Acta Mathematica Vietnamica (2020) 45: 555–570
https://doi.org/10.1007/s40306-018-0285-9

Baum-Katz's Type Theorems for Pairwise Independent Random Elements in Certain Metric Spaces

Nguyen Tran Thuan[1,2] · **Nguyen Van Quang**[1]

Received: 12 January 2018 / Accepted: 4 June 2018 / Published online: 27 August 2018

Abstract
In this study, some Baum-Katz's type theorems for pairwise independent random elements are extended to a metric space endowed with a convex combination operation. Our results are considered in the cases of identically distributed and non-identically distributed random elements. Some illustrative examples are provided to sharpen the results.

Keywords Baum-Katz theorem · Convex combination · Metric space

Mathematics Subject Classification (2010) 28B99 · 60F15 · 60F25 · 52A99

1 Introduction

The concept of complete convergence for a sequence of random variables (r.v.'s) was introduced by Hsu and Robbins [6]. A sequence $\{X_n, n \geqslant 1\}$ of real-valued r.v.'s converges completely to a constant θ if $\sum_{n=1}^{\infty} P(|X_n - \theta| > \varepsilon) < \infty$ for any $\varepsilon > 0$, and hence it follows from Borel-Cantelli's lemma that $X_n \to \theta$ almost surely. Also in [6], Hsu and Robbins proved that the sequence of arithmetic means of independent and identically distributed (i.i.d.) r.v.'s converges completely to the common expected value if their variance is finite. This result has been considered and extended by many authors. A noteworthy result was obtained by Katz [7] and Baum and Katz [2], that is:

Theorem 1.1 *Let $\{X, X_n, n \geqslant 1\}$ be a sequence of i.i.d. r.v.'s, and set $S_n = \sum_{i=1}^{n} X_i$. Given $p \geqslant 1$ and $0 < r < 2$. Then $E|X|^{pr} < \infty$ if and only if $\sum_{n=1}^{\infty} n^{p-2} P(\max_{1 \leqslant k \leqslant n} |S_k - km| > \varepsilon n^{1/r}) < \infty$ for every $\varepsilon > 0$, where $m = EX$ if $pr \geqslant 1$ and $m = 0$ if $0 < pr < 1$.*

✉ Nguyen Tran Thuan
thuan.nguyen@vinhuni.edu.vn

Nguyen Van Quang
nvquang@hotmail.com

[1] Department of Mathematics, Vinh University, Nghe An, Vietnam

[2] Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, P. O. Box 35, Finland

This result has been extensively studied for various classes of r.v.'s. Recently, Bai et al. [1] considered a particular case of above Baum-Katz's result when $r = 1$ and $1 \leqslant p < 2$, and in this situation, the condition of i.i.d. can be relaxed to be pairwise i.i.d. [1, Theorem 1.2]. Moreover, in Banach space setting, Bai et al. [1] also derived a similar result without any geometric property of the underlying Banach space [1, Theorem 3.2].

Besides considering Baum-Katz's type theorems for various classes of r.v.'s (e.g., pairwise i.i.d. r.v.'s in [1] or martingale, negatively associated r.v.'s, ρ^*-mixing r.v.'s in [8]), many researchers have also extended them into more abstract spaces such as Hilbert spaces [5] or Banach spaces [1]. Continuing this direction, we will discuss Baum-Katz's type theorems in a convex combination space, which is the certain metric space introduced in 2006 by Terán and Molchanov [12]. Roughly speaking, a convex combination space is a metric space endowed with a convex combination operation and the extension from linear spaces to convex combination spaces is not trivial. Some very basic sets, such as singletons and balls, may fail to be convex in this type of metric spaces. To illustrate this demonstration, Terán and Molchanov [12] provided many interesting examples for convex combination spaces, for example, the space of all cumulative distribution functions and the space of upper semicontinuous functions with t-norm. Furthermore, the authors also proved several basic properties of convex combination operation and used those to get the strong law of large numbers for pairwise i.i.d. random elements [12, Theorem 5.1], which extended [4, Theorem 1] of Etemadi. Since then, some limit theorems for random elements taking values in convex combination space were considered and extended (see [9, 11, 12, 14]). On the other hand, as shown recently in [13], it is fairly remarkable that although these spaces are not linear in general, they always contains a subspace which can be isometrically embedded into a Banach space and this embedding preserves the convex combination operation.

In this study, we establish the complete convergence for maximum partial sums of a sequence of random elements in a convex combination space, which gives us some new variants of Baum-Katz's type theorems. Notice that some usual techniques developed in Banach space are no longer applicable here because we are dealing with problems in a nonlinear space. For example, Lemma 2.2 in Section 2 is not necessary if one considers the problems in Banach space. Moreover, an illustrative example will be given to show that some conditions appearing in our results cannot be removed in general convex combination space while they become trivial in Banach space. This paper is organized as follows: In Section 2, we state and summarize some basic results about convex combination spaces, discuss the notion of compactly uniform integrability in Cesàro sense and present some auxiliary lemmas. Our main results regarding to Baum-Katz's type theorems for pairwise independent random elements taking values in convex combination space are established in Section 3.

2 Preliminaries

Throughout this paper, $(\Omega, \mathcal{A}, P)$ is a complete probability space. For $A \in \mathcal{A}$, the notation $I(A)$ (or I_A) is the indicator function of A, the symbol C denotes a general positive constant and it is probably not the same in each appearance.

For the reader's convenience, we now present a short introduction to the work given by Terán and Molchanov [12]. Let $(\mathcal{X}, d)$ be a metric space. Denote $\|x\|_u := d(u, x)$ for $u, x \in \mathcal{X}$. Based on $\mathcal{X}$, a *convex combination operation* is defined so that for all $n \geqslant 1$, numbers $\lambda_1, \ldots, \lambda_n > 0$ that satisfy $\sum_{i=1}^n \lambda_i = 1$, and all $u_1, \ldots, u_n \in \mathcal{X}$, this operation

produces an element in $\mathcal{X}$, which is denoted by $[\lambda_1, u_1; \ldots; \lambda_n, u_n]$ or $[\lambda_i, u_i]_{i=1}^n$. Assume that $[1, u] = u$ for every $u \in \mathcal{X}$ and the following properties are satisfied:

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

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

(CC.iii) (Continuity) If $u, v \in \mathcal{X}$ and $\lambda^{(k)} \to \lambda \in (0; 1)$ as $k \to \infty$, then $[\lambda^{(k)}, u; 1 - \lambda^{(k)}, v] \to [\lambda, u; 1 - \lambda, v]$;

(CC.iv) (Negative curvature) If $u_1, u_2, v_1, v_2 \in \mathcal{X}$ and $\lambda \in (0, 1)$, then

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

(CC.v) (Convexification) For each $u \in \mathcal{X}$, there exists $\lim_{n\to\infty}[n^{-1}, u]_{i=1}^n$, which is denoted by $K_{\mathcal{X}}u$ (or Ku without any confusion), and K is called the *convexification operator*.

Then, a metric space endowed with a convex combination operation is referred to the *convex combination space* (shortly, CC space). Notice that $[\lambda_1, u_1; \ldots; \lambda_n, u_n]$ and the shorthand $[\lambda_i, u_i]_{i=1}^n$ have the same intuitive meaning as the more familiar $\lambda_1 u_1 + \cdots + \lambda_n u_n$ and $\sum_{i=1}^n \lambda_i u_i$, but $\mathcal{X}$ is not assumed to have any addition. By induction and (CC.ii), the axiom (CC.iv) can be extended to convex combinations of n elements as follows: if $u_i, v_i \in \mathcal{X}$, $\lambda_i \in (0; 1)$ with $\sum_{i=1}^n \lambda_i = 1$, then $d([\lambda_i, u_i]_{i=1}^n, [\lambda_i, v_i]_{i=1}^n) \leqslant \sum_{i=1}^n \lambda_i d(u_i, v_i)$. The following properties (2.1)–(2.6) are implied from (CC.i)–(CC.v) above, and their proofs were given in [12]:

(2.1) For every $u_{11}, \ldots, u_{mn} \in \mathcal{X}$ and $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n > 0$ with $\sum_{i=1}^m \alpha_i = \sum_{j=1}^n \beta_j = 1$, we have $[\alpha_i, [\beta_j, u_{ij}]_{j=1}^n]_{i=1}^m = [\alpha_i \beta_j, u_{ij}]_{i=1,j=1}^{i=m,j=n}$.

