ADV MATH SCI JOURNAL
Advances in Mathematics: Scientific Journal **13** (2024), no.2, 145–180
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.13.2.3

AN APPROACH OF MEAN-SETS THEORY FOR NEGATIVELY CURVED CONVEX COMBINATION POLISH METRIC SPACES

Christophe Fotso[1], Duplex Elvis Houpa Danga, and Daniel Tieudjo

ABSTRACT. The choice of a convenient approach to be used is one important issue when attempting to develop methods or obtain results in the setting of Probability on general metric spaces. In this paper, we extend the mean-sets probabilistic approach formally introduced by Mosina on locally finite graphs, and hence (via Cayley graphs) on finitely generated groups, to the field of Negatively Curved Convex Combination Polish (NCCCP) metric spaces. We construct an appropriated Vertex-Weighted Metric (VWM) graph in the framework of this class of geometrical structures. We define a function called convexification function on the direct product of n copies of the vertex-set of this graph (for a given fixed integer $n \geq 2$), using the natural convexification operator of the metric space concerned. This function is then used to construct a weighted mean-set that generalizes the notion of convex combination (CC) mean in the sense of Terán and Molchanov, the mean-set concept according to Mosina and the ordinary notion of k-means ($k \geq 2$) of independent identically distributed (i.i.d.) random elements of the metric space. Two numerical examples are given for the cases when the metric space $X = [0, 1]$ and $X = \mathbf{R}^2$. Moreover, an analogue of the Strong Law of Large Numbers (SLLN), the consistency problem and the Chebyshev's inequality for NCCCP spaces are established.

[1]*corresponding author*
2020 *Mathematics Subject Classification.* Primary 05C12, 05C22, 60B05, 60D05, 60F15; Secondary 51K05.
Key words and phrases. NCCCP metric space, VWM graph, mean-set, convexification function.
Submitted: 17.05.2024; *Accepted:* 02.06.2024; *Published:* 05.06.2024.

1. Introduction

In the field of Probability Theory like in Statistics, real or vectorial random variables have been basically considered and several results have been established. Today, major papers about some concepts like expectation (average) and SLLN aim to generalize these to non linear classes of sets, devoid of vectorial structure, such as: graphs of finitely generated groups [12, 23, 25, 27], general graphs [2, 9, 18], trees [3], Polish metric spaces [1, 4, 5, 13, 14, 17, 21, 22, 28, 37], non-separable metric spaces [9], negatively curved metric spaces [13, 14, 31, 32], metric spaces endowed with a convex combination (CC) operation (also called convex combination metric spaces) [4, 5, 13, 34, 37], spaces of compact sets [9, 11] and regular topological spaces [36]. The SLLN in particular is in great demand for its many implications which, beyond Probability Theory and Statistics, are increasingly found in various fields like Engineering [24, 29, 30], Physical sciences [6] or Cryptography [16, 25–27], where randomness influences the dynamic behavior of so many phenomena.

In her 2009 PhD thesis, N. Mosina [25] introduced a new probabilistic approach called "Mean-sets Theory" that allows to define the concept of mean-set (in effect, the set of Fréchet means) of a graph/group random variable as the average-object (expectation) in the sense of Fréchet [17]. She also supplied an efficient algorithm for the computation of this mean-set, following Monte-Carlo's approximate calculation method for the field of real numbers. Accordingly, a version of the SLLN for a sequence of mutually independent and identically distributed (i.i.d.) graph/group valued random variables was established as well as an analogue of Chebyshev's inequality. These results make this mean-set a viable statistical tool for some algorithms involving groups, such as in group-based cryptography. Since then, the idea of Mosina has been used a few times, but seems to be relatively little explored compared to more well-known probabilistic approaches, such as random walks on groups.

Remind that giving a meaning to the notion of (mathematical) expectation or average of a sample of graphs/groups valued random variables is not an easy task. Things become more complex when some theoretical results obtained via the newly constructed mean-sets probabilistic approach are to be generalized to non vectorial spaces with no algebraic properties but having a compatible topology or compatible distance function like metric spaces. This problematic is still open for general metric spaces. Therefore, the main question that emerges wonders about a possible connection between the Mean-sets theory and general metric spaces. What changes with this theory if we ignore the algebraic properties of the groups and only consider their geometric structure as metric spaces? Can Mosina's results be extended to general metric spaces devoid of an algebraic structure and classical operations such as addition and product, but whose geometric conditions offer issues and tools for averaging? The openness given by one Mosina's endeavors coupled to our state of art synthesis, allow to tackle this question in this paper. Indeed, what impact Mosina's original probabilistic approach can have on definitional frameworks and properties of the mean (expectation) and SLLN concepts in metric spaces, whose already have a fairly rich literature on the subject? Having in mind that the Mean-sets theory is built on finitely generated ordinary groups, some fundamental principles used in geometric group theory have to be investigated. Notably, connections between algebraic, topological and geometric properties of the spaces on which these groups can act. This concerns NCCCP, i.e. negatively curved Polish metric spaces (by considering only its geometric aspect without recourse to algebraic structure which supports it) with the convex combination operation defined in the sense of Terán and Molchanov [34].

A motivating element when looking for an answer to this problematic is based on the fact that very often in such spaces (like separable Banach spaces), convex subsets appear as limits for the law of large numbers, since they are decomposable with respect to Minkowski addition [4, 5, 34]. Also, to each random variable with values in such a space, the negative curvature property assigns a unique barycenter or mean (expectation). Moreover, the concept of Fréchet mean used by Mosina gives way to generalize several notions of average in abstract spaces, while offering for some applications in many fields including cryptography [18, 25, 26]. Moreover, we assume that going from random variables taking values over countable sets, as in Mosina's work, to those defined over the metric graphs vertex-set of a NCCCP space does not matter at all.

This work is split into five sections. Section 2 is the state of art. Section 3 concerns some preliminaries. First comes a short presentation of the mean-sets probabilistic approach. After, some reminders on some crucial notions are recorded in order to build vertex-weighted metric graphs. In section 4, we present some useful results as tools that will help to build a theoretical mean-sets approach for NCCCP spaces. We also give a practical numerical example of the mean-set of a uniform distribution. Our main results are acheived in Section 5, including a new topological characterization of mean-sets, the SLLN and the Tchebychev inequality for NCCCP spaces.

2. State of art

A good remind of some existing definitions of the two probabilistic concepts namely mean-set (expectaton) and SLLN for several metric spaces has been taken up in [4, 5, 27]. In this section, we present some two additional one which are of great interest in our work.

2.1. **Some expectations (means) and SLLNs.**

2.1.1. *Expectation and SLLN according Herer.* Let (X, d_X) be a complete separable metric space with negative curvature and ξ a random X-variable point defined on $(\Omega, \mathcal{F}, \mathbb{P})$, equipped with Borel tribe $\mathcal{B}(X)$ and with distribution μ_ξ. Define the set $\pi = \{A_1, \dots, A_n\} \subset \mathcal{F}$ as a finite partition of Ω and put $x_i = \xi(A_i) \in X$. The random variable ξ takes the form of a μ_ξ simple step-function such that

$$\xi(\omega) = \sum_{i=1\dots n} x_i \cdot 1_{A_i}(\omega), \text{ for all } \omega \in \Omega.$$

Let $\prod(\xi)$ be the set of all non-empty finite partitions of Ω such that ξ is constant on A_i's compartments. For $\pi \in \prod(\xi)$, we define the set $\mathbb{E}_\pi\xi = \sum_{A\in\pi} \mathbb{P}(A)\xi(A)$. the Herer expectation [21] of a stepped random X-variable ξ, also called mathematical expectation of ξ, is the closed set

$$\mathbb{E}\xi = \overline{\bigcup_{\pi\in\prod(\xi)} \mathbb{E}_\pi\xi}.$$

With this concept, Herer establishes a version of strong law of large numbers for finitely compact NCCCP metric spaces (X, d_X) by using Hausdorff convergence on finite non-empty subsets of X, i.e. $\operatorname*{Lim}_n F_n = F$ if and only if $\lim_n d_{\text{Haus}}(F_n, F) = 0$.

Herer strong law of large numbers: Let $\{\xi_n\}_{n=1,\ldots,\infty}$ be a sequence of i.i.d. integrable random X-variables and write $F_n(\omega) = \sum_{i=1\ldots n} \frac{1}{n}\xi_i(\omega)$, for $\omega \in \Omega$ and $n = 1, 2, \ldots$. Then $\operatorname*{Lim}_n F_n(\omega) = \mathbb{E}\xi_1$ almost surely (a.s.) [22].

2.1.2. *Barycenter and convex combination mean according to De Fitte.* Refering to Herer's expectation, De Fitte [13, 14] extends the same concept to sequences of random independent and equidistributed points in NCCCP spaces. In his approach, he makes a difference between the mean (average) of a probability μ when the probabilistic space has atoms and when it has atoms as large as possible.

In the first case and under condition $\int_X d(a,x)d\mu(x) < \infty$, the barycenter of μ, denoted by $b(\mu)$, is equal to Herer expectation of a random point defined on $([0,1], \mathcal{B}([0,1]), \mathbb{P})$ with law μ, where $\mathbb{P}$ is the uniform probability on the Borel tribe $\mathcal{B}([0,1])$.

