform metric is not separable.

Theorem 6.3.   The convexification operator $K_\mathcal{U}$ uplifted to $\mathcal{U}$ from $\mathbb{E}$ is given by the convex hull in $\mathbb{E}$, i.e. $(K_\mathcal{U}F)(t) = \mathrm{co}K_\mathbb{E}(F(t))$, $t \geqslant 0$.

Proof.   By Molchanov's[15] Theorem 2, it suffices to show that

$$d_H\left([n^{-1}, F(t)]_{i=1}^n, \mathrm{co}K_\mathbb{E}(F(t))\right) \to 0$$

for all $t \in (0, 1]$ and

$$d_H\left([n^{-1}, F(t+)]_{i=1}^n, \mathrm{co}K_\mathbb{E}(F(t+))\right) \to 0$$

for all $t \in [0, 1)$, where $F(t+)$ is the limit of $F(s)$ as $s \downarrow t$. Both these statements follow from Theorem 6.2 applied to $F(t)$ and to $F(t+)$. $\square$

7. CONVEX COMBINATIONS AND THE DOSS EXPECTATION

It was shown in Section 4 that the convex combination concept naturally leads to a definition of the expectation. In the other direction, a concept of the expectation leads to the definition of the convex combination by defining $[\lambda_i, u_i]_{i=1}^n$ to be the expectation of the simple random element that takes values $u_1, \ldots, u_n$ with probabilities $\lambda_1, \ldots, \lambda_n$. Note that (ii) corresponds to the total probability formula for the expectations of discrete random elements, (iii) is the continuity property and (iv) means that the expectation is non-expansive, i.e. $d(\mathbf{E}\xi, \mathbf{E}\eta) \leqslant \mathbf{E}d(\xi, \eta)$.

The operator K from (v) is the identity operator if and only if the expectation of each deterministic singleton $u \in \mathbb{E}$ coincides with u. In this case, we call both the expectation operator and the convex combination operation *unbiased*. In its turn, a metric space $\mathbb{E}$ will be called *convexifiable* if $\mathbb{E}$ admits an unbiased convex combination operation or, equivalently, an unbiased expectation operator.

Let ξ be an integrable random element in $\mathbb{E}$. The *Doss expectation* of ξ is the set $\mathbf{E}_{\text{Doss}}\xi$ of all points $a \in \mathbb{E}$ such that $d(x, a) \leqslant \mathbf{E}d(x, \xi)$ for all $x \in \mathbb{E}$. Note that the Doss expectation may be empty or may contain more than one point in $\mathbb{E}$. The random element ξ is said to be *Doss integrable* if its Doss expectation is non-empty. The following proposition shows that if $\mathbb{E}$ is a Banach space, then the Doss expectation becomes the conventional expectation of ξ.

Theorem 7.1.   (see Ref. 6) If $\mathbb{E}$ is a Banach space with the metric generated by the norm, then each Bochner integrable random element ξ satisfies $\mathbf{E}_{\text{Doss}}\xi = \{\mathbf{E}\xi\}$, where $\mathbf{E}\xi$ is the Bochner expectation of ξ.

A metric space $(\mathbb{E}, d)$ is called *convex in the sense of Doss* if for any two elements $x_1, x_2 \in \mathbb{E}$ there exists an element $a \in \mathbb{E}$ such that

$$d(x, a) \leqslant \frac{1}{2}(d(x, x_1) + d(x, x_2)) \tag{7.1}$$

for all $x \in \mathbb{E}$. Clearly, each Banach space is convex in the sense of Doss. It is easy to see that (7.1) implies the Doss integrability of all random elements which take two values with equal probabilities.

Theorem 7.2.   If $\mathbb{E}$ is a convexifiable metric space, then every integrable random element $\xi \in L_{\mathbb{E}}^1$ is Doss integrable and $\mathbf{E}\xi \in \mathbf{E}_{\text{Doss}}\xi$.