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

(2.3) The convexification operator K is linear, that is $K([\lambda_j, u_j]_{j=1}^n) = [\lambda_j, Ku_j]_{j=1}^n$.

(2.4) If $u \in \mathcal{X}$ and $\lambda_1, \ldots, \lambda_n > 0$ with $\sum_{j=1}^n \lambda_j = 1$, then $K([\lambda_j, u]_{j=1}^n) = Ku = [\lambda_j, Ku]_{j=1}^n$. Hence, K is an idempotent operator on $\mathcal{X}$.

(2.5) For $\lambda_1, \lambda_2, \lambda_3 > 0$ with $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and $u, v \in \mathcal{X}$,

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

(2.6) The mapping K is non-expansive, that is $d(Ku, Kv) \leqslant d(u, v)$.

Let $\lambda_k \subset (0; 1)$, $\lambda_k \to 0$ and $u, v \in \mathcal{X}$. By (CC.iv) and property (2.4), we have

$$d([\lambda_k, Ku; 1 - \lambda_k, Kv], Kv) = d([\lambda_k, Ku; 1 - \lambda_k, Kv], [\lambda_k, Kv; 1 - \lambda_k, Kv]) \leqslant \lambda_k d(Ku, Kv) \to 0$$

as $k \to \infty$. It follows $[\lambda_k, Ku; 1 - \lambda_k, Kv] \to Kv$ and this allows us to extend weights λ_i from $(0; 1)$ to $[0; 1]$ for elements in $K(\mathcal{X})$, it means that we can define $[\lambda_i, x_i]_{i\in\{i:\lambda_i\geqslant 0\}} = [\lambda_i, x_i]_{i\in\{i:\lambda_i>0\}}$, where $x_i \in K(\mathcal{X})$, $\sum_i \lambda_i = 1$.

Suppose that $(\mathcal{X}, d)$ is a metric space. A mapping $X\colon \Omega \to \mathcal{X}$ is called an $\mathcal{X}$-valued random element (or $\mathcal{A}$-measurable) if $X^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}(\mathcal{X})$, where $\mathcal{B}(\mathcal{X})$ is the Borel σ-algebra on $\mathcal{X}$. When an $\mathcal{X}$-valued random element X takes finite values, it is called a simple random element.

The distribution P_X of an $\mathcal{X}$-valued random element X is defined by $P_X(B) = P(X^{-1}(B))$, $\forall B \in \mathcal{B}(\mathcal{X})$, and two $\mathcal{X}$-valued random elements X, Y are said to be identically distributed if $P_X = P_Y$. The collection of $\mathcal{X}$-valued random elements $\{X_i, i \in I\}$ is said to be independent (resp. pairwise independent) if the collection of σ-algebras $\{\sigma(X_i), i \in I\}$ is independent (resp. pairwise independent), where $\sigma(X) = \{X^{-1}(B), B \in \mathcal{B}(\mathcal{X})\}$.

In the sequel, we assume that $(\mathcal{X}, d)$ is a separable and complete CC space. According to (CC.v), the set $K(\mathcal{X})$ is nonempty, and hence an element $u_0 \in K(\mathcal{X})$ is fixed. Since $\mathcal{X}$ is separable, there exists a countable dense subset $\{u_n, n \geqslant 1\}$ of $\mathcal{X}$. For each $k \geqslant 1$, we define the mapping $\varphi_k : \mathcal{X} \to \mathcal{X}$ by setting $\varphi_k(x) = u_{m_k(x)}$, where $m_k(x) = \min\{i \in \{0, \ldots, k\} : d(u_i, x) = \min_{0 \leqslant j \leqslant k} d(u_j, x)\}$.

The expectation for an integrable $\mathcal{X}$-valued random element is constructed via approximation as follows. For a simple random element $X = [I_{\Omega_i}, x_i]_{i=1}^n$, the *expectation* of X is defined by $EX = [P(\Omega_i), Kx_i]_{i=1}^n$. It is easy to prove that if X, Y are simple random elements, then $d(EX, EY) \leqslant Ed(X, Y)$. A random element $X : \Omega \to \mathcal{X}$ is said to be *integrable* if $d(u, X)$ is an integrable real-valued random variable for some $u \in \mathcal{X}$, and the space of all integrable $\mathcal{X}$-valued random elements is denoted by $L^1_{\mathcal{X}}$. Since $\mathcal{X}$ is separable and complete, any integrable $\mathcal{X}$-valued random element can be approximated by a sequence of simple random elements. Namely, if $X \in L^1_{\mathcal{X}}$ then $X = \lim_{k\to\infty} \varphi_k(X)$, and the *expectation* of X is defined by $EX := \lim_{k\to\infty} E\varphi_k(X)$. Based on the approximation, we can also prove that $d(EX, EY) \leqslant Ed(X, Y)$ whenever $X, Y \in L^1_{\mathcal{X}}$.

A set $A \subset \mathcal{X}$ is said to be *convex* if $[\lambda_i, u_i]_{i=1}^n \in A$ for all $u_i \in A$ and any positive numbers λ_i that sum up to 1. The *convex hull* of $A \subset \mathcal{X}$, denoted by coA, is the smallest convex subset of $\mathcal{X}$ containing A, and $\overline{co}A$ denotes the closed convex hull of A. Let $k(\mathcal{X})$ denote the set of nonempty compact subsets of $\mathcal{X}$. It follows from [12, Theorem 6.2] that if $\mathcal{X}$ is a separable complete CC space, then the space $k(\mathcal{X})$ with the convex combination

$$[\lambda_i, A_i]_{i=1}^n = \{[\lambda_i, u_i]_{i=1}^n : u_i \in A_i, \text{ for all } i\}$$

and the Hausdorff metric d_H

$$d_H(A, B) = \max\left\{\sup_{a\in A}\inf_{b\in B} d(a, b), \sup_{b\in B}\inf_{a\in A} d(b, a)\right\}$$

is a separable complete CC space as well, where the convexification operator $K_{k(\mathcal{X})}$ is given by

$$K_{k(\mathcal{X})}A = \overline{co}K_{\mathcal{X}}(A) = \overline{co}\{K_{\mathcal{X}}u : u \in A\}.$$

This is a nice feature of CC space. Based on this property, if a result holds in CC space, then it can be uplifted to the space of nonempty compact subsets. Further details can be found in [11–13].

The notion of compactly uniform integrability in Cesàro sense for a collection of random elements taking values in Banach spaces was discussed by many authors (see, e.g., [1, 3, 15]). We now introduce this notion in metric spaces, which is also naturally extended from Banach spaces. Let $r > 0$, then a sequence $\{X_n, n \geqslant 1\}$ of $\mathcal{X}$-valued random elements is said to be *compactly uniformly r-th order integrable in Cesàro sense* (Cesàro r-th CUI) if there is a $u \in \mathcal{X}$ such that for every $\varepsilon > 0$, there exists a compact subset $\mathcal{K}_\varepsilon$ (depending on u) of $\mathcal{X}$ with

$$\sup_{n\geqslant 1}\left(n^{-1}\sum_{i=1}^n E\left(\|X_i\|_u^r I(X_i \notin \mathcal{K}_\varepsilon)\right)\right) \leqslant \varepsilon.$$

When $r = 1$, we also use the terminology *Cesàro compactly uniformly integrable* or *Cesàro CUI* for the sake of simplicity. The following proposition shows that the notion of Cesàro r-th CUI does not depend on the selection of u.