In the second case, the mean of μ, called convex combination mean of μ and denoted by $c(\mu)$ is the Herer expectation of ξ relative to $(\Omega, \mathcal{F}_\xi, \mathbb{P}_\xi)$, where $\mathcal{F}_\xi$ represents the tribe generated by ξ (i.e. the poorest in terms of open sets on which ξ is measurable) and $\mathbb{P}_\xi$ is the restriction of $\mathbb{P}$ to $\mathcal{F}_\xi$.

In general, the double inclusion $c(\mu_\xi) \subset \mathbb{E}\xi \subset b(\mu_\xi)$ holds (See [21], Remark 1.1.).

To this two definitions, we can add the version according to Terán and Molchanov. But we shall come back to it with more details in Subsection 3.2, when talking about the convex combination metric spaces.

The next section represents the Preliminaries for our work including some notations and background theory which will be used in Sections 4 and 5.

3. Preliminaries

We start with a short overview on the Mean-sets probability theory introduced by Mosina. After this, we define the CC operation and some general properties of CC spaces. We end with a brief description of the convergence of minimizing sequences as well as the notion of τ-convergence.

3.1. **Mean-sets Probability Theory by Mosina.** Given a group G and a probability distribution $\mathbb{P}$ induced by random G-variables ξ. We endow Cayley graphs of G with word metric and we define a weight-function on these graphs. Let S be a non-empty generating subset of the group G. Let $\Gamma = (V\Gamma, E\Gamma)$ be the Cayley graph of G, $(\Omega, \mathcal{F}, \mathbb{P})$ a probabilistic space.

3.1.1. *Some mean-sets probability tools.* A probability measure on G attached to a random G-variable ξ is a function defined on G by $\mu_\xi(g) = \mathbb{P}(\{\xi^{-1}\{g\} \in \Omega\})$, for all $g \in G$.

Weight-function: The weight-function, denoted $M_\xi : V\Gamma \to \overline{\mathbb{R}}_+ = [0, +\infty]$, is defined on $V\Gamma$ by $M_\xi(g) = \sum_{s\in G} d_G^2(g,s)\mu_\xi(s)$, for $g \in G$, where d_G is the Cayley distance (or word distance) on $C_G(S)$.

The domain of M_ξ is the set $\mathrm{dom}(M_\xi) = \{g \in V\Gamma \mid M_\xi(g) < \infty\}$. The weight-function M_ξ is totally defined if $\mathrm{dom}(M_\xi) = V\Gamma$.

Mean-set: In the case when M_ξ is totally defined, the expectation or mean-set of ξ is defined to be the set of objects $g \in V\Gamma$ that minimize the weight-function M_ξ, i.e.

$$\mathbb{E}\xi = \{g \in V\Gamma \mid M_\xi(g) \leq M_\xi(h), \text{for all}\, h \in V\Gamma\}.$$

Let $(\xi_i)_{i=1,\ldots,n}$ be a finite sequence of mutually i.i.d. random $V\Gamma$-variables on $(\Omega, \mathcal{F}, \mathbb{P})$.

Sampled weight-function: The probability distribution $\mu_n : V\Gamma \to [0,1]$ on G is defined by the relative frequency

$$\mu_n(g) = \mu_n(g,\omega) = n^{-1} \times |\{i \mid \xi_i(\omega) = g,\, 1 \leq i \leq n\}|.$$

The sampled weight-function $M_n : V\Gamma \to \overline{\mathbb{R}}_+$ is defined by:

$$M_n(g) = \sum_{s\in G} d_G^2(g,s)\mu_n(s),$$

where d_G is a metric on Γ.

Sampled mean-set: The sampled mean-set of $(\xi_i)_{i=1,\ldots,n}$ is the set $\mathbb{S}_n$ defined by: $\mathbb{S}_n = \mathbb{S}(\xi_1, \ldots, \xi_n) = \{g \in G/M_n(g) \leq M_n(h), \text{ for all}\, h \in G\}$.

Shift property: The function $\xi_g : \Omega \to G$ defined by $\xi_g(\omega) = g\xi(\omega)$, for $\omega \in G$, satisfies the property $\mathbb{E}\xi_g = g\mathbb{E}\xi$. This equality is similar to the linearity property of a classical mean in the domain of real variables.

3.1.2. *SLLN for graphs/groups.*

Theorem 3.1. ***(SLLN)*** *([25,27]) Let $\Gamma = C_G(S)$ be a locally-finite connected graph of $G = \langle S \rangle$, for a finite subset S, and $(\xi_i)_{i=1\ldots\infty}$ a sequence of mutually i.i.d. random $V\Gamma$-variables such that M_{ξ_1} is totally defined on $V\Gamma$ and $\mathbb{E}\xi_1$ is a singleton of $V\Gamma$. Then $\mathbb{S}(\xi_1, \ldots, \xi_n) \longrightarrow \mathbb{E}\xi_1$ almost surely when $n \to \infty$.*

3.1.3. *Chebyshev's inequality for graphs/groups.*

Theorem 3.2. *([25, 27]) Let $\Gamma = C_G(S)$ be a locally-finite connected graph of $G = \langle S \rangle$, for a finite subset S, and $(\xi_i)_{i=1\ldots\infty}$ a sequence of mutually i.i.d. random $V\Gamma$-variables such that the weight-function M_{ξ_1} is totally defined on $V\Gamma$. Then there exists a constant $C = C(\Gamma, \xi_1) > 0$ such that*

$$\mathbb{P}\left(\{\exists u \in V\Gamma \setminus \mathbb{E}\xi_1 \, \textit{s.t.} \; M_n(u) \leq M_n(v), v \in \mathbb{E}\xi_1\}\right) \leq Cn^{-1}.$$

This short presentation only partially reveals the deepness of the mathematical background attached to the Mean-sets probabilistic approach and its fields of applications, particularly in the world of Digital computing through group-based Cryptography. For more details, see e.g., [16, 25–27].

We now give a short reminders on Polish metric spaces endowed with a CC operation also called convex combination Polish spaces. One can also see in [11, 37].

3.2. **Convex combination Polish spaces.** Consider a Polish metric space (X, d_X) provided or not with a usual algebraic addition operation. We will focus on Terán and Molchanov [34] approach of the notion convex combination. The version of Herer can be found in [21, 22].

3.2.1. *Convex combination according to Terán and Molchanov.* A convex combination on X is an operation, denoted by $[\cdot, \cdot]$, which assigns to any finite sequence of elements $x_1, x_2, \ldots x_n \in X$ $(n \geq 2)$ and reals $\lambda_1, \ldots \lambda_n > 0$ such that $\sum_{i=1}^n \lambda_i = 1$, a special element of X, denoted by

$$[\lambda_i, x_i]_{i=1}^n = [\lambda_i, x_i]_{i \in \{1,\ldots,n\}} = [\lambda_1, x_1; \ldots ; \lambda_n, x_n],$$

which satisfies a number of algebraic, analytic and topological properties notably: commutativity, associativity, continuity, negative curvature, convexification. The reader can refer to [4, 5, 34, 37] and references therein for more details and proofs.

A point $z \in [\lambda_i, x_i]_{i=1}^n$ if z can take the form

$$z = \left[m_1 \left[\frac{\lambda_j}{m_1}, x_j \right]_{j \in I_1} ; m_2 \left[\frac{\lambda_j}{m_2}, x_j \right]_{j \in I_2} \right]$$

where I_1 and I_2 are a partition of the set $\{1, \ldots, n\}$ and $m_\ell = \sum_{i \in I_\ell} \lambda_i$ for $\ell \in 1, 2$.

The metric space (X, d_X) is called convex (respectively strictly convex) if the set $\lambda x + (1 - \lambda)y$ is non-empty (respectively a singleton), for all $x, y \in X$ and $\lambda \in [0, 1]$. In the case of strict convexity, the set $\lambda x + (1 - \lambda)y$ is simply identified with its single element.

A metric space endowed with a CC operation that satisfies the five properties listed above is called a convex combination (CC) space [34, 37]. And any negatively curved Polish metric space with a CC operation is just a NCCCP space.

Convexification operator $\mathcal{K}$: Let (X, d) be a convex combination Polish space, i.e. a Polish metric space endowed with a CC operation (in the sense of Terán and Molchanov). The convexification operator of X is a linear operator $\mathcal{K} : X \longrightarrow X$, which assigns to every $x \in X$, the limit of an iteration of the convex combination of the same element with uniform weights n^{-1} $(n \geq 2)$, i.e. for all $x \in X$,

$$\mathcal{K}x = \lim_{n \to +\infty} [n^{-1}, x]_{i=1}^n.$$

The image of x, i.e. $\mathcal{K}x$, is therefore called the convexifier of x. An element $x \in X$ admits a convex decomposition if and only if it can take the form $x = [\lambda_i, x]_{i=1}^n$, with $n \geq 2$, $\lambda_1, \ldots, \lambda_n > 0$ and $\sum_{i=1}^n \lambda_i = 1$. The convexifiable domain of X, denoted by $\mathcal{K}(X)$, is the set of elements of X admitting convex decompositions. Therefore, $\mathcal{K}(X) \subseteq X$. The space X is called convexifiable if $\mathcal{K}(X) = X$. In this case, the attached CC operation $[\cdot, \cdot]$ is called unbiased.

The operator $\mathcal{K}$ is *linear* in X (i.e. $\mathcal{K}[\lambda_i, x_i]_{i=1}^n = [\lambda_i, \mathcal{K}x_i]_{i=1}^n$), *idempotent* in X (i.e. $\mathcal{K}^2 = \mathcal{K}$) and *non-expansive* for the metric d_X (i.e. $d_X(\mathcal{K}x, \mathcal{K}y) \leq d_X(x, y)$, for all $x, y \in X$). For the proofs (see, e.g., [4, 5, 34, 37]).