Proof. Assume that $\mathbb{E}$ is equipped with the unbiased convex combination operation. By Proposition 3.2, every $y \in \mathbb{E}$ is convexely decomposable. Then $\mathbf{E}y = Ky = y$ and

$$d(\mathbf{E}\xi, y) = d(\mathbf{E}\xi, \mathbf{E}y) \leqslant \mathbf{E}d(\xi, y).$$

Hence $\mathbf{E}\xi$ belongs to the Doss expectation of ξ, i.e. $\mathbf{E}_{\mathrm{Doss}}\xi$ is non-empty. □

Corollary 7.3. If $\mathbb{E}$ is any convex combination space, then $K(\mathbb{E})$ is Doss convex. Furthermore, every convexifiable metric space is Doss convex.

Proof. Since K is idempotent by Proposition 3.4, the convex combination operation restricted to $K(\mathbb{E})$ is unbiased. Then Doss integrability of two-valued functions yields the Doss convexity. The second statement follows from the fact that $K(\mathbb{E}) = \mathbb{E}$ in a convexifiable space. □

Proposition 7.4. Let ξ be a random element in a convexifiable metric space $\mathbb{E}$. Then $\xi \in L^1_{\mathbb{E}}$ if and only if $\mathbf{E}_{\mathrm{Doss}}\xi$ is non-empty and bounded.

Proof. If $\xi \in L^1_{\mathbb{E}}$, then $\mathbf{E}d(u_0, \xi) < \infty$, whence by its definition $\mathbf{E}_{\mathrm{Doss}}\xi \subset B(u_0, \mathbf{E}d(u_0, \xi))$.

Conversely, if $\xi \notin L^1_{\mathbb{E}}$, then $\mathbf{E}d(u_0, \xi) = \infty$, whence $\mathbf{E}d(u, \xi) = \infty$ for every $u \in \mathbb{E}$. Hence $\mathbf{E}_{\mathrm{Doss}}\xi = \mathbb{E}$, but $\mathbb{E}$ cannot be bounded since $\mathbf{E}d(u_0, \xi) = \infty$. □

In principle, a metric space may admit several unbiased convex combination operations. However, this is not possible if the Doss expectation is single-valued, e.g. if $\mathbb{E}$ is a Banach space.

Theorem 7.5. A Banach space admits the unique unbiased convex combination operation, which is given by the usual convex combination $[\lambda_i, u_i]_{i=1}^n = \sum \lambda_i u_i$. The same holds in any Polish space $\mathbb{E}$ if the corresponding Doss expectation in $\mathbb{E}$ is always single-valued.

Proof. Let ξ be a Bochner integrable random element of that Banach space. By Theorem 7.1, $\mathbf{E}_{\mathrm{Doss}}\xi = \{\mathbf{E}\xi\}$ where $\mathbf{E}\xi$ is the Bochner integral of ξ. Since the single-valued Doss expectation yields an unbiased convex combination operation,

$$\mathbf{E}\xi \in \mathbf{E}_{\mathrm{Doss}}\xi = \{\mathbf{E}\xi\}.$$

The result follows by choosing ξ to be a simple function. □

A subset of a metric space $\mathbb{E}$ is called *universally convex* if it is convex with respect to every unbiased convex combination operation on $\mathbb{E}$. By Theorem 7.5, convex sets and universally convex sets do not differ in a Banach space. Notice the following consequence of Proposition 3.8.

Proposition 7.6. Let $\mathbb{E}$ be a metric space. Then, every singleton and every ball are universally convex.

8. UNIFORM STRONG LAW OF LARGE NUMBERS

Let $\mathcal{O}$ be the class of all unbiased convex combination operations that may be defined on $\mathbb{E}$. A convex combination operation will be denoted by $_o[\cdot, \cdot]$ for $o \in \mathcal{O}$. Its associated expectation will be denoted by $_o\mathbf{E}$.

Note that the space $L_{\mathbb{E}}^1$ only depends on the metric of $\mathbb{E}$ and as such it stays the same for every convex combination operation o that may be defined in $\mathbb{E}$. Hence, in every convexifiable metric space $\mathbb{E}$ a set-valued expectation

$$\mathbf{E}^*\xi = \mathrm{cl}\{_o\mathbf{E}\xi \mid o \in \mathcal{O}\} \tag{8.1}$$

and the corresponding operation

$$[[\lambda_i, u_i]]_{i=1}^n = \mathrm{cl}\{_o[\lambda_i, u_i]_{i=1}^n \mid o \in \mathcal{O}\}$$

make sense because $\mathcal{O}$ is non-empty. In general, the operation $[[\cdot, \cdot]]$ does not satisfy property (ii). Then we have the following corollary to Proposition 7.2.