Proposition 2.1 *A sequence $\{X_n, n \geqslant 1\}$ of $\mathcal{X}$-valued random elements is Cesàro r-th CUI with respect to some element u if and only if it is Cesàro r-th CUI with respect to any element $a \in \mathcal{X}$.*

Proof Suppose that $\{X_n, n \geqslant 1\}$ is Cesàro r-th CUI with respect to u and let a be another element of $\mathcal{X}$. For $\varepsilon > 0$ and for each $m \in \mathbb{N}$, there exists $\mathcal{K}_{\varepsilon,m} \in k(\mathcal{X})$ such that

$$n^{-1}\sum_{i=1}^{n} E\left(\|X_i\|_u^r I(X_i \notin \mathcal{K}_{\varepsilon,m})\right) \leqslant \frac{\varepsilon}{m2^m}, \forall n \geqslant 1.$$

Let $B(x,\delta)$ denote the open ball with center $x \in \mathcal{X}$ and radius δ. We have

$$\begin{aligned}
&n^{-1}\sum_{i=1}^{n} m^{-1} P\left(X_i \notin \left(\mathcal{K}_{\varepsilon,m} \cup B\left(u, m^{-1/r}\right)\right)\right)\\
\leqslant\ & n^{-1}\sum_{i=1}^{n} E\|X_i\|_u^r I\left(X_i \notin \left(\mathcal{K}_{\varepsilon,n} \cup B\left(u, m^{-1/r}\right)\right)\right)\\
\leqslant\ & n^{-1}\sum_{i=1}^{n} E\|X_i\|_u^r I\left(X_i \notin \mathcal{K}_{\varepsilon,m}\right) \leqslant \frac{\varepsilon}{m2^m}.
\end{aligned}$$

Therefore,

$$n^{-1}\sum_{i=1}^{n} P\left(X_i \notin \left(\mathcal{K}_{\varepsilon,m} \cup B\left(u, m^{-1/r}\right)\right)\right) \leqslant \frac{\varepsilon}{2^m}.$$

The compactness of $\mathcal{K}_{\varepsilon,m}$ follows that it can be covered by a finite number of open balls with equal radii $m^{-1/r}$, and so is $\mathcal{K}_{\varepsilon,m} \cup B(u, m^{-1/r})$. Set $\mathcal{K}_\varepsilon = \cap_{m\geqslant 1}(\mathcal{K}_{\varepsilon,m} \cup B(u, m^{-1/r}))$, then $\mathcal{K}_\varepsilon$ is totally bounded in $\mathcal{X}$. Since $\mathcal{X}$ is complete, the closure $\mathrm{cl}\mathcal{K}_\varepsilon$ of $\mathcal{K}_\varepsilon$ in $\mathcal{X}$ is compact. Thus, we obtain

$$\begin{aligned}
&n^{-1}\sum_{i=1}^{n} P\left(X_i \notin \mathrm{cl}\mathcal{K}_\varepsilon\right)\\
\leqslant\ & n^{-1}\sum_{i=1}^{n} P\left(X_i \notin \mathcal{K}_\varepsilon\right) \leqslant n^{-1}\sum_{i=1}^{n} P\left(X_i \notin \cap_{m\geqslant 1}\left(\mathcal{K}_{\varepsilon,m} \cup B\left(u, m^{-1/r}\right)\right)\right)\\
\leqslant\ & n^{-1}\sum_{i=1}^{n} P\left(\cup_{m\geqslant 1}\left(X_i \notin \left(\mathcal{K}_{\varepsilon,m} \cup B\left(u, m^{-1/r}\right)\right)\right)\right) \leqslant \sum_{m=1}^{\infty}\frac{\varepsilon}{2^m} = \varepsilon.
\end{aligned}$$

Denote $\mathbf{K}_\varepsilon = \mathrm{cl}\mathcal{K}_\varepsilon \cup \mathcal{K}_{\varepsilon,1}$, then $\mathbf{K}_\varepsilon \in k(\mathcal{X})$. Hence for all $n \geqslant 1$,

$$\begin{aligned}
&n^{-1}\sum_{i=1}^{n} E\left(\|X_i\|_a^r I(X_i \notin \mathbf{K}_\varepsilon)\right)\\
\leqslant\ & C_r n^{-1}\sum_{i=1}^{n} E(\|X_i\|_u^r I(X_i \notin \mathbf{K}_\varepsilon)) + C_r d^r(u,a)\left(n^{-1}\sum_{i=1}^{n} P(X_i \notin \mathbf{K}_\varepsilon)\right)\\
\leqslant\ & C_r\left(\varepsilon/2 + d^r(u,a)\varepsilon\right).
\end{aligned}$$

By the arbitrariness of $\varepsilon > 0$, the proof is complete. □

By Jensen's inequality, it is easy to see that for $0 < r \leqslant p$,

$$\left(n^{-1}\sum_{i=1}^{n} E\|X_i\|_u^r I(X_i \notin \mathcal{K})\right)^{p/r} \leqslant n^{-1}\sum_{i=1}^{n} E\|X_i\|_u^p I(X_i \notin \mathcal{K}),$$

and this implies that if $\{X_n, n \geqslant 1\}$ is Cesàro p-th CUI, then it is also Cesàro r-th CUI for $0 < r \leqslant p$. Further details about CUI, the readers can refer to [15].

Lemma 2.2 [9, Lemma 3.3] *Let* $\{a_i, b_i, 1 \leqslant i \leqslant n\} \subset [0,1]$ *be a collection of nonnegative constants with* $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1$. *Then* $d([a_i, Kx_i]_{i=1}^n, [b_i, Kx_i]_{i=1}^n) \leqslant \sum_{i=1}^n |a_i - b_i| d(x_i, u)$, *where* $x_1, \ldots, x_n, u \in \mathcal{X}$ *are arbitrary.*

Notice that the inequality $d([a_i, x_i]_{i=1}^n, [b_i, x_i]_{i=1}^n) \leqslant \sum_{i=1}^n |a_i - b_i| d(x_i, u)$ does not hold for $x_1, \ldots, x_n \in \mathcal{X}$ in general as shown in [13, Example 1].

Lemma 2.3 [16, Lemma A.6] *Let* $A_1, \ldots, A_n$ *be events satisfying* $\mathrm{Var}(\sum_{i=1}^n I_{A_i}) \leqslant \alpha \sum_{i=1}^n P(A_i)$, *where positive* α *does not depend on* n. *Then* $\left(1 - P\left(\bigcup_{i=1}^n A_i\right)\right)^2 \sum_{i=1}^n P(A_i) \leqslant \alpha P\left(\cup_{i=1}^n A_i\right)$.

3 Baum-Katz's Type Theorems for Pairwise Independent $\mathcal{X}$-Valued Random Elements

Throughout this section, $\mathcal{X}$ is a complete and separable CC space and u_0 is the fixed element of $K(\mathcal{X})$ as mentioned in Section 2.

In the first theorem, we establish a similar result to [1, Theorem 1.2] in CC space. However, the version in CC space has a significant difference compared to the corresponding version in Banach space that is condition (3.1) below. It becomes trivial when one considers in Banach space with usual convex combination operation; moreover, we also show immediately after the proof that it cannot be removed in general CC space, even when considered random elements are independent.

Theorem 3.1 *Let* $1 \leqslant p < 2$, $a \in K(\mathcal{X})$ *and let* $\{X, X_n, n \geqslant 1\}$ *be a sequence of pairwise i.i.d.* $\mathcal{X}$*-valued random elements. Then* $EX = a$ *and* $E\|X\|_a^p < \infty$ *if and only if*

$$Ed^p(X, KX) < \infty \tag{3.1}$$

and

$$\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1 \leqslant k \leqslant n} kd\left([k^{-1}, X_i]_{i=1}^k, a\right) > n\varepsilon\right) < \infty \text{ for all } \varepsilon > 0. \tag{3.2}$$

Proof **Necessity:** If $EX = a$ and $E\|X\|_a^p < \infty$, then

$$\begin{aligned} Ed^p(X, KX) &\leqslant 2^{p-1}(E\|X\|_a^p + E\|KX\|_a^p) \\ &= 2^{p-1}(E\|X\|_a^p + Ed^p(KX, Ka)) \leqslant 2^p E\|X\|_a^p < \infty. \end{aligned}$$

We now prove that (3.2) holds.