Thuan [37] used the approach of Terán and Molchanov to show that, for a complete metric space X, there exists an isometric embedding of $\mathcal{K}(X)$ in a Banach space and this embedding preserves CC structure (see Theorem 3.3).

It is proved that in CC spaces, one can construct a single point-valued expectation called convex combination expectation which satisfies the SLLN (see [34]).

A situation that is quite different from the case in the general setting of metric spaces, where the expectation is frequently defined as a set which may be of multiple point-valued.

CC expectation and SLLN according to Terán and Molchanov: Let ξ be a measurable X-valued function defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. There exists a partition $\{\Omega_1, \dots, \Omega_m\}$ of non-empty measurables of Ω, such that ξ takes a constant value x_j on each subset Ω_j, for $j = 1, \dots, m$.

The CC mean is defined to be the set

$$\mathbb{E}\xi = [\mathbb{P}(\Omega_j), \mathcal{K}x_j]_{j=1}^{m}.$$

Theorem 3.3. *([34]) Let $\xi \in L^1_X$ and let $(\xi_1, \xi_2, \dots, \xi_n)$ be a sequence of random elements with the same distribution as ξ. Then the convergence $[n^{-1}, \xi_i]_{i=1}^{n} \xrightarrow{a.s.} \mathbb{E}\xi$ when $n \to \infty$, is feasible with a probability equals to 1.*

Proof. See [34]. □

Remark 3.1. *In a Polish metric space X equipped with a CC operation $[\cdot,\cdot]$, the invariant elements of the convexification operator $\mathcal{K}$ of X appear as limits for the SLLN [34, 37]. The idea of defining the mean-set of a random distribution using convex combination of finite set of k elements has led to approximating this distribution by a sequence of k points ($k \geq 2$ a fixed integer). There arise the problem of convergence of minimizing sequences of k-means.*

The problem of convergence of minimizing sequences of k-means involved the notion of set convergence, with two well adapted tools, actually Hausdorff convergence and τ-convergence. The Hausdorff convergence often refers to the notion of strong convergence, while the τ-convergence, which is not always comparable with metric, but under certain conditions on the structure of space, generalizes the convergence in the sense of Hausdorff. If the case, one talks about τ-Hausdorff topology [10, 24].

3.3. **Convergence of minimizing sequences and the notion of paths.**

3.3.1. *τ-convergence.* A sequence $H_{n,\ell}$ of n-elements sets of converges (weakly) to $H_{n,1}$ when $\ell \to \infty$ if there is a subsequence of points, denoted by $\{h^{\ell}_{i_1}, \dots, h^{\ell}_{i_n}\}$, extracted from $H_{n,\ell}$ for $\ell \geq 1$, such that $h^{\ell}_{i_j} \longrightarrow h^{1}_{i_j}$ (weak convergence) when $\ell \to \infty$, for $j = 1, \dots, n$.

Also, a sequence $(x_n) \subset X$ converges weakly to $x \in X$ if $f(x_n) \to f(x)$, for all $f \in X^\star$, where $X^\star$ represents the dual space of X (see [35]).

If (X, d_X, τ) is a metric space endowed with a τ-Hausdorff topology, then the map $d_X(\cdot, y) : (E, \tau) \longrightarrow \mathbb{R}$ is sequentially weakly semi-continuous. If in addition, $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous and strictly monotone function such that $\phi(0) = 0$, then $\phi \circ d_X$ is also sequentially weakly semi-continuous. And more if $x_n \xrightarrow{\tau} x$ then $\liminf \phi(d_X(x_n, y)) \geq \phi(d_X(x, y))$, for all $y \in X$. Similarly, if $x_n \xrightarrow{\tau} x$ in X and $\|x_n\| \longrightarrow \|x\|$ then $x_n \longrightarrow x$.

For normed vector spaces $(X, \|\cdot\|)$, the notion of τ-convergence coincides with the weak convergence topology. For more details about these concepts, see e.g., [24, 29, 30, 35].

Now, talking about sets convergence, let $\mathcal{E}$ be the set of non-empty finite subsets of (X, τ). Hausdorff metric on $\mathcal{E}$ is denoted by d_{Haus} and defined by

$$d_{\text{Haus}}(A, B) = \max \left\{ \sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A) \right\}.$$

If $\mathcal{U} \subset \mathcal{E}$, then the convergence in the Hausdorff sense of A_n's to a set $A \in \mathcal{U}$ is traduced by

$$\left(A_n \overset{d_{\text{Haus}}}{\longrightarrow} A \in \mathcal{U} \right) \iff \left(\inf_{B \in \mathcal{U}} d_{\text{Haus}}(A_n, B) \longrightarrow 0 \right).$$

That is, $A_n \overset{d_{\text{Haus}}}{\longrightarrow} A \in \mathcal{U}$ if every subsequence of $\{A_n\}$ admits a subsubsequence $\{A_{n_r}\}$ which converges to the element A of $\mathcal{U}$. The converse of this statment is not always true.

3.3.2. *Quantization and minimizing sequences of k-means.* Consider $k = n$ and let $H_n = \{h_1, \ldots, h_n\}$ represents the general term of the consecutive observations sequence necessary for the computation of the n-mean of a random variable ξ with distribution μ. The convergence of sequences $(H_{n,\ell})_{\ell \geq 1}$ of finite sets $H_{n,\ell} = \{h_1^\ell, \ldots, h_n^\ell\}$ of n elements, that minimize a weight-function of multi-vertex arguments is studied in [24, 30]. To acheive this, a measure of probability $\mathbb{P}$ with support T and a functional $\theta(\cdot, \mathbb{P}) : \mathcal{E} \longrightarrow \overline{\mathbb{R}}_+$, called "loss function" are considered. And the function $\theta(A, \mathbb{P}) = \int_X \phi \circ d(x, A) \mathbb{P}(dx)$, is defined on the set $\mathcal{E}$ of finite parts $A \subset X$, where $d(x, A) = \min_{a \in A} d(x, a)$ is the distance from point x to the set A and ϕ is a suitable function (i.e. a continuous and strictly increasing function such that $\phi(0) = 0$).

The finiteness assumption for the function $\theta(\cdot, \mathbb{P})$ should be an analogous condition to the totally defined property required for the Mosina's weight-function. In fact, this property is simply a combination of Lember's (P1) and (P2) properties [24]. Pollard [29] studied strong consistency of "k-means" using as Lember, the same functionals $\theta(\cdot, \mathbb{P})$ and $\theta(\cdot, \mathbb{P}_n)$. As a result, he found that for all $A \in \mathcal{E}$, the value $\theta(A, \mathbb{P}) = \int_X \min\limits_{a \in A} \|x - a\|^2 \mathbb{P}(dx)$ is the limit under strong convergence of the sequence with general term $\theta(A, \mathbb{P}_n)$, when $n \to +\infty$.

The next subsection is a short reminder on the notions of paths and homotopy of paths.

3.3.3. *Paths and homotopy of paths.* Let (X, d_X) be a pseudometric space, i.e. a set X equipped with a numerical application d_X which satisfies the symmetric, triangular inequality properties and the condition $d_X(x, x) = 0$ (meaning that d_X does not necessarily respect the axiom of coÃŕncidence). A path in X is a continuous mapping $\gamma : [0, 1] \longrightarrow X$ whose origin is $\gamma(0)$ and $\gamma(1)$ the terminus. If $\gamma(0) = x$ and $\gamma(1) = y$, then we say that γ is a path joining x to y. If for all $x, y \in X$ there exists such a path γ such that $\gamma(0) = x$ and $\gamma(1) = y$, then the space (X, d_X) is said to be connected by arcs.

We denote by $C_{x,y}([0, 1], X)$ the set of all paths joining x to y in X. For all $x \in X$, the constant path γ_x with origin and terminus x is defined, for all $t \in [0, 1]$, by $\gamma_x(t) = x$. Any two paths $\gamma, \gamma' \in C_{x,y}([0, 1], X)$ are said to be homotopic if there exists a continuous map $F : [0, 1] \times [0, 1] \longrightarrow X$, which satisfies both of the following properties:

(i) $F(t, 0) = \gamma(t)$ and $F(t, 1) = \gamma'(t)$, for all $t \in [0, 1]$.
(ii) $F(0, s) = x$ and $F(1, s) = y$, for every $s \in [0, 1]$.

Then the map F is a homotopy of paths joining x to y.

On the set $C_{x,y}([0, 1], X)$, the relation denoted by $\mathcal{R}$ and defined by $\gamma \mathcal{R} \gamma'$ if and only γ and γ' are homotopic, is an equivalence relation. Let $\Pi_{x,y}(X)$ denotes all of its homotopy classes. Write $\Pi(X) = \bigcup\limits_{x,y \in X} \Pi_{x,y}(X)$ and $[\gamma]$ an element of $\Pi_{x,y}(X)$. On $\Pi(X)$, define an internal composition law, denoted by "$\cdot$" and defined by $[\gamma] \cdot [\gamma'] = [\gamma\gamma']$, where $\gamma\gamma'$ is the composed path of γ and γ' when it makes sense, i.e. when $\gamma(1) = \gamma'(0)$.

4. Notation and Useful Results

We start by the construction of vertex-weighted metric graphs for NCCCP metric spaces which will played the same role as Cayley graphs for finitely generated groups [25]. After, we define the concept of mean-sets expectation for this class of metric spaces supplimented with a computational example.