Corollary 8.1. If $\xi \in L_{\mathbb{E}}^1$, then $\mathbf{E}_{\mathrm{Doss}}\xi \supset \mathbf{E}^*\xi$.

Note that (8.1) is a definition of expectation in a metric space, which is not tied to any particular algebraic structure on $\mathbf{E}$. Corollary 8.1 signals that $\mathbf{E}^*$ is interesting, since it is in general set-valued but always smaller than the Doss expectation. Since the reason of the appearance of the set-valued expectations is that we are not able to distinguish a single central point of a distribution, $\mathbf{E}^*$ is *more discriminating* than $\mathbf{E}_{\mathrm{Doss}}$.

We are going to obtain the strong law of large numbers for this set-valued expectation as a consequence of a stronger and maybe surprising result, namely that the law of large numbers proven in Section 5 is in fact uniform over the family of all unbiased convex combination operations.

To begin with, we note that $\mathbf{E}^*\xi$ is always bounded, hence it makes sense to use convergence in the Hausdorff metric in the study of this expectation.

Corollary 8.2. Let $\mathbb{E}$ be a convexifiable metric space. Then $\mathbf{E}^*\xi$ is bounded and $d(u, \mathbf{E}^*\xi) \leqslant \mathbf{E}d(u, \xi)$ for every $u \in \mathbb{E}$.

Proof. Use Corollary 8.1 and Proposition 7.4. $\qquad\square$

Proposition 8.3. If $\mathbb{E}$ is a convexifiable space, then property (iii) holds uniformly over all unbiased convex combination operations $o \in \mathcal{O}$.

Proof. Let $\lambda^{(k)} \to \lambda \in (0, 1)$. We assume that $\lambda^{(k)} \leqslant \lambda$ for all k, but the reasoning is the same in all cases.

For all $o \in \mathcal{O}$ every element of $\mathbb{E}$ is convexely decomposable (see Proposition 3.2). Applying Proposition 3.5 and then property (iv), we obtain

$$\sup_{o\in\mathcal{O}} d({}_o[\lambda^{(k)}, u; 1 - \lambda^{(k)}, v], {}_o[\lambda, u; 1 - \lambda, v])$$

$$= \sup_{o\in\mathcal{O}} d\big({}_o[\lambda^{(k)}, u; |\lambda^{(k)} - \lambda|, v; 1 - \lambda, v], {}_o[\lambda^{(k)}, u; |\lambda^{(k)} - \lambda|, u; 1 - \lambda, v]\big)$$

$$\leqslant \sup_{o\in\mathcal{O}} |\lambda^{(k)} - \lambda| d(u, v) = |\lambda^{(k)} - \lambda| d(u, v),$$

where the latter converges to 0. $\qquad\square$

We can now derive the uniform strong law of large numbers. Of course, it is a vacuous statement if $\mathbb{E}$ is not convexifiable, i.e. does not admit any convex combination operation for which K is the identity mapping.

Theorem 8.4. Let $\xi \in L_{\mathbb{E}}^1$, and let $\xi_1, \xi_2, \ldots$ be a sequence of pairwise-independent random elements distributed as ξ. Then

$$\sup_{o\in\mathcal{O}} d({}_o[n^{-1}, \xi_i]_{i=1}^n, {}_o\mathbf{E}\xi) \to 0$$

almost surely as $n \to \infty$.