Step 1 Assume that X is simple with values $x_1, x_2, \ldots, x_m$ on non-null sets $\Omega_1, \Omega_2, \ldots, \Omega_m$ respectively. Since $\{X, X_n, n \geqslant 1\}$ is identically distributed, each X_n also takes values $x_1, \ldots, x_m$ a.s. For each $j = 1, \ldots, m$, set

$$\begin{aligned} Z_n^j(\omega) &= \sum_{i=1}^n I(X_i = x_j)(\omega) = \mathrm{card}\{i \in [1,n] : X_i(\omega) = x_j\}, \\ \mathcal{T}_n(\omega) &= \{j : 1 \leqslant j \leqslant m, Z_n^j(\omega) > 0\}, n \geqslant 1. \end{aligned}$$

Then $\{Z_n^j(\omega)\}_{n=1}^{\infty}$ is a non-decreasing sequence for each j and each ω. By (CC.i) and property (2.1),

$$[k^{-1}, X_i]_{i=1}^k = \left[k^{-1} Z_k^j, \left[(Z_k^j)^{-1}, x_j\right]_{i=1}^{Z_k^j}\right]_{j \in \mathcal{T}_k} \text{ a.s.}$$

Therefore,

$$
\begin{aligned}
d\left([k^{-1}, X_i]_{i=1}^k, EX\right) &= d\left(\left[k^{-1}Z_k^j, \left[(Z_k^j)^{-1}, x_j\right]_{i=1}^{Z_k^j}\right]_{j\in\mathcal{T}_k}, \left[P(\Omega_j), Kx_j\right]_{j=1}^m\right)\\
\leqslant\ & d\left(\left[k^{-1}Z_k^j, \left[(Z_k^j)^{-1}, x_j\right]_{i=1}^{Z_k^j}\right]_{j\in\mathcal{T}_k}, \left[k^{-1}Z_k^j, Kx_j\right]_{j\in\mathcal{T}_k}\right)\\
&+d\left(\left[k^{-1}Z_k^j, Kx_j\right]_{j\in\mathcal{T}_k}, \left[P(\Omega_j), Kx_j\right]_{j=1}^m\right)\\
\leqslant\ & \sum_{j\in\mathcal{T}_k} k^{-1}Z_k^j d\left(\left[(Z_k^j)^{-1}, x_j\right]_{i=1}^{Z_k^j}, Kx_j\right)\\
&+\sum_{j=1}^m \left|k^{-1}Z_k^j - P(\Omega_j)\right| \|Kx_j\|_{u_0} \ \text{(by (CC.iv) and Lemma 2.2)}\\
\leqslant\ & \sum_{j\in\mathcal{T}_k} k^{-1}Z_k^j d\left(\left[(Z_k^j)^{-1}, x_j\right]_{i=1}^{Z_k^j}, Kx_j\right) + M\sum_{j=1}^m \left|k^{-1}Z_k^j - P(\Omega_j)\right|,
\end{aligned}
$$

where $M := \max_{1\leqslant j\leqslant m}\|x_j\|_{u_0}$. It implies

$$
\begin{aligned}
n^{-1}\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, X_i]_{i=1}^k, EX\right) \leqslant\ & \max_{1\leqslant k\leqslant n}\sum_{j\in\mathcal{T}_k}\frac{Z_k^j}{n}d\left(\left[(Z_k^j)^{-1}, x_j\right]_{i=1}^{Z_k^j}, Kx_j\right)\\
&+M\sum_{j=1}^m n^{-1}\max_{1\leqslant k\leqslant n}\left|Z_k^j - kP(\Omega_j)\right|\\
:=\ & (I_1)+(I_2).
\end{aligned}
$$

Next, we show that $(I_1) < \varepsilon/2$ for all $\omega\in\Omega$ when n is sufficiently large. Indeed, by the definition of the operator K,

$$\lim_{n\to\infty} d([n^{-1}, x_j]_{i=1}^n, Kx_j) = 0$$

for each $j = 1, \ldots, m$. Thus, there exists $n_0(\varepsilon, m)\in\mathbb{N}$ such that for all $n\geqslant n_0(\varepsilon, m)$ and for all $j = 1, \ldots, m$,

$$d([n^{-1}, x_j]_{i=1}^n, Kx_j) < \frac{\varepsilon}{2m}. \tag{3.3}$$

Denote

$$N(\varepsilon, m) = \max_{1\leqslant j\leqslant m}\max_{1\leqslant k<n_0(\varepsilon,m)} d\left([k^{-1}, x_j]_{i=1}^k, Kx_j\right)$$

and let $n(\varepsilon, m)$ be the smallest integer number such that $n(\varepsilon, m)\geqslant 2\varepsilon^{-1}mN(\varepsilon, m)n_0(\varepsilon, m)$. Now, for any $n\geqslant n(\varepsilon, m)$ and for each $k = 1, \ldots, n$, each $\omega\in\Omega$:
If $Z_k^j(\omega)\geqslant n_0(\varepsilon, m)$, then it follows from (3.3) and $n^{-1}Z_k^j(\omega)\leqslant n^{-1}Z_n^j(\omega)\leqslant 1$ that

$$\frac{Z_k^j(\omega)}{n}d\left(\left[(Z_k^j(\omega))^{-1}, x_j\right]_{i=1}^{Z_k^j(\omega)}, Kx_j\right) < \frac{\varepsilon}{2m}.$$

If $0 < Z_k^j(\omega) < n_0(\varepsilon, m)$, then

$$\frac{Z_k^j(\omega)}{n} d\left(\left[(Z_k^j(\omega))^{-1}, x_j\right]_{i=1}^{Z_k^j(\omega)}, Kx_j\right) < \frac{n_0(\varepsilon, m)}{n(\varepsilon, m)} N(\varepsilon, m) \leqslant \frac{\varepsilon}{2m}.$$

Hence, for $n \geqslant n(\varepsilon, m)$

$$\frac{Z_k^j(\omega)}{n} d\left(\left[(Z_k^j(\omega))^{-1}, x_j\right]_{i=1}^{Z_k^j(\omega)}, Kx_j\right) \leqslant \frac{\varepsilon}{2m}.$$

This implies that

$$(I_1) = \max_{1\leqslant k\leqslant n} \sum_{j\in\mathcal{T}_k} \frac{Z_k^j}{n} d\left(\left[(Z_k^j)^{-1}, x_j\right]_{i=1}^{Z_k^j}, Kx_j\right) \leqslant \max_{1\leqslant k\leqslant n} \sum_{j\in\mathcal{T}_k} \frac{\varepsilon}{2m} \leqslant \max_{1\leqslant k\leqslant n} \sum_{j=1}^{m} \frac{\varepsilon}{2m} = \frac{\varepsilon}{2}$$

for all $n \geqslant n(\varepsilon, m)$. Therefore, $\sum_{n=1}^{\infty} n^{p-2} P((I_1) > \varepsilon/2) < \infty$.
For (I_2), we have

$$P((I_2) > \varepsilon/2) \leqslant \sum_{j=1}^{m} P\left(n^{-1} \max_{1\leqslant k\leqslant n} \left|Z_k^j - kP(\Omega_j)\right| > \frac{\varepsilon}{2Mm}\right).$$

For each $j = 1, \ldots, m$, applying [1, Theorem 1.2] for sequence $\{I(X = x_j), I(X_n = x_j), n \geqslant 1\}$ of pairwise i.i.d. and uniformly bounded r.v.'s to get

$$\sum_{n=1}^{\infty} n^{p-2} P\left(n^{-1} \max_{1\leqslant k\leqslant n} \left|Z_k^j - kP(\Omega_j)\right| > \frac{\varepsilon}{2Mm}\right) < \infty.$$

Combining above arguments,

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, X_i]_{i=1}^k, EX\right) > n\varepsilon\right) \\
&\leqslant \sum_{n=1}^{\infty} n^{p-2} P((I_1) > \varepsilon/2) + \sum_{n=1}^{\infty} n^{p-2} P((I_2) > \varepsilon/2) \\
&\leqslant \sum_{n=1}^{\infty} n^{p-2} P((I_1) > \varepsilon/2) + \sum_{j=1}^{m} \sum_{n=1}^{\infty} n^{p-2} P\left(n^{-1} \max_{1\leqslant k\leqslant n} \left|Z_k^j - kP(\Omega_j)\right| > \frac{\varepsilon}{2Mm}\right) < \infty,
\end{aligned}$$

which means that (3.2) holds for this case.