4.1. **Conctruction of metric graphs for NCCCP spaces.**

4.1.1. *A compatible word metric on NCCCP spaces.* Let (X, d_X) be a NCCCP (pseudo) metric space and $c > 0$ a real. For any $x, y \in X$ and $n \geq 0$, a c-path of n steps from x to y in X is a sequence $x = h_0, h_1, \ldots, h_n = y$ of points in X such that $0 < d_X(h_{i-1}, h_i) \leq c$ for $i = 1, \ldots, n$. We also talk about c-path of n steps from x to y and with length-step at most or equal to c in X. Let E be the subset of $X \times X$ containing the diagonal $\mathrm{diag}(X)$.

Let R_E be the equivalence relation in X attached to E and defined by: $(x, y) \in R_E$ if there exists a finite sequence $(h_0 = x, h_1, \ldots, h_n = y)$ of points in X such that $(h_{i-1}, h_i) \in E$, for $i = 1, \ldots, n$. The space X is said to be an E-chained if the relation R_E is trivial, i.e. $(x, y) \in R_E \Longleftrightarrow (x = y)$.

To define a generalized metric on X that is very closed to the ordinary word-metric used by mosina on Cayley-graphs, we started with the map $\delta_E : X \times X \longrightarrow \mathbb{N}$ by

$$\delta_E(x, y) = \inf \left\{ n \geq 0 \mid \begin{array}{c} \exists (h_0 = x, h_1, \ldots, h_n = y) \in X^{n+1} \text{with} \\ (h_{i-1}, h_i) \in E \text{ for } i = 1, \ldots, n \end{array} \right\}.$$

But this metric δ_E doesn't take into account the metric d_X of the original space X. A situation that doesn't really satisfy enough, since the topology of a metric graph should depend fundamentally on that of the attached metric space. Therefore, we need more, i.e. finding a suitable compatible metric of the graph that meets this requirement.

Under the same assumptions on set E and relation R_E as fixed previously, we define a compatible word metric on a NCCCP space (X, d_X) as follows.

Definition 4.1. *([12]) A compatible word metric on the (X, d_X) is the function, denoted by $\delta_{E,d_X} : X \times X \longrightarrow \mathbb{R}_+$ and defined by:*

$$\delta_{E,d_X}(x,y) = \inf \left\{ \sum_{i=1}^{n} d_X(h_{i-1}, h_i) \mid \begin{array}{c} \exists (h_0 = x, h_1, \ldots, h_n = y) \in X^{n+1} \textit{ with} \\ (h_{i-1}, h_i) \in E \textit{ for } i = 1, \ldots n \end{array} \right\}.$$

Observe that, for all $x, y \in X$, the inequality $d_X(x,y) \leq \delta_{E,d_X}(x,y)$ holds. And, infinimas of δ_{E,d_X} can only be achieved on c-good E-paths having their elementary origins and ends in the subset E.

A path $(h_0 = x, h_1, \ldots, h_n = y)$ is said to be c-good [12] if $n \leq 1$ or, as soon as $n \geq 2$, $d_X(h_{i-1}, h_i) + d_X(h_i, h_{i+1}) > c$, for all $i = 1, \ldots, n-1$. Hence, if $x, y \in X$ with $\delta_{E,d_X}(x,y) < c$ or $d_X(x,y) < c$ then $\delta_{E,d_X}(x,y) = d_X(x,y)$.

Remark 4.1. *Coming up to geometric group theory as we are looking up to generalize Mean-sets theory to general metric spaces, notice that if $X = G = \langle S \rangle$ is a group with S a symmetric generator set containing the identity element 1_G and if $E = \{(g,h) \in G \times G \mid g^{-1}h \in S\}$, then δ_{E,d_X} is just the weighted word metric which extends the usual word distance defined in Cayley graphs of G.*

The function δ_{E,d_X} is a word-weighted compatible metric on X which thereby shares some interesting topological properties with d_X, thanks to the notion of control.

The notion of (c, C)-control. The notion of control here helps to insure the conservation of topological structure of (X, δ_E) while embedding in (X, δ_{E,d_X}).

Let (X, d_X) be an E-chained NCCCP space with $E \subset X \times X$ symmetric containing the diagonal diag(X) of X. Let $c,\, C \in \mathbb{R}_+^\star$. The triple (X, d_X, E) is said to be (c, C)-controlled E-chained NCCCP space if for all $x \in X$,

$$\overline{B}_X(x,c) \subset \{y \in X \mid (x,y) \in E\} \subset \overline{B}_X(x,C)$$

holds, where $\overline{B}_X(x,r) = \{y \in X \mid d_X(x,y) \leq r\}$ is the closed ball in X with center x and radius $r > 0$.

Some geometric properties under (c, C)-control assumptions of the metric space X are recorded in the next Lemma.

Lemma 4.1. *([12]) Let (X, d_X) be an E-chained NCCCP space with $E \subset X \times X$ symmetric containing the diagonal diag(X). Let $c, C \in \mathbb{R}_+^\star$ such that (X, d_X, E) is (c, C)-controlled. Then*

(i) *If* $x, y \in X$ *are such that* $\delta_{E,d_X}(x,y) < c$ *or* $d_X(x,y) < c$ *then* $\delta_{E,d_X}(x,y) = d_X(x,y)$.

(ii) *For all* $x, y \in X$, $\frac{c}{2}(\delta_E(x,y) - 1) \leq \delta_{E,d_X}(x,y) \leq C\delta_E(x,y)$ *hold.*

(iii) *The functions* δ_E *and* δ_{E,d_X} *are quasi-isometric pseudometrics on* X.

(iv) *When* X *is a topological space with* d_X *continuous then* δ_{E,d_X} *is also continuous.*

Proof. See [12] (Lemma 4.B.7) for more details. □

In Lemma 4.1 above, property (ii) implies that the map $(X, \delta_E) \longrightarrow (X, \delta_{E,d_X})$ is a Lipschitz self embedding of X. Therefore, there exist two real constants $c_1, c_2 > 0$ such that for all $x, y \in X$, we have $\frac{1}{c_2}\delta_E(x,y) \leq \delta_{E,d_X}(x,y) \leq c_1\delta_E(x,y)$. Indeed, just consider $c_1 = C$ and $c_2 = \frac{2}{c}$.

4.1.2. *Metric graph equipped with a weighted word metric.* Let (X, d_X) be an E-chained NCCCP space and $(X_{\text{Haus}}, d_{\text{Haus}})$ the greatest Hausdorff quotient of X, endowed with its natural metric d_{Haus}, where $X_{\text{Haus}} = X/\sim$ is the quotient of X by the relation

$$x \sim y \quad \text{iff} \quad \delta_{E,d_X}(x,y) = 0.$$

The metric graph $\mathit{\Gamma}_X = (V\mathit{\Gamma}_X,\ E\mathit{\Gamma}_X)$ attached to X is the chained (or connected) graph having $V\mathit{\Gamma}_X = X_{\text{Haus}}$ as the set of vertices and $E\mathit{\Gamma}_X = E_{\text{Haus}}$ as the set of edges, consisting of all the connections between pairs $(\widetilde{x}, \widetilde{y}) \in V \times V$, with $0 < d_{\text{Haus}}(\widetilde{x}, \widetilde{y}) \leq c$ for a fixed $c > 0$.

Consider the diagram

$$(X, d_X) \underset{\varphi}{\longrightarrow} (X, \delta_{E,d_X}) \underset{\psi}{\longrightarrow} (X_{\text{Haus}}, d_{\text{Haus}})$$

in which the embedding φ and the isomorphism ψ are both lipschitz. Therefore, $\psi \circ \varphi$ is a lipschitz embedding which preserves the topological structure of X. From Lemma 4.1 and according to the fact that (X, d_X, E) is (c, C)-controlled, we have

$$d_{\text{Haus}}(\widetilde{x}, \widetilde{y}) = \delta_{E,d_X}(x,y) = d_X(x,y).$$

Hence, the metrics d_{Haus}, δ_{E,d_X} and d_X are equivalent.

Remark 4.2. *Under the same assumptions as in Remark 4.1, i.e. If we set* $X = G = \langle S \rangle$ *a group, with* S *a symmetric subset of* G *containing the identity element* 1_G, $N = \{g \in G : \quad g \sim 1_G\}$ *the normal compact subgroup of* G *and take* $E = \{(g,h) \in$

$G \times G \mid g^{-1}h \in S\}$, then the E-chained metric graph $\mathit{\Gamma}_X = (X_{Haus}, E_{Haus})$ is attached to X and weightable on vertices.

We now define a weight-function on the vertex-set of this metric graph.

4.2. **A weight-function on the newly constructed metric graph.** Let denote by $\mathcal{P}_{(a,x,c,E)}(V\mathit{\Gamma}_X)$, or by $\mathcal{P}_{(a,x,c,E)}(\mathit{\Gamma}_X)$ to make simple, the set of elementary ends of a homotopy class of c-good E-paths in $\mathit{\Gamma}_X$, joining a given fixed point a to x in X. Observe that $\mathcal{P}_{(a,x,c,E)}(\mathit{\Gamma}_X) \subset \Pi_{a,x}(X)$, the set of homotopy classes of paths joining point a to x in X. Let $\bigcup\limits_{a,x \in X} \Pi_{a,x}(X) = \Pi(X)$, the set of homotopy classes of X.