Proof. If we examine the proof of the strong law of large numbers, we see that among the three quantities (**I**)–(**III**), only (**II**) depends on the choice of $o \in \mathcal{O}$. Hence it suffices to prove the result for simple functions. This amounts to showing, with the notation used in that proof, that

$$\sup_{o\in\mathcal{O}} d({}_o[B_{n,j}n^{-1}, [B_{n,j}^{-1}, u_j]_{i=1}^{B_{n,j}}]_{j=1}^m, {}_o[\mathbf{P}(\Omega_j), Ku_j]_{j=1}^m) \to 0.$$

But $o \in \mathcal{O}$ means that $K u_j = u_j$ and u_j is convexely decomposable for all j, hence $[B_{n,j}^{-1}, u_j]_{i=1}^{B_{n,j}} = u_j$ and there remains to prove that

$$\sup_{o \in \mathcal{O}} d({}_o[B_{n,j} n^{-1}, u_j]_{j=1}^m, {}_o[\mathbf{P}(\Omega_j), u_j]_{j=1}^m) \to 0.$$

This follows from an application of Proposition 8.3 together with Etemadi's law of large numbers. Notice that the latter is only required in order to obtain $B_{n,j} n^{-1} \to \mathbf{P}(\Omega_j)$ and hence is valid outside of a null set, whatever the cardinality of $\mathcal{O}$ is.   □

Corollary 8.5. Let **E** be a convexifiable metric space. If $\xi_1, \xi_2, \dots,$ is a sequence of pairwise-independent random elements distributed as $\xi \in L_{\mathbb{E}}^1$, then

$$[[n^{-1}, \xi_i]]_{i=1}^n \to \mathbf{E}^* \xi \qquad \text{a.s. as } n \to \infty,$$

where the convergence is understood in the Hausdorff metric.

Proof. It follows from the definition (6.2) of the Hausdorff metric that

$$d_{\mathrm{H}}(\mathrm{cl}(\cup_j A_j), \mathrm{cl}(\cup_j C_j)) \leqslant \sup_j d_{\mathrm{H}}(A_j, C_j).$$

Hence,

$$d_{\mathrm{H}}([[n^{-1}, \xi_i]]_{i=1}^n, \mathbf{E}^* \xi) \leqslant \sup_{o \in \mathcal{O}} d({}_o[n^{-1}, \xi_i]_{i=1}^n, {}_o\mathbf{E}\xi).$$

Then apply Theorem 8.4.   □

9. EXAMPLES

Example 1. (Banach spaces). If $\mathbb{E}$ is a Banach space with the conventional definition of the convex combination operation, then our concepts turn into the conventional Bochner expectation and the strong law of large numbers in a Banach space.

It is interesting to consider an uplift of this operation to the family of all compact subsets of $\mathbb{E}$, i.e. Minkowski sums of compact sets. By Theorem 6.2, the uplifted operator $K_{\mathcal{K}}$ is the convex hull operator. The convexification property (v) in this setting has been considered in Ref. 1, while the finite-dimensional case easily follows from the Shapley–Folkman-Starr theorem (see Ref.2 and 14 Section 3.1). The corresponding uplifted strong law of large numbers then becomes the strong law of large numbers for random compact sets in a Banach space (see Ref. 1).

Example 2. (Herer expectation).   The *Herer expectation* of a random compact set X in a metric space[9] is related to the Doss expectation and is given by

$$\mathbf{E}_{\mathrm{H}}X = \{x : d(x, a) \leqslant d_{\mathrm{H}}(X, \{a\}) \text{ for all } a \in \mathbb{E}\}.$$

This expectation induces the convex combination operation on the family $\mathcal{K}$ of compact sets in $\mathbb{E}$, namely $[\lambda_i, A_i]_{i=1}^n$ is the Herer expectation of a random compact set taking values A_i with probabilities λ_i, $i = 1, \ldots, n$.

Let $\mathbb{E}$ be a separable Banach space. For any subset $A \subset \mathbb{E}$, the intersection of all closed balls that cover A is denoted by $\mathrm{blh}\, A$ and called the *ball hull* of A. By Theorem 41 in Ref.[26], the Herer expectation of a random compact set X equals the ball hull of its selection expectation. Therefore, the Herer convex combination $[\lambda_i, A_i]_{i=1}^n$ is given by the ball hull of the Minkowski sum $\sum_{i=1}^n \lambda_i A_i$.