Step 2 Let us consider the general case when $X \in L_{\mathcal{X}}^1$. For $\varepsilon > 0$ arbitrarily, [12, Proposition 4.1] implies that there exists a natural number h large enough such that $Ed(\varphi_h(X), X) \leqslant \varepsilon/6$, where the function φ_h was mentioned in Section 2. Then $\{\varphi_h(X), \varphi_h(X_n), n \geqslant 1\}$ is a collection of pairwise i.i.d. and simple random elements with common expectation $E\varphi_h(X) := b$. Moreover,

$$E\|\varphi_h(X)\|_b^p \leqslant 2^{p-1}(E\|\varphi_h(X)\|_{u_0}^p + d^p(u_0, b)) \leqslant 2^{2p-1} E\|X\|_{u_0}^p + 2^{p-1} d^p(u_0, b) < \infty.$$

It follows from the first case that

$$\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, \varphi_h(X_i)]_{i=1}^k, E\varphi_h(X)\right) > n\varepsilon/3\right) < \infty.$$

Similarly, $\{d(X, \varphi_h(X)), d(X_n, \varphi_h(X_n)), n \geqslant 1\}$ is also a collection of pairwise i.i.d. real-valued r.v.'s satisfying

$$Ed^p(X, \varphi_h(X)) \leqslant 2^{p-1}(E\|X\|_{u_0}^p + E\|\varphi_h(X)\|_{u_0}^p) \leqslant CE\|X\|_{u_0}^p < \infty.$$

As a corollary of [1, Theorem 1.2],

$$\sum_{n=1}^{\infty} n^{p-2} P\left(\left|\sum_{i=1}^{n}(d(X_i, \varphi_h(X_i)) - Ed(X_i, \varphi_h(X_i)))\right| > n\varepsilon/6\right) < \infty.$$

By the triangle inequality,

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, X_i]_{i=1}^{k}, EX\right) > n\varepsilon\right)\\
\leqslant &\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, X_i]_{i=1}^{k}, [k^{-1}, \varphi_h(X_i)]_{i=1}^{k}\right) > n\varepsilon/3\right)\\
&+\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, \varphi_h(X_i)]_{i=1}^{k}, [k^{-1}, E\varphi_h(X_i)]_{i=1}^{k}\right) > n\varepsilon/3\right)\\
&+\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, E\varphi_h(X_i)]_{i=1}^{k}, EX\right) > n\varepsilon/3\right)\\
\leqslant &\sum_{n=1}^{\infty} n^{p-2} P\left(\left|\sum_{i=1}^{n}(d(X_i, \varphi_h(X_i)) - Ed(X_i, \varphi_h(X_i)))\right| > n\varepsilon/6\right)\\
&+\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, \varphi_h(X_i)]_{i=1}^{k}, [k^{-1}, E\varphi_h(X_i)]_{i=1}^{k}\right) > n\varepsilon/3\right)\\
&+\sum_{n=1}^{\infty} n^{p-2} P(d(EX, E\varphi_h(X)) > \varepsilon/3)\\
< &\infty,
\end{aligned}$$

and this completes the necessity part.

Sufficiency: Suppose that $Ed^p(X, KX) < \infty$ and (3.2) holds for some $a \in K(\mathcal{X})$. By the triangle inequality,

$$\begin{aligned}
&\|KX_n\|_a\\
\leqslant &\, d(KX_n, [n^{-1}, X_i]_{i=1}^{n}) + d([n^{-1}, X_i]_{i=1}^{n}, a)\\
= &\, d\left(\left[\frac{n-1}{n}, KX_n; \frac{1}{n}, KX_n\right], \left[\frac{n-1}{n}, \left[\frac{1}{n-1}, X_i\right]_{i=1}^{n-1}; \frac{1}{n}, X_n\right]\right) + d([n^{-1}, X_i]_{i=1}^{n}, a)\\
\leqslant &\, \frac{n-1}{n} d([(n-1)^{-1}, X_i]_{i=1}^{n-1}, KX_n) + \frac{1}{n} d(X_n, KX_n) + d([n^{-1}, X_i]_{i=1}^{n}, a)\\
\leqslant &\, \frac{n-1}{n} d([(n-1)^{-1}, X_i]_{i=1}^{n-1}, a) + \frac{n-1}{n}\|KX_n\|_a + \frac{1}{n} d(X_n, KX_n) + d([n^{-1}, X_i]_{i=1}^{n}, a).
\end{aligned}$$

This is equivalent to

$$\|KX_n\|_a \leqslant d(X_n, KX_n) + (n-1)d([(n-1)^{-1}, X_i]_{i=1}^{n-1}, a) + nd([n^{-1}, X_i]_{i=1}^{n}, a).$$

Thus

$$\max_{1\leqslant k\leqslant n}\|KX_k\|_a \leqslant \max_{1\leqslant k\leqslant n} d(X_k, KX_k) + 2\max_{1\leqslant k\leqslant n} kd([k^{-1}, X_i]_{i=1}^k, a).$$

On the other hand, for any $\alpha > 0$, it follows from $Ed^p(X, KX) < \infty$ that

$$\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} d(X_k, KX_k) > n\alpha\right) \leqslant \sum_{n=1}^{\infty} n^{p-2} \sum_{k=1}^{n} P(d(X_k, KX_k) > n\alpha)$$

$$= \sum_{n=1}^{\infty} n^{p-1} P(d(X, KX) > n\alpha) \leqslant CEd^p(X, KX) < \infty.$$

Therefore, combining with (3.2) we obtain

$$\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n}\|KX_k\|_a > n\beta\right)$$

$$\leqslant \sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} d(X_k, KX_k) > n\beta/3\right)$$

$$+ \sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, X_i]_{i=1}^k, a\right) > n\beta/3\right) < \infty$$

for every $\beta > 0$. It follows

$$P\left(\max_{1\leqslant k\leqslant n}\|KX_k\|_a > n\right) \to 0 \text{ as } n \to \infty,$$

and hence $P(\max_{1\leqslant k\leqslant n}\|KX_k\|_a > n) \leqslant \frac{1}{2}$ for all $n \geqslant N_0$. Applying Lemma 2.3 for sequence $\{(\|KX_k\|_a > n), 1 \leqslant k \leqslant n, n \geqslant 1\}$ of pairwise independent events with $\alpha = 1$, we have that for every $n \geqslant N_0$,

$$\sum_{k=1}^{n} P(\|KX_k\|_a > n) \leqslant \frac{P\left(\max_{1\leqslant k\leqslant n}\|KX_k\|_a > n\right)}{\left(1 - P\left(\max_{1\leqslant k\leqslant n}\|KX_k\|_a > n\right)\right)^2} \leqslant 4P\left(\max_{1\leqslant k\leqslant n}\|KX_k\|_a > n\right).$$