For an arbitrarily fixed point $a \in X$, the results that we present in the following generalize the analogous results for separable Banach spaces, by taking $a = o$ the origin of the metric space and replacing the distance $d_X(a,x)$ by the standard norm $\|x-a\| = \|x-o\| = \|x\|$, for all $x \in X$.

From now on, one can take point a to be the origin o of the space X and write $\mathcal{P}_{(x,c,E)}(\mathit{\Gamma}_X)$ for an easier perception in separable Banach spaces.

4.2.1. *Minimizing the mean discrepancy function.* Let $H_n \in \mathcal{P}_{(a,x,c,E)}(\mathit{\Gamma}_X)$ be a subset of X with n elements defined by:

$$H_n = \left\{ h_1, \ldots, h_n \in X \left| \begin{array}{c} \exists (h_0 = a, h_1, \ldots, h_n = x) \in X^{n+1} \text{ with } \ (h_{i-1}, h_i) \in E, \\ \delta_{E,d_X}(h_{i-1}, h_i) < c, \text{ for } \quad i = 1, \ldots, n \end{array} \right. \right\}.$$

This corresponds to the set of the n elementary ends a finite path of n steps joining a given fixed point a to x in X. It therefore depends on the pair $(a,x) \in V\mathit{\Gamma}_X \times V\mathit{\Gamma}_X$. Such a subset of X, i.e. H_n, is always non-empty and finite, because X is an arc-connected space. It is also a bounded closed set and therefore a compact of X, as (X, d_X, E) is (c, C)-controlled and space X is finitely compact.

Now, let

$$\mathcal{P}_{(c,E)}(\mathit{\Gamma}_X) = \bigcup_{a,x \in X} \mathcal{P}_{(a,x,c,E)}(\mathit{\Gamma}_X),$$

be the set of elementary ends of a class of homotopies of c-good E-paths joining any two distinct points in $V\mathit{\Gamma}_X$. We have $\mathcal{P}_{(c,E)}(\mathit{\Gamma}_X) \subset \Pi(X)$.

For every element $H_n \in \mathcal{P}_{(c,E)}(\mathit{\Gamma}_X)$, define a finite partition $\pi_{H_n} = \{\Omega_i\}_{i=1,\ldots n}$ of Borel-measurables subsets of X attached to H_n by

$$\Omega_1 = \left\{ z \in X \mid \delta_{E,d_X}(z, h_1) \leq \min_{1 \leq j \leq n} \delta_{E,d_X}(z, h_j) \right\}$$

and for $i = 2, \ldots, n$,

$$\Omega_i = \left\{ z \in X \mid \delta_{E,d_X}(z, h_i) \leq \min_{1 \leq j \leq n} \delta_{E,d_X}(z, h_j) \right\} \setminus (\Omega_1 \cup \ldots \cup \Omega_{i-1}).$$

Working in separable Banach spaces, Cuesta and Matrán [10] considered, for a given such subset H_n and a random $\mathit{\Gamma}_X$-variable ξ taking a finite number of values $\{h_1, \ldots, h_n\} = H_n \subset X$, the simple function $\eta_{H_n} : X \longrightarrow X$ which, to any element $x = \xi(\omega)$ of X assigns $\eta_{H_n}(x) = \sum_{i=1}^{n} h_i 1_{\Omega_i}(x)$, where 1_{Ω_i} is the usual indicator function of Ω_i. And they looked to minimize the mean discrepancy between x and $\eta_{H_n}(x)$. A process that is sometimes called quantization [30].

We emphasize that our strategy is quite similar to that of Cuesta and Matrán [10]. But, our reasoning relies on tools adapted to the context of NCCCP spaces equipped with a CC operation.

The astute the reader will notice is that we replace the simple function η_{H_n} used by Cuesta and Matrán, by considering the function of the same name, defined for any element $x = \xi(\omega)$ of X by

$$\eta_{H_n}(x) = [p_i, \, \mathcal{K}h_i]_{i=1}^{n},$$

where $\mathcal{K}$ is the convexification operator of X, $\mathcal{K}h_i = \xi(\Omega_i)$ and $p_i = \frac{1}{n} \times 1_{\Omega_i}(x)$.

Note that, for given two partitions π_{H_n} and π'_{H_n} of X such that π_{H_n} is finer than π'_{H_n}, i.e. each part of X belonging to π'_{H_n} is included in a part of X belonging to π_{H_n}, we have $\xi(\omega) = \xi(\omega')$ for all $\Omega_i \in \pi_{H_n}$ and all $\omega, \omega' \in \Omega_i$. The function η_{H_n} is therefore well defined on X and $\eta_{H_n}(\Omega_i) = \mathcal{K}h_i$ on Ω_i's.

For any subset $H_n \in \mathcal{P}_{(c,E)}(\mathit{\Gamma}_X)$, the attached partition π_{H_n} defined above allows to minimize the mean discrepancy between x and $\eta_{H_n}(x)$, represented by the real value $\phi \circ \delta_{E,d_X}(x, \eta_{H_n}(x))$, where $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ is a function of the general form $\phi(x) = x^p$ with $p \in \{1, 2\}$. It is worth noticing that the discrepancy between x and $\eta_{H_n}(x)$ can arise in a non-convexifiable metric space, i.e. when the convexification operator $\mathcal{K}$ is biaised (and traduced by $\mathcal{K}X \subset X$ et $\mathcal{K}X \neq X$).

We now define a weight-function on the set of vertices of our newly constructed metric graph.

4.2.2. *Weight-function on $V\mathit{\Gamma}_X$.* Let $\mathit{\Gamma}_X = (X_{\text{Haus}}, E_{\text{Haus}})$ be an E-chained metric graph attached to X. For all $H_n \in \mathcal{P}_{(c,E)}(\mathit{\Gamma}_X)$ and any probability measure μ on X, we define a $\mathbb{R}_+$-valued weight-function $\theta_\mathit{\Gamma}$, called weight-function attached to $\mathit{\Gamma}_X$, on $\mathcal{P}_{(c,E)}(\Gamma_X)$ by:

$$\theta_\Gamma(H_n) = \int_{x\in X} \phi \circ \delta_{E,d_X}(x, \eta_{H_n}(x))d\mu(x).$$

This weight-function is just a special case of loss-function where set A has been replaced by $\eta_{H_n}(x)$.

The condition $\int_{x\in X} \phi \circ \delta_{E,d_X}(x, \eta_{H_n}(x))d\mu(x) < \infty$, for all $H_n \in \mathcal{P}_{(c,E)}(\Gamma_X)$ traduces the finiteness property of this weight-function which is also guaranteed by the following more general condition, for all constant $K \geq 0$,

$$\int_{x\in X} \phi \left(\delta_{E,d_X}(x, \eta_{H_n}(x)) + K\right) d\mu(x) < \infty.$$

This more general condition is always verified for power functions $\phi(x) = x^p$ with $p \in \{1, 2\}$.

The next proposition shows that the weight-function θ_Γ is totally defined in the sense of Mosina.

Proposition 4.1. *Consider the weight-function θ_Γ defined above. Then $\theta_\Gamma(H_n) < \infty$ holds, for all $H_n \in \mathcal{P}_{(c,E)}(\Gamma_X)$.*

Proof. Take $H_n \in \mathcal{P}_{(c,E)}(\Gamma_X)$. For every $x \in X$, $\mathcal{K}x = x$ holds since all singletons are convex in X. We can therefore write $x = [p_i, \mathcal{K}x]_{i=1}^n$. But, according to the negative curvature property of space, the non expansiveness property of the metric and the compatibility property of the convexification operator $\mathcal{K}$ of X, we have the following:

$$\begin{aligned}
\delta_{E,d_X}(x, \eta_{H_n}(x)) &= \delta_{E,d_X}([p_i, \mathcal{K}x]_{i=1}^n, [p_i, \mathcal{K}h_i]_{i=1}^n) \\
&\leq \sum_{i=1\dots n} p_i \delta_{E,d_X}(\mathcal{K}x, \mathcal{K}h_i) \\
&\leq \sum_{i=1\dots n} p_i \delta_{E,d_X}(x, h_i) \\
&\leq n^{-1} \times (\tfrac{n(n+1)}{2})c = \tfrac{(n+1)c}{2}.
\end{aligned}$$

The inequalities $\phi \circ \delta_{E,d_X}(x, \eta_{H_n}(x)) \leq \phi(\frac{(n+1)c}{2}) < \infty$ hold, for all $x \in X$, since ϕ is continuous and strictly increasing. Hence, $\phi \circ \delta_{E,d_X}$ is bounded and

$$\theta_\Gamma(H_n) \leq \phi(\frac{(n+1)c}{2})\mu(X) \leq \phi(\frac{(n+1)c}{2}) < \infty$$

as needed. □

Moreover, the mean-set $H_{n,0}$ of order $k = n$ always exists as the function $\phi \circ \delta_{E,d_X}$ is bounded.

Remark 4.3.

(i) *The graph $\Gamma_X = (X_{\mathit{Haus}}, E_{\mathit{Haus}}, \theta_\Gamma)$ as constructed above is a VWM graph of X.*

(ii) *The weight-function θ_Γ depends on the set H_n and the probability measure μ. So we can either write $\theta_\Gamma(H_n, \mu)$.*

With this notation, if $\mathcal{E} = \{H \subset X \mid |H| \leq \infty\}$ is the set of finite subsets of X, and for all $n \geq 1$, $\mathcal{E}_n = \{H \subset X \,:\, |H| \leq n\}$ the set of parts of X having at most n elements, then we define the real

$$\theta_{\Gamma,n}(\mu) = \inf_{H \in \mathcal{E}_n} \theta_\Gamma(H, \mu),$$

called n-variance or variance of order $k = n$ of μ.