Assume that $\mathrm{blh}(\mathcal{K})$ is stable under Minkowski addition of balls, in the sense that $K + C$ is an intersection of balls whenever $K \in \mathcal{K}$ is so and C is a ball. It follows then from the definition of the Hausdorff metric that blh is non-expansive (see Ref. 26 Prop. 21), whence properties (i)–(iv) follow from their analogous for the uplifted convex combination operation in Section 6. Property (v) holds with $K = \mathrm{blh}$. The ball hull equals the convex hull if the space $\mathbb{E}$ has the *Mazur Intersection Property* (see Ref. 8), but in general it is larger.

Example 3. (Upper semicontinuous functions).   Let $\mathbb{E}$ be a convex combination space. Consider upper semicontinuous functions $u \colon \mathbb{E} \mapsto [0, 1]$ such that $\inf u = 0$, $\sup u = 1$ and $\mathrm{supp}\, u = \{x : u(x) > 0\}$ is relatively compact in $\mathbb{E}$. The convex combination of $u_1, \ldots, u_n$ is a function v such that, for each $t \in (0, 1]$, its *level set* $\{x : v \geqslant t\}$ is the uplifted convex combination of $\{x : u_i(x) \geqslant t\}$, $i = 1, \ldots, n$ (cf. Theorem 6.3). In other words,

$$v(y) = \sup_{[\lambda_i, x_i]_{i=1}^n = y} \min(u_1(x_1), \ldots, u_n(x_n)). \tag{9.1}$$

By Theorem 6.3, the corresponding convexification operator K transforms an upper semicontinuous function u into the function Ku whose level sets are convex hulls of the level sets of u, i.e. $\{x : (Ku)(x) \geqslant t\} = \mathrm{co}\{x : u(x) \geqslant t\}$ for all $t \in (0, 1]$. Theorem 5.1 yields the strong law of large numbers for convex combinations of random upper semicontinuous functions.

If $\mathbb{E}$ is a linear space with the conventional convex combination operation, then this definition turns into the definition of level sums of upper semicontinuous functions (see Ref. 20). The resulting function v has the level sets being the weighted Minkowski averages of the level sets of

$u_1, \ldots, u_n$. Theorem 5.1 then implies the strong law of large numbers for level sums of random upper semicontinuous functions (see Refs. 5,13,15 and 19) for the corresponding results for $\mathbb{E}$ being a Banach space.

Example 4. (T-normed sums of upper semicontinuous functions). It is possible to replace the minimum in the right-hand side of (9.1) with another operation given by a triangular norm (usually called *t-norm*). A *triangular norm* $\top$ is an associative and commutative operation in $[0, 1]$ which is non-decreasing as a bivariate function and has 1 as its neutral element.

Define the convex combination operation of upper semicontinuous functions $u_1, \ldots, u_n$ as

$$[\lambda_i, u_i]_{i=1}^n(x) = \sup_{[\lambda_i, x_i]_{i=1}^n = x} u_1(x_1) \top \cdots \top u_n(x_n)$$

(see also Ref. 27). If we restrict ourselves to the space $\mathcal{U}$ of upper semi-continuous functions such that the level set mapping $u \mapsto \{x : u(x) \geqslant t\}$ is continuous in the Hausdorff metric, then $\mathcal{U}$ is separable with the uniform metric given by (6.3) and properties (i)–(iv) hold.

Property (v) (see Ref. 27, Prop. 4) also holds if the limit is defined in the weak topology generated by all rational level set mappings, i.e. by the pseudometrics $\rho_t(u, v) = d_{\mathrm{H}}(\{x : u(x) \geqslant t\}, \{x : v(x) \geqslant t\})$ where t is any rational number from $(0, 1]$. The convexification operator K is given by

$$(Ku)(x) = \sup \left\{ a \in \left[0, \sup_{x \in \mathrm{co}\{x', x''\}} \min\{u(x'), u(x'')\} \right] : a \top a = a \right\}.$$

The method of proof of Theorem 5.1 works for each pseudometric if one bounds ρ_t from above by the uniform metric (6.3). The strong law of large numbers obtained this way is stronger than Ref. 27, Proposition 9. Theorem 13 in the same paper drops the requirement of continuity of level mappings.