Hence

$$E\|KX\|_a^p \leqslant C\sum_{n=1}^{\infty} n^{p-1} P(\|KX\|_a > n) = C\sum_{n=1}^{\infty} n^{p-2} \sum_{k=1}^{n} P(\|KX_k\|_a > n)$$

$$\leqslant C\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n}\|KX_k\|_a > n\right) < \infty$$

and $E\|X\|_a^p \leqslant 2^{p-1}(Ed^p(X, KX) + E\|KX\|_a^p) < \infty$. This implies the existence of EX, then applying the necessity part

$$\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, X_i]_{i=1}^k, EX\right) > n\varepsilon\right) < \infty. \tag{3.4}$$

Combining (3.2) with (3.4), we obtain $EX = a$. □

Remark 3.2 If $\mathcal{X}$ is a separable Banach space with usual convex combination $[\lambda_i, x_i]_{i=1}^n = \sum_{i=1}^n \lambda_i x_i$, then the condition (3.1) is trivial due to $KX = X$. However, in general CC space, the condition (3.1) cannot be removed as shown in the example below:

Example 3.3 Let p, r be real numbers such that $1 < r < p < 2$, $r(r-1) \neq 1$ and $r^2 > p$. Assume that $(\mathcal{X}, \|.\|)$ is a Rademacher type r Banach space and denote by d the metric associated with its norm $\|.\|$. An operation ${}^r[.,.]$ is defined based on $\mathcal{X}$ as follows: ${}^r[\lambda_i, x_i]_{i=1}^n = \sum_{i=1}^n \lambda_i^r x_i$. As shown in [12, Example 5], ${}^r[.,.]$ is the convex combination operation (r-th power combination) and the corresponding convexification operator $K_r u = 0$ for all $u \in \mathcal{X}$. Assume that $\{X, X_n, n \geqslant 1\}$ is a collection of i.i.d. $\mathcal{X}$-valued random elements satisfying $E\|X\|^r < \infty$ but $E\|X\|^p = \infty$. Now we show that the condition (3.2) holds while (3.1) does not. It follows from $K_r X = 0$ that $Ed^p(X, K_r X) = E\|X\|^p = \infty$, thus (3.1) fails. Since $E\|X\| \leqslant 1 + E\|X\|^r < \infty$, there exists the expectation of X with respect to ${}^r[.,.]$, denoted by $E_r X$, and $a := E_r X = 0$. Now for $\varepsilon > 0$, applying the Hájek-Rényi inequality for the collection $\{X, X_n, n \geqslant 1\}$ of i.i.d. $\mathcal{X}$-valued random elements with $EX = 0$ as in [10],

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd({}^r[k^{-1}, X_i]_{i=1}^k, a) > n\varepsilon\right) = \sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} \frac{1}{k^{r-1}} \left\|\sum_{i=1}^k X_i\right\| > n\varepsilon\right)\\
&\leqslant \sum_{n=1}^{\infty} n^{p-2} \cdot \frac{C}{n^r \varepsilon^r} \sum_{k=1}^{n} \frac{E\|X_i\|^r}{k^{r(r-1)}} = CE\|X\|^r \sum_{n=1}^{\infty} n^{p-r-2} \sum_{k=1}^{n} \frac{1}{k^{r(r-1)}}\\
&\leqslant CE\|X\|^r \sum_{n=1}^{\infty} n^{p-r-2} \cdot \frac{1}{n^{r(r-1)-1}} = CE\|X\|^r \sum_{n=1}^{\infty} \frac{1}{n^{r^2-p+1}} < \infty,
\end{aligned}$$

which means that (3.2) holds.

Remark 3.4 (a) Since the condition (3.2) implies $d([n^{-1}, X_i]_{i=1}^n, a) \to 0$ a.s. as $n \to \infty$, we immediately obtain the strong law of large numbers for $\{X, X_n, n \geqslant 1\}$ of pairwise i.i.d. random elements in CC space from the condition $X \in L^1_{\mathcal{X}}$. Therefore, [12, Theorem 5.1] is a particular case of Theorem 3.1.
(b) Notice that some arguments in the proof of Theorem 3.1 can be also obtained by combining the embedding theorem [13, Theorem 3.3] and corresponding results in Banach space. However, this embedding is not too straightforward since its proof requires several intermediate results, while the proof of Theorem 3.1 presented above is more direct.

Using the similar technique as in Theorem 3.1 and applying Theorem 1.1 with $r = 1$, we obtain the following result for the case of i.i.d. random elements:

Theorem 3.5 *Let $p \geqslant 1$, $a \in K(\mathcal{X})$ and let $\{X, X_n, n \geqslant 1\}$ be a sequence of i.i.d. $\mathcal{X}$-valued random elements. Then $EX = a$ and $E\|X\|_a^p < \infty$ if and only if*

$$\begin{aligned}
&Ed^p(X, KX) < \infty\\
\text{and}\quad &\sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, X_i]_{i=1}^k, a\right) > n\varepsilon\right) < \infty \text{ for all } \varepsilon > 0.
\end{aligned}$$

Next, we establish some results on complete convergence and L^r-convergence for non-identically distributed random elements.

Proposition 3.6 *Let $\{X_n, n \geqslant 1\}$ be a sequence of pairwise independent $\mathcal{X}$-valued random elements. If there is a compact subset $\mathcal{K}$ of $\mathcal{X}$ such that $P(X_n \in \mathcal{K}) = 1$ for all n, then the following statements hold:*

(i) $\sum_{n=1}^{\infty} n^{-1} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, X_i]_{i=1}^{k}, [k^{-1}, EX_i]_{i=1}^{k}\right) > n\varepsilon\right) < \infty$ *for every* $\varepsilon > 0$.

(ii) $Ed^r\left([n^{-1}, X_i]_{i=1}^{n}, [n^{-1}, EX_i]_{i=1}^{n}\right) \to 0$ *as* $n \to \infty$ *for* $r \geqslant 1$.

Proof For $\varepsilon > 0$, by the compactness of $\mathcal{K}$, there exists $\{c_1, c_2, \ldots, c_m\} \subset \mathcal{K}$ such that

$$\mathcal{K} \subset \bigcup_{j=1}^{m} B(c_j, \varepsilon/4), \text{ where } B(c_j, r) = \{x \in \mathcal{X} : d(x, c_j) < r\}.$$

For $n \geqslant 1$, define a sequence of $\mathcal{X}$-valued random elements as follows:

$$Y_n(\omega) = \begin{cases} c_0 := u_0 & \text{if } X_n(\omega) \notin \mathcal{K} \\ c_1 & \text{if } X_n(\omega) \in B(c_1, \varepsilon/4) \cap \mathcal{K} \\ c_j & \text{if } X_n(\omega) \in B(c_j, \varepsilon/4) \cap \{\cup_{t=1}^{j-1} B(c_t, \varepsilon/4)\}^c \cap \mathcal{K},\ j = 2, \ldots, m. \end{cases}$$

It is obvious that the sequence $\{Y_n, n \geqslant 1\}$ is also pairwise independent. By the triangle inequality,

$$\begin{aligned} & d\left([n^{-1}, X_i]_{i=1}^{n}, [n^{-1}, EX_i]_{i=1}^{n}\right) \\ \leqslant\ & d\left([n^{-1}, X_i]_{i=1}^{n}, [n^{-1}, Y_i]_{i=1}^{n}\right) + d\left([n^{-1}, Y_i]_{i=1}^{n}, [n^{-1}, KY_i]_{i=1}^{n}\right) \\ & + d\left([n^{-1}, KY_i]_{i=1}^{n}, [n^{-1}, EY_i]_{i=1}^{n}\right) + d\left([n^{-1}, EY_i]_{i=1}^{n}, [n^{-1}, EX_i]_{i=1}^{n}\right) \\ :=\ & A_n + B_n + C_n + D_n. \end{aligned}$$