(iii) *In the case when $X = \mathbb{R}$ and $\phi(x) = x^2$, the real $\theta_{\Gamma,1}(\mu)$ is just the classic variance of μ. And since $\phi \circ \delta_{E,d_X}$ is bounded, we have $\phi(\infty) < \infty$.*

This condition allows to have the decreaseness property of the finite sequence of weight-functions $\theta_{\Gamma,\ell}(\mu)$, for $\ell = 1, \ldots, n$, i.e.

$$\theta_{\Gamma,n}(\mu) < \theta_{\Gamma,n-1}(\mu) < \ldots < \theta_{\Gamma,1}(\mu) < \phi(\infty).$$

A situation that remains valid as soon as $\theta_{\Gamma,n-1}(\mu) > 0$ (see, e.g., [24]). If $\mathcal{P}_{(c,E)}(\Gamma_X) \subset \mathcal{E}_n$ then the previous inequalities obtained in $\mathcal{E}_n$ remain valid in $\mathcal{P}_{(c,E)}(\Gamma_X)$.

As our strategy is to define and use meaningful metric tools compatible with CC operation properties. For a random X-variable taking a finite number of values, we built a framework that allows to compute the moving average of order n. After, we look at how to convexify this average-object, i.e. to transform it into a convex objet, by using an appropriate function we called convexification function.

4.3. **Convexification function.** Let (X, d_X) be a NCCCP space. Write X^n, the direct product of n copies of X and Prob_n the set of probability measures on $\{1, 2, \ldots, n\}$ which counts exactly n elements, equipped with the norm $\|\mu - \nu\| = \sum\limits_{i=1,\ldots n} |\mu(i) - \nu(i)|$ of ℓ_1 space.

Definition 4.2. *The convexification function of X with n X-valued entries attached to a measure $\mu \in Prob_n$ is defined to be the continuous function $\Psi_\mu : X^n \longrightarrow X$ by*

$$\Psi_\mu(x_1, \ldots, x_n) = [\mu(i), \mathcal{K}x_i]_{i=1}^n ,$$

where $\mu \in Prob_n$ and $\mu(i) = p_i$ for $i = 1, \ldots n$.

This function is well defined since the space X is complete and negatively curved. The single output element in X represents the convex combination of at most n different values âĂŃâĂŃreceived in input.

One question to be ask and which the answer depends on the properties of Ψ_μ is: " do this single output element (i.e. the convex mean) share the same properties with the input elements ? "

Some general properties of the convexification function Ψ_μ. Notice that the function Ψ_μ acts as a projection of the CC operation $[0,1]^n \times X^n \longrightarrow X$ on the cartesian space X^n. For its n arguments, Ψ_μ shares the same topological, metric and algebraic properties on $[0,1]^n \times X^n$ with the convex combination on X, in the same way as if X was a convex subset of a bounded ball in a normed vector space [6].

Proposition 4.2. *If there exists a function Ψ_μ attached to a metric space X such that*

$$\Psi_\mu(x_1, \cdots, x_n) = \mathcal{K}[\mu(i), x_i]_{i=1}^n,$$

then the space X is convexifiable.

Proof. Let suppose there exists a function Ψ_μ attached to the metric space X such that $\Psi_\mu(x_1, \cdots, x_n) = \mathcal{K}[\mu(i), x_i]_{i=1}^n$. For all $x \in X$, the invariance $\mathcal{K}x = x$ holds since every singleton X is convex. We therefore have

$$x = [p_i,\, \mathcal{K}x]_{i=1}^n = \mathcal{K}[p_i,\, x]_{i=1}^n = \Psi_\mu(x, \cdots, x) \in \mathcal{K}(X).$$

Therefore $X \subseteq \mathcal{K}(X)$. Hence $X = \mathcal{K}(X)$ and X is convexifiable. □

One therefore says that the metric space X equipped with a convexification function Ψ_μ has a convex-like structure. Thuan [37] used this notion of convex-like structure to show that any complete (and if separable) metric space endowed with a CC operation can be embeded in a Banach (and also separable) space. More informations on the notion of convex-like metric structure are available in [6, 7].

Proposition 4.3. *Let X be a convexifiable space. For μ_0 a uniform distribution defined by $\mu_0(i) = \frac{1}{n}$, the mapping Ψ_{μ_0} is a surjection.*

Proof. Indeed, for all $x \in X$, we have $x = \left[\frac{1}{n}, x\right]_{i=1}^{n} = \Psi_{\mu_0}\left(\{x\}_{i=1}^{n}\right)$. □

Remark 4.4. *The convexification operator $\mathcal{K}$ and the convexification function Ψ of X are two concepts that differ at the levels of their arguments and their values.*

The next proposition shows a link between these two concepts.

Proposition 4.4. *Let (X, d_X) be a NCCCP space and $\mathcal{K}$ the convexification operator of X. If Ψ_{μ_0} is the convexification function of X wth n entries attached to μ_0 (i.e. the uniform distribution over the set $\{1, \ldots, n\}$), then for all $x \in X$, $\mathcal{K}x = \lim_{n\to\infty}(\Psi_{\mu_0}\{(x)_{i=1}^{n}\})$.*

Proof. Let $x \in X$, we have $x = \left[\frac{1}{n}, x\right]_{i=1}^{n} = \Psi_{\mu_0}\{(x)_{i=1}^{n}\}$. By definition of the operator $\mathcal{K}$, we have

$$\mathcal{K}x = \lim_{n\to\infty}([\frac{1}{n}, x]_{i=1}^{n}) = \lim_{n\to\infty}(\Psi_{\mu_0}\{(x)_{i=1}^{n}\}).$$

□

With the tools gathered in subsection 3.2, we can now construct in the framework of Mean-set probability theory an expectation operator $\mathbb{E}$ satisfying the law of large numbers for NCCCP spaces.

4.4. **Weighted mean-set (expectation) for NCCCP spaces.** Let (X, d_X) be a NCCCP (pseudo)metric space. Assume the space is E-chained with the set $E \subset X \times X$ symmetric containing the diagonal diag(X). Also assume that X is finitely compact, arc connected and the CC operation is unbiased.

Let $c, C \in \mathbb{R}_+^\star$ such that (X, d_X, E) is (c, C)-controlled, $n \geq 1$ and let $\mathit{\Gamma}_X = (X_{\text{Haus}}, E_{\text{Haus}}, \theta_\mathit{\Gamma})$ the E-chained VWM graph attached to X. As $H_n = \{h_1, \ldots, h_n\}$ represents the general term of the consecutive observations sets sequence necessary for the computation of the n-mean of a random variable ξ with distribution μ, it is on such set-elements that weight-function $\theta_\mathit{\Gamma}$ takes its arguments and reaches its global minimum.

The moving average of order n of a random variable ξ is defined to be a set $H_{n,0} = \{h_1^0, \ldots, h_n^0\}$ of elementary ends of homotopy classes of c-good E-paths in X, which minimizes weight-function $\theta_\mathit{\Gamma}$ on $\mathcal{P}_{(c,E)}(\mathit{\Gamma}_X)$ [10, 24, 30].

Write,

$$H_{n,0} = \arg \min_{H_n \in \mathcal{P}_{(c,E)}(\Gamma_X)} \theta_\Gamma(H_n).$$

The mean-set $H_{n,0}$ of order $k = n$ always exists as the function $\phi \circ \delta_{E,d_X}$ is bounded. This finite subset of n elements is compact and therefore non empty since ϕ is continuous on $\mathbb{R}_+$ and infinimas of δ_{E,d_X} are reached only on $\mathcal{P}_{(c,E)}(\Gamma_X)$. In general, this set is not necessarily reduced to a singleton, except in the case when the metric space X is strictly convex or negatively curved [13, 14, 21, 22].

4.4.1. *Weighted mean-set (expectation) of a random $V\Gamma_X$-variable.* Let (X, d_X) be a NCCCP space and $\Gamma_X = (X_{\text{Haus}}, E_{\text{Haus}}, \theta_\Gamma)$ a E-chained VWM graph attached to X. Let ξ be a random variable taking a finite number of values $h_1^0, \ldots, h_n^0$, with probabilities $p_1, \ldots, p_n$ respectively, such that $H_{n,0} = \{h_1^0, \ldots, h_n^0\}$ is the limit set of minimizers $H_{n,\ell}$ of weight-function θ_Γ on $\mathcal{P}_{(c,E)}(\Gamma_X)$, when $\ell \to \infty$.

Definition 4.3. *The weighted mean-set (expectation) or mean-set (to make short) of ξ, denoted by $\mathbb{E}_c\xi$, is the set*

$$\mathbb{E}_c\xi = \Psi_{\mu_\xi}(H_{n,0}) = \left[p_i,\, \mathcal{K}h_i^0\right]_{i=1}^n,$$

where Ψ_{μ_ξ} is the convexification function attached to μ_ξ and $\mathcal{K}$ the convexification operator of X.

Remark 4.5. *The set $\mathbb{E}_c\xi$ is well defined, thank to the uniqueness of the limit set $H_{n,0}$. Under the same assumptions as in Remarks 4.1 and 4.2, this definition generalizes, in the discrete case, the concept of mean-set (expectation) according to Mosina.*

The next proposition shows that this mathematical expectation for a random variable with values âĂŃâĂŃin a NCCCP metric space X always belongs to the convexifiable domain of X.