Example 5. (Power combinations). Define the convex combination operation on a linear normed space $\mathbb{E}$ as

$$[\lambda_i, u_i]_{i=1}^n = \sum \lambda_i^p u_i \tag{9.2}$$

for some $p > 1$. The corresponding convexification operator satisfies $Ku = 0$ for every u, and the convex hull of a singleton $\{u\}$ is the segment joining u and the origin.

Example 6. (Extremes). The space $\mathbb{E}=[0,\infty)$ with the Euclidean distance and the maximum operation

$$[\lambda_i, u_i]_{i=1}^n = \max(\lambda_1 u_1, \ldots, \lambda_n u_n)$$

is a convex combination space. The convexification operator K satisfies $Ku = 0$ for every u. Thus, the only convex singleton is $\{0\}$. The corresponding expectation is equal to zero for all simple random variables and all integrable (in the usual sense) random variables. The strong law of large numbers means that $n^{-1}\max(X_1, \ldots, X_n)$ converges to zero almost surely for i.i.d. copies of any integrable random variable X.

Example 7. (Cumulative distribution functions). Let $\mathbb{E}$ be the space of all distributions of integrable real-valued random variables (alternatively, the space of all cumulative distribution functions) metrised by any *ideal probability metric* $\mathfrak{m}$ of order $r = 1$ that is continuous with respect to the weak convergence (see Ref. 21 and Section 1.4 of Ref. 32). This means that $\mathfrak{m}(\xi + \zeta, \eta + \zeta) \leqslant \mathfrak{m}(\xi, \eta)$ for any two random variables ξ and η and another independent random variable ζ, and $\mathfrak{m}(c\xi, c\eta) = |c|^r \mathfrak{m}(\xi, \eta)$ for all $c \neq 0$. The convex combination of the random variables $u_1, \ldots, u_n$ is defined as the random variable distributed as $\sum \lambda_i u_i$, where $u_1, \ldots, u_n$ are chosen to be independent. Then (i)–(iv) hold, particularly the latter follows from the ideal property of the metric (see Ref. 32, 1.4.4). The strong law of large numbers for random variables implies that Ku is the expectation of u. Therefore, convex singletons in $\mathbb{E}$ correspond to deterministic random variables. A set $A \subset \mathbb{E}$ is convex if A is closed under taking finite convolutions of the rescaled cumulative distribution functions corresponding to the elements of A. For instance, all normal distributions form a convex set.

Since a random element in $\mathbb{E}$ corresponds to a random probability measure,[4,16] Theorem 5.1 yields a strong law of large numbers for convolutions of random probability measures.

Example 8. (Stable probability distributions). Let $\mathbb{E}$ be the space of random variables with an ideal metric $\mathfrak{m}$ of order $\alpha \in (0, 2]$. The convex combination of independent random variables $u_1, \ldots, u_n$ is defined as the random variable distributed as $\sum_{i=1}^n \lambda_i^{1/\alpha} u_i$. Then $Ku = u$ if and only if u has the α-stable distribution (see Ref. 31).

Max-stable random variables appear as convex singletons if $\mathfrak{m}$ is the ideal metric in the max-scheme (see Chapter 18 in Ref. 21) and the convex combination of $u_1, \ldots, u_n$ is the random variable distributed as $\max(\lambda_1 u_1, \ldots, \lambda_n u_n)$.

Example 9. (Curved spaces). Let $\mathbb{E}$ be a global NPC space, i.e.

$$\inf_{u\in\mathbb{E}} \mathbf{E}d^2(\xi,u) \leqslant \frac{1}{2}\mathbf{E}d^2(\xi,\xi') \tag{9.3}$$

for any two i.i.d. discrete $\mathbb{E}$-valued random elements ξ and ξ' (see Ref. 24). This property implies that the expectation of a square integrable random element ξ is uniquely defined as the minimiser of $\mathbf{E}d^2(\xi,u)$ over $u\in\mathbb{E}$ (see Theorem 2.1 in Ref. 24. It is also possible to define the conditional expectations and martingales. If $\mathbb{E}$ is a linear space, then the right-hand side of (9.3) is the variance of ξ, and (9.3) turns into an equality.