Modifying the proof of [14, Proposition 3.1], we obtain $\operatorname{ess\,sup}_{\omega\in\Omega} A_n(\omega) \leqslant \varepsilon/4$, $\operatorname{ess\,sup}_{\omega\in\Omega} D_n(\omega) \leqslant \varepsilon/4$ for all n. Using the same arguments as in the proof of (I_1) in Theorem 3.1 (necessity part) to get $\operatorname{ess\,sup}_{\omega\in\Omega} B_n(\omega) \leqslant \varepsilon/4$ for n large enough. For C_n, by property (2.1) and Lemma 2.2, we have

$$\begin{aligned} C_n &= d\left([n^{-1}, KY_i]_{i=1}^{n}, [n^{-1}, EY_i]_{i=1}^{n}\right) \\ &= d\left([n^{-1}, [I(Y_i = c_j), Kc_j]_{j=0}^{m}]_{i=1}^{n}, [n^{-1}, [P(Y_i = c_j), Kc_j]_{j=0}^{m}]_{i=1}^{n}\right) \\ &\leqslant \sum_{j=0}^{m} \left|\frac{1}{n}\sum_{i=1}^{n}\left(I(Y_i = c_j) - P(Y_i = c_j)\right)\right| \|c_j\|_{u_0} \\ &\leqslant M \sum_{j=1}^{m} \left|\frac{1}{n}\sum_{i=1}^{n}\left(I(Y_i = c_j) - P(Y_i = c_j)\right)\right|, \end{aligned}$$

where $M := \max_{1\leqslant j\leqslant m} \|c_j\|_{u_0}$.

For conclusion (i): Notice that $\left\{I(Y_n = c_j) - P(Y_n = c_j), n \geqslant 1\right\}$ is a sequence of pairwise independent and uniformly bounded r.v.'s. It follows from [1, Theorem 1.1] that

$$\sum_{n=1}^{\infty} n^{-1} P\left(\max_{1\leqslant k\leqslant n} \left|\sum_{i=1}^{k}\left(I(Y_k = c_j) - P(Y_k = c_j)\right)\right| > n\varepsilon/4Mm\right) < \infty.$$

Thus,

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^{-1} P\left(\max_{1\leqslant k\leqslant n} kC_k > n\varepsilon/4\right)\\
&\leqslant \sum_{n=1}^{\infty} n^{-1} P\left(\sum_{j=1}^{m} \max_{1\leqslant k\leqslant n}\left|\sum_{i=1}^{k}\left(I(Y_k = c_j) - P(Y_k = c_j)\right)\right| > n\varepsilon/4M\right)\\
&\leqslant \sum_{j=1}^{m}\sum_{n=1}^{\infty} n^{-1} P\left(\max_{1\leqslant k\leqslant n}\left|\sum_{i=1}^{k}\left(I(Y_k = c_j) - P(Y_k = c_j)\right)\right| > n\varepsilon/4Mm\right) < \infty.
\end{aligned}$$

Combining the parts above, we have the desired result.
For conclusion (ii): By Jensen's inequality and for n large enough,

$$\begin{aligned}
&Ed^r\left([n^{-1}, X_i]_{i=1}^n, [n^{-1}, EX_i]_{i=1}^n\right)\\
&\leqslant 4^{r-1}\left(3(\varepsilon/4)^r + M^r m^{r-1}\sum_{j=1}^{m} E\left|\frac{1}{n}\sum_{i=1}^{n}\left(I(Y_i = c_j) - P(Y_i = c_j)\right)\right|^r\right)\\
&\leqslant 3\varepsilon^r/4 + 4^{r-1}M^r m^{r-1}\sum_{j=1}^{m} 2^{r-1} E\left|\frac{1}{n}\sum_{i=1}^{n}\left(I(Y_i = c_j) - P(Y_i = c_j)\right)\right|\\
&\leqslant 3\varepsilon^r/4 + C\sum_{j=1}^{m}\frac{1}{n}\left(E\left|\sum_{i=1}^{n}\left(I(Y_i = c_j) - P(Y_i = c_j)\right)\right|^2\right)^{1/2}\\
&\leqslant 3\varepsilon^r/4 + Cn^{-1/2}.
\end{aligned}$$

Letting $n \to \infty$ and by the arbitrariness of $\varepsilon > 0$, we derive the desired conclusion. □

Theorem 3.7 *Let $\{X_n, n \geqslant 1\}$ be a sequence of pairwise independent $\mathcal{X}$-valued random elements.*
(a) *If $\{X_n, n \geqslant 1\}$ is Cesàro CUI and $\sum_{n=1}^{\infty} n^{-p} E\|X_n\|_u^p < \infty$ for some $p \in [1, 2]$, some $u \in \mathcal{X}$, then*

$$\sum_{n=1}^{\infty} n^{-1} P\left(\max_{1\leqslant k\leqslant n} kd\left([k^{-1}, X_i]_{i=1}^k, [k^{-1}, EX_i]_{i=1}^k\right) > n\varepsilon\right) < \infty$$

for all $\varepsilon > 0$. In particular, $d([n^{-1}, X_i]_{i=1}^n, [n^{-1}, EX_i]_{i=1}^n) \to 0$ a.s. as $n \to \infty$.
(b) *If $\{X_n, n \geqslant 1\}$ is Cesàro r-th CUI ($r \geqslant 1$), then $Ed^r\left([n^{-1}, X_i]_{i=1}^n, [n^{-1}, EX_i]_{i=1}^n\right) \to 0$ as $n \to \infty$.*

Proof Given $\varepsilon > 0$. By Cesàro r-th CUI assumption (with $r = 1$ in the case (a)), there exists a compact subset $\mathcal{K} := \mathcal{K}_{\varepsilon,u}$ of $\mathcal{X}$ satisfying

$$n^{-1}\sum_{i=1}^{n} E(\|X_i\|_u^r I(X_i \notin \mathcal{K})) \leqslant \varepsilon/4 \text{ for all } n.$$

For each $n \geqslant 1$, we define a sequence of $\mathcal{X}$-valued random elements by setting

$$Y_n(\omega) = \begin{cases} X_n(\omega) & \text{if } X_n(\omega) \in \mathcal{K} \\ u & \text{if } X_n(\omega) \notin \mathcal{K}. \end{cases}$$

It is obvious that the sequence $\{Y_n, n \geqslant 1\}$ is also pairwise independent. By the triangle inequality,

$$\begin{aligned} d\left([n^{-1}, X_i]_{i=1}^n, [n^{-1}, EX_i]_{i=1}^n\right) &\leqslant d\left([n^{-1}, X_i]_{i=1}^n, [n^{-1}, Y_i]_{i=1}^n\right) \\ &\quad + d\left([n^{-1}, Y_i]_{i=1}^n, [n^{-1}, EY_i]_{i=1}^n\right) \\ &\quad + d\left([n^{-1}, EY_i]_{i=1}^n, [n^{-1}, EX_i]_{i=1}^n\right) \\ &:= Q_n + R_n + S_n. \end{aligned}$$

Let us bound three parts above as follows:
For Q_n: We have

$$Q_n \leqslant \frac{1}{n}\sum_{i=1}^n d(X_i, Y_i) = \frac{1}{n}\sum_{i=1}^n \|X_i\|_u I(X_i \notin \mathcal{K}).$$

For R_n: It is clear that $\{Y_n, n \geqslant 1\}$ is a sequence of pairwise independent $\mathcal{X}$-valued random elements and $Y_n \in \mathcal{K} \cup \{u\}$ for all n. Applying Proposition 3.6, we get that

$$\sum_{n=1}^\infty n^{-1} P\left(\max_{1\leqslant k\leqslant n} kR_k > n\varepsilon/4\right) < \infty \text{ and } E(R_n)^r \to 0 \text{ as } n \to \infty.$$

For S_n: Since $\{X_n, n \geqslant 1\}$ is Cesàro r-th CUI ($r \geqslant 1$), Jensen's inequality yields

$$\begin{aligned} (S_n)^r &\leqslant \left(\frac{1}{n}\sum_{i=1}^n Ed(Y_i, X_i)\right)^r = \left(\frac{1}{n}\sum_{i=1}^n E\left(\|X_i\|_u I(X_i \notin \mathcal{K})\right)\right)^r \\ &\leqslant \frac{1}{n}\sum_{i=1}^n E\left(\|X_i\|_u^r I(X_i \notin \mathcal{K})\right) \leqslant \varepsilon/4. \end{aligned}$$