Proposition 4.5. *Given a weighted mean-set $\mathbb{E}_c\xi$ as previously defined in 4.3, then $\mathbb{E}_c\xi \in \mathcal{K}(X)$.*

Proof. Let $\mathbb{E}_c\xi$ as defined previously. Using linearity property of $\mathcal{K}$, we have

$$\mathbb{E}_c\xi = \left[p_i,\, \mathcal{K}h_i^0\right]_{i=1}^n = \mathcal{K}\left[p_i,\, h_i^0\right] \in \mathcal{K}(X).$$

$\square$

Possible generalizations of $\mathbb{E}_c\xi$: Let $\kappa \in \mathbb{N} \setminus \{0\}$ and Ψ_μ be the convexification function with n X-valued entries of X attached to μ. By analogy with the weight-function θ_Γ, we define a weight-function of class κ on $\mathcal{P}_{(c,E)}(\Gamma_X)$ by:

$$\theta_\Gamma^\kappa(H_n) = \int_{x\in X} \phi^\kappa \circ \delta_{E,d_X}(x, \eta_{H_n}(x))d\mu(x),$$

for all $H_n \in \mathcal{P}_{(c,E)}(\Gamma_X)$. And the mean-set $\mathbb{E}_c^\kappa\xi$ of order κ by

$$\mathbb{E}_c^\kappa\xi = \Psi_{\mu_\xi}\left(\arg \min_{H_n\in\mathcal{P}_{(c,E)}(\Gamma_X)} \theta_\Gamma^\kappa(H_n)\right).$$

It is obvious that the properties of the weight-function θ_Γ and the set $\mathbb{E}_c\xi$ also hold respectively for θ_Γ^κ and $\mathbb{E}_c^\kappa\xi$.

4.4.2. *Two numerical examples of mean-sets computation.*

- *When the metric space* $X = [0, 1]$. In this example, we consider $n = 2$, $c = 1$ and set $E = [0, 1] \times [0, 1]$. We are interested by 1-good E-paths of $n = 2$ steps in X. Let ξ be a random variable with the uniform distribution $\mu_\xi(x) = \frac{1}{2}$ on the bounded and closed interval $X = [0, 1]$ of $\mathbb{R}$, equipped with the real distance induced by the usual metric, $d_{[0,1]}(x, y) = |x - y|$ on $\mathbb{R}$. Also consider $\phi(x) = x^2$ as the discrepancy function.

To calculate the mean-set $\mathbb{E}_{c=1}\xi$, we start by checking the pair of reals $H_{2,0} = \{a, b\}$ that minimizes the weight-function θ_Γ, hereby defined as a cluster sum of squares by:

$$\theta_\Gamma(H_{2,0}) = \theta_\Gamma(a, b) = \frac{1}{2}\int_0^1 \min\left\{|x - a|^2, |x - b|^2\right\} dx.$$

Suppose $0 \le a \le b$ and consider as a partition of $[0, 1]$, the two disjoint intervals $\Omega_1 = [0, \frac{1}{2}(a + b)[$ and $\Omega_2 = [\frac{1}{2}(a + b), 1]$.

The explicite form of θ_Γ comes out after the following transformations:

$$\begin{aligned}
\theta_\Gamma(a,b) &= \tfrac{1}{2}\textstyle\int_{\Omega_1} \min\left\{|x-a|^2, |x-b|^2\right\} dx + \tfrac{1}{2}\int_{\Omega_2} \min\left\{|x-a|^2, |x-b|^2\right\} dx \\
&= \tfrac{1}{2}\textstyle\int_{\Omega_1} |x-a|^2 dx + \tfrac{1}{2}\int_{\Omega_2} |x-b|^2 dx \\
&= \tfrac{1}{2}\left[\tfrac{1}{3}(x-a)^3\right]_0^{\frac{1}{2}(a+b)} + \tfrac{1}{2}\left[\tfrac{1}{3}(x-b)^3\right]_{\frac{1}{2}(a+b)}^1 \\
&= \tfrac{1}{6}(a^3 + (1-b)^3 + \tfrac{1}{4}(b-a)^3).
\end{aligned}$$

A Critical point of θ_Γ is every point (a,b) solution of the following partial derivatives equations system:

$$\begin{array}{ll} \frac{\partial\theta_\Gamma}{\partial a}(a,b) & = \frac{1}{8}[4a^2-(b-a)^2]=0 \\ \frac{\partial\theta_\Gamma}{\partial b}(a,b) & = \frac{1}{8}[-4(1-b)^2+(b-a)^2]=0 \end{array}$$

This optimization problem has as unique solution, the point $(a=\frac{1}{4}, b=\frac{3}{4})$ at which the value of θ_Γ is $\frac{1}{96}$. The Hessian matrix of θ_Γ at (a,b) gives

$$H_{\theta_\Gamma}(a,b) = \begin{pmatrix} a+\frac{1}{4}(b-a) & \frac{1}{4}(a-b) \\ \frac{1}{4}(a-b) & 1-b+\frac{1}{4}(a-b) \end{pmatrix}.$$

At the point $(\frac{1}{4},\frac{3}{4})$, The Hessian matrix is $H_{\theta_\Gamma}(\frac{1}{4},\frac{3}{4}) = \begin{pmatrix} \frac{3}{8} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{3}{8} \end{pmatrix}$. The calculations give $\det H_{\theta_\Gamma}(\frac{1}{4},\frac{3}{4}) = \frac{1}{8} > 0$ and tr$H_{\theta_\Gamma}(\frac{1}{4},\frac{3}{4}) = \frac{3}{4} > 0$. Hence, the point $(\frac{1}{4},\frac{3}{4})$ is the minimum of θ_Γ.

Therefore, the pair $H_{2,0} = \{\frac{1}{4},\frac{3}{4}\}$ is the unique minimizer of θ_Γ. This subset is actually the "2-mean" of the random variable ξ.

Moreover, the uniform distribution of ξ on X allows us to consider the probability values $p(\Omega_i) = p_i = \frac{1}{2}$, for $i \in \{1,2\}$. If Ψ_ξ is the convexification function and $\mathcal{K}$ the convexification operator of $X=[0,1]$, then the mean-set of ξ is defined to be the set:

$$\mathbb{E}_1\xi = \Psi_\xi(H_{2,0}) = \left[\frac{1}{2}, \mathcal{K}\{\frac{1}{4}\}; \frac{1}{2}, \mathcal{K}\{\frac{3}{4}\}\right].$$

From the fact that $\mathbb{R}$ is a separable and strictly convex Banach space, hence negatively curved [13, 14], we can conclude that its subset $X=[0,1]$ is convexifiable, hence invariant for the convexification operator $\mathcal{K}$.

Therefore,

$$\mathbb{E}_1\xi = \left[\frac{1}{2}, \frac{1}{4}; \frac{1}{2}, \frac{3}{4}\right] = \{\frac{1}{2}\}.$$

Hence, the weighted mean-set is reduced to the singleton of element-set $\frac{1}{2}$ which is the ordinary classic theoretical mean of a uniform distribution over $[0,1]$.

This result shows that our definition of the weighted mean-set (expectation) is coherent with the analogous notion of the classic mean (average) of a Bochner-integrable random variable which, in general, is a singleton in any Banach space or negatively curved metric space.

- *When the metric space* $X = \mathbb{R}^2$.

We consider the random $\mathbb{R}^2$-variable ξ with distribution μ defined by $\mu(0,0) = \mu(0,3) = \mu(3,0) = \frac{1}{3}$ and for all other points (a,b), $\mu(a,b) = 0$. Equip $\mathbb{R}^2$ with the Euclidean distance d such that

$$d^2(x,y) = \sum_{i=1,\dots 2} (x_i - y_i)^2$$

and consider $\phi(x) = x^2$ to be the discrepancy function.

Now considering the distribution on $\mathbb{Z}^2$, the weight-function is defined at every point (a,b) of $\mathbb{Z}^2$ by

$$\begin{aligned} \theta_\Gamma(a,b) &= \textstyle\sum_{i=1,2,3} \left((a-x_i)^2 + (b-y_i)^2\right) \mu(x_i,y_i) \\ &= \tfrac{1}{3}\left(a^2 + b^2 + (a-3)^2 + b^2 + a^2 + (b-3)^2\right) \\ &= \tfrac{1}{3}\left(2a^2 + 2b^2 + (a-3)^2 + (b-3)^2\right). \end{aligned}$$

After computing the weight of each point of the triple, we find that $\theta_\Gamma(0,0) = 6$ which is less than $\theta_\Gamma(3,0) = 9$ and $\theta_\Gamma(0,3) = 9$. Therefore, the mean-set $\mathbb{E}_{c=1}\xi$ in $\mathbb{Z}^2$ is reduced to the point $(0,0)$, as Mosina noted in [25] (Remark 3.10).

Coming back on $\mathbb{R}^2$, recall that the minimum of θ_Γ is among the critical point (a,b) solution of the PDE following system:

$$\begin{aligned} \tfrac{\partial\theta_\Gamma}{\partial a}(a,b) &= \tfrac{1}{3}[4a + 2(a-3)] = 0 \\ \tfrac{\partial\theta_\Gamma}{\partial b}(a,b) &= \tfrac{1}{3}[4b + 2(b-3)] = 0 \end{aligned}$$

We find as solution of this system, the unique point $(a,b) = (1,1)$. The Hessian matrix of θ_Γ at (a,b) is constant $H_{\theta_\Gamma}(a,b) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. Hence $H_{\theta_\Gamma}(1,1) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. But, since $\det H_{\theta_\Gamma}(1,1) = 4 > 0$ and tr$H_{\theta_\Gamma}(1,1) = 4 > 0$. We conclude that $(1,1)$ is the minimum of θ_Γ.