This definition of the expectation gives rise to the unbiased convex combination operation. Theorem 5.1 implies the strong law of large numbers for square-integrable random elements (cf Theorem 2.6 in Ref. 24. Note that the square integrability is essential, since otherwise the expectation is not well defined. The uniform strong law of large numbers (Theorem 8.4) yields the uniform convergence result over all possible metrics on $\mathbb{E}$ that turn $\mathbb{E}$ into a global NPC space. Theorem 6.2 describes the uplifted convex combination operation, and thereupon provides the strong law of large numbers for elementwise sums of random sets in $\mathbb{E}$.

Single-valued expectations appear also in Ref. 17 as barycentres of probability measures on manifolds. They may provide example of several unbiased convex combination operations.

When averaging elements of general curved spaces it is often natural to assume that the result of the convex combination operation is not a single element, but a set in $\mathbb{E}$ (see Refs. 10 and 22). The convex combination $[\lambda_1,u_1;\lambda_2,u_2]$ of two points in $\mathbb{E}$ is then defined as the set of all points z such that $d(z,u_1)=\lambda_2 d(u_1,u_2)$ and $d(z,u_2)=\lambda_1 d(u_1,u_2)$ (see Refs. 10 and 22). General convex combinations are then defined by induction. If $[\lambda_i,u_i]_{i=1}^k$ is defined for $k<n$, then $[\lambda_i,u_i]_{i=1}^n$ is defined as the set of all $[p_1,v_1;p_2,v_2]$, where $p_j=\sum_{i\in I_j}\lambda_i$, $v_j=[\lambda_i/p_j,u_i]_{i\in I_j}$ for $j=1,2$ and any disjoint sets I_1 and I_2 such that $I_1\cup I_2=\{1,\ldots,n\}$ (see Definition 1.1 in Ref. 10. The space $\mathbb{E}$ is called *convex* (resp. *strictly convex*) if such defined convex combination operation $[\lambda_1,u_1;\lambda_2,u_2]$ is always nonempty (resp. has exactly one element). If $\mathbb{E}$ is strictly convex, then condition (iv) coincides with the condition imposed in Definition 1.2 in Ref. 10. This setting corresponds to an extension of the uniquely defined convex combinations by assuming that both the result of $[\lambda_i,u_i]_{i=1}^n$ and its arguments $u_1,\ldots,u_n$ may be set-valued. Note that this setting is different from uplifting a single-valued convex combination operation in Section 6.

A generic multi-valued combination of $u_1, \ldots, u_n$ may also be defined as the set of all z such that

$$\sum_{i=1}^{n} \lambda_i d(z, u_i)^p = \inf_{x \in \mathbb{E}} \sum_{i=1}^{n} \lambda_i d(x, u_i)^p ,$$

where $p > 0$ is a fixed number. If $p = 2$, then one obtains the Fréchet mean of a random element that takes values $u_1, \ldots, u_n$ with the probabilities $\lambda_1, \ldots, \lambda_n$ (see Ref. 18). The corresponding strong law of large numbers has been obtained in Ref. 30, while the case of a general $p > 0$ is considered in Ref. 25. If $p = 1$, one obtains a variant of the spatial median (see Ref. 23). These results however cannot be obtained as corollaries of our framework, since property (ii) does not necessarily hold.