Proof of part (a). It is sufficient to show $\sum_{n=1}^\infty n^{-1} P(\max_{1\leqslant k\leqslant n} kQ_k > n\varepsilon/2) < \infty$. Indeed,

$$\begin{aligned} &\sum_{n=1}^\infty n^{-1} P\left(\max_{1\leqslant k\leqslant n} kQ_k > n\varepsilon/2\right) \leqslant \sum_{n=1}^\infty n^{-1} P\left(\sum_{i=1}^n \|X_i\|_u I(X_i \notin \mathcal{K}) > n\varepsilon/2\right) \\ &\leqslant \sum_{n=1}^\infty n^{-1} P\left(\left|\sum_{i=1}^n \left(\|X_i\|_u I(X_i \notin \mathcal{K}) - E\left(\|X_i\|_u I(X_i \notin \mathcal{K})\right)\right)\right| > n\varepsilon/4\right) < \infty, \end{aligned}$$

where the last quantity is finite by applying [1, Lemma 2.1] for sequence $\{\|X_n\|_u I(X_n \notin \mathcal{K}), n \geqslant 1\}$.
Proof of part (b). By using Jensen's inequality again,

$$\begin{aligned} &Ed^r\left([n^{-1}, X_i]_{i=1}^n, [n^{-1}, EX_i]_{i=1}^n\right) \leqslant 3^{r-1}(E(Q_n)^r + E(R_n)^r + E(S_n)^r) \\ &\leqslant 3^{r-1}\left(\frac{1}{n}\sum_{i=1}^n E\|X_i\|_u^r I(X_i \notin \mathcal{K}) + E(R_n)^r + \varepsilon/4\right) = 3^{r-1}\varepsilon/2 + o(1) \text{ as } n \to \infty. \end{aligned}$$

By the arbitrariness of $\varepsilon > 0$, we obtain the conclusion (b). □

Remark 3.8 In Theorem 3.7, conclusion (a) extends [1, Theorem 3.1] to CC space and conclusion (b) extends [3, Theorem 1.2] to CC space.

As mentioned in Section 2, $(k(\mathcal{X}), d_H)$ is a separable and complete CC space. Let $E_{k(\mathcal{X})}X$ denote the expectation of an integrable random element X in $(k(\mathcal{X}), d_H)$. By applying Theorem 3.1 and Theorem 3.7, we obtain immediately the following corollaries:

Corollary 3.9 *Let $1 \leqslant p < 2$, $A \in K_{k(\mathcal{X})}(k(\mathcal{X}))$ and let $\{X, X_n, n \geqslant 1\}$ be a sequence of pairwise i.i.d. $k(\mathcal{X})$-valued random elements. Then $E_{k(\mathcal{X})}X = A$ and $Ed_H^p(X, A) < \infty$ if and only if*

$$Ed_H^p(X, K_{k(\mathcal{X})}X) < \infty$$

$$\text{and} \quad \sum_{n=1}^{\infty} n^{p-2} P\left(\max_{1\leqslant k\leqslant n} k d_H\left([k^{-1}, X_i]_{i=1}^k, A\right) > n\varepsilon\right) < \infty \text{ for all } \varepsilon > 0.$$

Corollary 3.10 *Let $\{X_n, n \geqslant 1\}$ be a sequence of pairwise independent $k(\mathcal{X})$-valued random elements.*
(a) *If $\{X_n, n \geqslant 1\}$ is Cesàro CUI and $\sum_{n=1}^{\infty} n^{-p} Ed_H^p(X_n, U) < \infty$ for some $p \in [1, 2]$, some $U \in k(\mathcal{X})$, then for all $\varepsilon > 0$*

$$\sum_{n=1}^{\infty} n^{-1} P\left(\max_{1\leqslant k\leqslant n} k\, d_H\left([k^{-1}, X_i]_{i=1}^k, [k^{-1}, E_{k(\mathcal{X})}X_i]_{i=1}^k\right) > n\varepsilon\right) < \infty.$$

In particular, $d_H([n^{-1}, X_i]_{i=1}^n, [n^{-1}, E_{k(\mathcal{X})}X_i]_{i=1}^n) \to 0$ a.s. as $n \to \infty$.
(b) *If $\{X_n, n \geqslant 1\}$ is Cesàro r-th CUI ($r \geqslant 1$), then*

$$Ed_H^r\left([n^{-1}, X_i]_{i=1}^n, [n^{-1}, E_{k(\mathcal{X})}X_i]_{i=1}^n\right) \to 0 \text{ as } n \to \infty.$$

Acknowledgements This work was completed while the authors were visiting the Vietnam Institute for Advanced Study in Mathematics (VIASM), and the authors would like to thank the VIASM for their support and hospitality.

Funding Information This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 101.03-2017.24.

References

1. Bai, P., Chen, P.Y., Sung, S.H.: On complete convergence and the strong law of large numbers for pairwise independent random variables. Acta Math. Hungar. **142**(2), 502–518 (2014)
2. Baum, L.E., Katz, M.: Convergence rates in the laws of large numbers. Trans. Am. Math. Soc. **120**, 108–123 (1965)
3. Chen, P.Y., Wang, D.C.: L^r, convergence for B-valued random elements. Acta Math. Sin. Engl. Ser. **28**(4), 857–868 (2012)
4. Etemadi, N.: An elementary proof of the strong law of large numbers. Z. Wahrsch. Verw. Gebiete **55**(1), 119–122 (1981)
5. Huan, N.V., Quang, N.V., Thuan, N.T.: Baum-Katz type theorems for coordinatewise negatively associated random vectors in Hilbert spaces. Acta Math. Hungar. **144**(1), 132–149 (2014)
6. Hsu, P.L., Robbins, H.: Complete convergence and the law of large numbers. Proc. Nat. Acad. Sci. USA **33**, 25–31 (1947)
7. Katz, M.L.: The probability in the tail of a distribution. Ann. Math. Statist. **34**, 312–318 (1963)
8. Kuczmaszewska, A., Lagodowski, Z.A.: Convergence rates in the SLLN for some classes of dependent random fields. J. Math. Anal. Appl. **380**(2), 571–584 (2011)
9. Quang, N.V., Thuan, N.T.: On the strong laws of large numbers for double arrays of random variables in convex combination spaces. Acta Math. Hungar. **134**(4), 543–564 (2012)

10. Shixin, G.: The Hàjek-Rènyi inequality for Banach space valued martingales and the p smoothness of Banach spaces. Statist. Probab. Lett. **32**(3), 245–248 (1997)
11. Terán, P.: Algebraic, metric and probabilistic properties of convex combinations based on the t-normed extension principle: the strong law of large numbers. Fuzzy Sets Syst. **223**, 1–25 (2013)
12. Terán, P., Molchanov, I.: The law of large numbers in a metric space with a convex combination operation. J. Theoret. Probab. **19**(4), 875–898 (2006)
13. Thuan, N.T.: Approach for a metric space with a convex combination operation and applications. J. Math. Anal. Appl. **435**(1), 440–460 (2016)
14. Thuan, N.T., Quang, N.V., Nguyen, P.T.: Complete convergence for arrays of rowwise independent of random variables and fuzzy random variables in convex combination spaces. Fuzzy Sets Syst. **250**, 52–68 (2014)
15. Wang, X.C., Bhaskara Rao, M.: Some results on the convergence of weighted sums of random elements in separable Banach spaces. Studia Math. **86**(2), 131–153 (1987)
16. Zhang, L.X., Wen, J.W.: Strong laws for sums of B-valued mixing random fields. Chinese Ann. Math. Chinese Series **20**, 205–216 (2001)