REFERENCES

1. Artstein, Z., and Hansen, J. C. (1985). Convexification in limit laws of random sets in Banach spaces. *Ann. Probab.* **13**, 307–309.
2. Artstein, Z., and Vitale, R. A. (1975). A strong law of large numbers for random compact sets. *Ann. Probab.* **3**, 879–882.
3. Aumann, R. J. (1965). Integrals of set-valued functions. *J. Math. Anal. Appl.* **12**, 1–12.
4. Bloomer, L., and Hill, T. P. (2002). Random probability measures with given mean and variance. *J. Theoret. Probab.* **15**, 919–937.
5. Colubi, A., López-Díaz, M., Domínguez-Menchero, J. S., and Gil, M. A. (1999). A generalized strong law of large numbers. *Probab. Theor. Relat. Fields* **114**, 401–417.
6. Doss, S. (1949). Sur la moyenne d'un élément aléatoire dans un espace distancié. *Bull. Sci. Math.* **73**, 48–72.
7. Etemadi, N. (1981). An elementary proof of the strong law of large numbers. *Z. Wahrsch. verw. Gebiete* **55**, 119–122.
8. Granero, A. S., Moreno, J. P., and Phelps, R. R. (2004). Convex sets which are intersections of closed balls. *Adv. Math.* **183**, 183–208.
9. Herer, W. (1991). Mathematical expectation and martingales of random subsets of a metric space. *Probab. Math. Statist.* **11**, 291–304.
10. Herer, W. (1992). Mathematical expectation and strong law of large numbers for random variables with values in a metric space of negative curvature. *Probab. Math. Statist.* **13**, 59–70.
11. Hess, Ch. (1999). The distribution of unbounded random sets and the multivalued strong law of large numbers in nonreflexive Banach spaces. *J. Convex Anal.* **6**, 163–182.
12. Hiai, F. (1984). Strong laws of large numbers for multivalued random variables. In Salinetti, G. (ed.), *Multifunctions and Integrands*, Vol. 1091. Lecture Notes in Mathamatical, Springer, Berlin, pp.160–172.
13. Klement, E. P., Puri, M. L., and Ralescu, D. A. (1986). Limit theorems for fuzzy random variables. *Proc. R. Soc. Lon. A* **407**, 171–182.
14. Molchanov, I. (2005). *Theory of Random Sets*, Springer, London.
15. Molchanov, I. S. (1999). On strong laws of large numbers for random upper semicontinuous functions. *J. Math. Anal. Appl.* **235**, 349–355.
16. Monticino, M. (2001). How to construct a random probability measure. *Int. Statist. Rev.* **69**, 153–167.

17. Picard, J. (1994). Barycentres et martingales sur une varété. *Ann. Inst. H.Poincaré, Sect. B, Prob. et Stat.* **30**, 647–702.

18. Pick, R. (1987). Expectation in metric spaces. *Stud. Sci. Math. Hungarica* **22**, 347–350.

19. Proske, F. N., and Puri, M. L. (2002). Strong law of large numbers for Banach space valued fuzzy random variables. *J. Theoret. Probab.* **15**, 543–551.

20. Puri, M. L., and Ralescu, D. A. (1986). Fuzzy random variables. *J. Math. Anal. Appl.* **114**, 409–422.

21. Rachev, S. T. (1991). *Probability Metrics and the Stability of Stochastic Models*, Wiley, Chichester.

22. Raynaud de Fitte, P. (1997). Théoreme ergodique ponctuel et lois fortes des grands nombres pour des points aléatoires d'un espace métrique à courbure négative. *Ann. Probab.* **25**, 738–766.

23. Small, C. G. (1990). A survey on multidimensional medians. *Int. Statist. Rev.* **58**, 263–277.

24. Sturm, K.-T. (2002). Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature. *Ann. Probab.* **30**, 1195–1222.

25. Sverdrup-Thygeson, H. (1981). Strong law of large numbers for measures of central tendency and dispersion of random variables in compact metric spaces. *Ann. Statist.* **9**, 141–145.

26. Terán, P. (2005). Intersections of balls and expectations of random sets. Unpublished manuscript.

27. Terán, P. (2005). A strong law of large numbers for t-normed arithmetics. Submitted.

28. Uemura, T. (1993). A law of large numbers for random sets. *Fuzzy Sets Sys.* **59**, 181–188.

29. Vakhaniya, N. N., Tarieladze, V. I., and Chobanyan, S. A. (1987). *Probability Distributions on Banach Spaces*. D. Reidel Publ. Co., Dordrecht.

30. Ziezold, H. (1977). On expected figures and a strong law of large numbers for random elements in quasi-metric spaces. In Trans. 7th Prague Conf. Inf. Th., Statist. Dec. Func., Random Processes (Prague, 1974), Vol. A, Dordrecht, Reidel, pp. 591–602.

31. Zolotarev, V. M. (1986). *One-dimensional Stable Distributions*. American Mathematical Society, Providence, RI.

32. Zolotarev, V. M. (1997). *Modern Theory of Summation of Independent Random Variables*. VSP, Utrecht.