Finally, the unique minimizer $(a,b) = (1,1)$ of θ_Γ is naturally the classical mean defined coordinate-wise in $\mathbb{R}^2$ by $(\mathbb{E}X, \mathbb{E}Y)$ for a given finite set of i.i.d. points $(x_1,y_1), \cdots, (x_n,y_n)$.

Moreover, the computation of the weight of this point $(1,1)$ shows $\theta_\Gamma(1,1) = 4$, which is smaller than $\theta_\Gamma(0,0) = 6$ and $\theta_\Gamma(3,0) = \theta_\Gamma(0,3) = 9$. Therefore, the mean-set of ξ is reduced to this single element-set $(1,1)$.

Remark 4.6. *A short verification of this result can be done in two ways. But before this, let remind that the convex combination operation defined on a metric space X can be naturally extended (uplifted) to act on subsets $A_1, \cdots, A_n$ of X by letting:*

$$[\lambda_i, A_i]_{i=1}^n \quad \textit{or} \quad [\lambda_i, A_i]_{i\in\{1,\dots,n\}} = \textit{cl}\left\{[\lambda_i, a_i]_{i=1}^n : a_i \in A_i, i = 1, \cdots, n\right\},$$

where cl in the right-hand side denotes the closure in X (see, e.g., [34] for more details). Note that taking the closure is necessary, since the Minkowski sum of two non-compact closed sets is not necessarily closed.

Coming back to our verification, we show the:

- First: Let $\mathcal{A} = \{O(0,0), A(3,0), B(0,3)\}$ be a finite three elements subset of $\mathbb{R}^2$. This can be used to generate $\mathbb{R}^2$. By definition, the Convex hull of $\mathcal{A}$, denoted by $\mathbf{H}(\mathcal{A})$, consists of all convex combinations of the generators, i.e. the elements in $\mathcal{A}$.

When computing in $\mathbb{R}^2$ the convex combination of the three points $O(0,0)$, $A(3,0)$ and $B(0,3)$ of $\mathcal{A}$ with parameter $\frac{1}{3}$ each, we find that,

$$\left[\frac{1}{3}, O; \frac{1}{3}, A; \frac{1}{3}, B\right] = \{(1,1)\},$$

as expected (see, e.g., [19, 20]).

- Second: Using our approach, let endow $\mathbb{R}^2$ with the metric induced by the infinity norm denoted by $\|(x,y)\|_\infty = \max(|x|, |y|)$. From the triangle OAB, define the open balls

$$\Omega_i = OAB \bigcap \mathcal{B}\left(M_i, \frac{3}{2}\right)$$

for $i \in 1,2,3$, with $M_1 = O$, $M_2 = A$ and $M_3 = B$. One can easily show that $(\Omega_i)_{i=1,2,3}$ represent a finite partition of OAB. For a uniform distribution ξ on the NCCCP metric space $\mathbb{R}^2$, consider $p(\Omega_i) = p_i = \frac{1}{3}$, for $i \in \{1,2,3\}$. Let Ψ_ξ be the convexification function and $\mathcal{K}$ the convexification operator of $\mathbb{R}^2$. Then the mean-set of ξ is defined to be the set:

$$\mathbb{E}_1\xi = \Psi_\xi(\mathcal{A}) = \left[\frac{1}{3}, \mathcal{K}\{O\}; \frac{1}{3}, \mathcal{K}\{A\}; \frac{1}{3}, \mathcal{K}\{B\}\right] = \{(1,1)\},$$

as expected.

This shows that for a distribution on a topological space, the mean-set can change dramatically according to the nature of the typical underlying space: discrete for $\mathbb{Z}^2$ and continue or convex for $\mathbb{R}^2$.

We now define the Empirical weighted mean-sets for NCCCP metric spaces with a CC operation.

4.4.3. *Empirical weighted mean-sets for NCCCP spaces.* Let $\mathbb{S}_\ell = (\xi_i)_{i=1\ldots\ell}$ be a finite sample of Γ_X-mutually i.i.d. random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and Ψ_μ the convexification function of X-valued n entries attached to the measure μ. The empirical measure attached to this sample is denoted by μ_ℓ and defined on the borel sets $B \in \mathcal{B}(X)$ by

$$\mu_\ell(B) = \frac{1}{\ell} \times |\{i \mid \xi_i(\omega) \in B,\ 1 \leq i \leq \ell\}|.$$

We attach to this measure the empirical weight-function, denoted by $\theta_{\Gamma,\ell} : \mathcal{P}_{(c,E)}(\Gamma_X) \to \mathbb{R}_+$, defined by

$$\theta_{\Gamma,\ell}(H_{n,\ell}) = \int_{x \in X} \phi \circ \delta_{E,d_X}(x, \eta_{H_{n,\ell}}(x)) d\mu_\ell(x),$$

where $\phi :\ \mathbb{R}^+ \to \mathbb{R}^+$ is a good function (i.e. positive, continuous, strictly increasing function and such that $\phi(0) = 0$).

The empirical moving average of order n of the sample $\mathbb{S}_\ell = (\xi_i(\omega))_{i=1,\ldots,\ell}$ is defined to be the set, denoted by $H_{n,0,\ell} = \{h_j^{0,\ell} \mid j = 1, \ldots, n\}$, of elements that minimize the empirical weight-function $\theta_{\Gamma,\ell}$. That is,

$$H_{n,0,\ell} = \underset{H_{n,\ell} \in \mathcal{P}_{(c,E)}}{\text{arg min}}\, \theta_{\Gamma,\ell}(H_{n,\ell}).$$

If the measure μ_ℓ has a probability distribution $p_i^{(\ell)}$, $i = 1, \ldots, n$ such that $\sum\limits_{i=1,\ldots n} p_i^{(\ell)} = 1$ (i.e. if μ_ℓ takes its values $h_1^{0,\ell}, \ldots,\ h_n^{0,\ell}$ on the Ω_i's with respective probabilities $p_1^{(\ell)}, \ldots,\ p_n^{(\ell)}$), then the empirical weighted mean set of the sample, denoted by $\overline{\mathbb{S}_\ell}$, is defined to be

$$\overline{\mathbb{S}_\ell} = \Psi_{\mu_\ell}(H_{n,0,\ell}) = \left[p_i^{(\ell)},\ \mathcal{K}h_i^{0,\ell}\right]_{i=1}^n.$$

Remark 4.7. *The set $\overline{\mathbb{S}_\ell}$ is not necessarily unique. Therefore arise the problem of minimization and set convergence of sequences of such subsets (see Subsection 3.3.1). In fact, it is a rather complex set optimization problem, since it does not obey the* *same existence or uniqueness properties required by the classical limit problem. For this reason, we can't study the SLLN with the usual facilities of classical convergence.*

Our main results can now be presented.

5. Main Results

From now on, $(X, d_X, [\cdot,\cdot])$ is a E-chained NCCCP (pseudo)metric space with $E \subset X \times X$ a symmetric containing the diagonal diag(X). The graph $\mathit{\Gamma}_X = (X_{\text{Haus}}, E_{\text{Haus}}, \theta_{\mathit{\Gamma}_X})$ is a chained vertex-weighted metric graph attached to X with $\theta_{\mathit{\Gamma}_X}$ its attached weight-function defined on $\mathcal{P}_{(c,E)}(\mathit{\Gamma}_X)$, and $\mathcal{P}_{(c,E)}(\mathit{\Gamma}_X) \subset \mathcal{E}_n$.

5.1. **Topological characterization of mean-sets for NCCCP spaces.** In the context of topological geometry, the weight of a mean-element can lead to formulate a characterization of this element as a membership of the mathematical expectation set. This is based on the fact that, for any $z \in \mathbb{E}\xi$ and every fixed vertex $u \in V\mathit{\Gamma}_X$, the closed ball of minimal radius $r = \theta_{\mathit{\Gamma}_X}(z)$ centered at u is lies in limsup of the cocentric closed balls of the same center u but with radius r running in the product $[1,2]\theta_{\mathit{\Gamma}_X}(\mathbb{E}\xi)$ of sets $[1,2]$ and $\theta_{\mathit{\Gamma}_X}(\mathbb{E}\xi)$.

Theorem 5.1. ***(Topological characterization)*** *Let ξ be a random $\mathit{\Gamma}_X$-variable with mean-set $\mathbb{E}_c\xi = \Psi_\xi(H_{n,0}) = [p_i,\, \mathcal{K}h_i^0]_{i=1}^n$, where Ψ_ξ is the convexification function of n entries, $H_{n,0} = \{h_1^0, \ldots, h_n^0\}$ is the limit minimizing set of $\theta_{\mathit{\Gamma}_X}$ and $\mathcal{K}$ the convexification operator of X. Then, for all fixed $u \in V\mathit{\Gamma}_X$, the implication*

$$(z \in \mathbb{E}_c\xi) \Longrightarrow \left(\overline{B}(u, \theta_{\mathit{\Gamma}_X}(\Psi_\xi^{-1}(z))) \subseteq \limsup_{\ell\to\infty}\{\overline{B}(u, (1+\frac{1}{\ell})\theta_{\mathit{\Gamma}_X}(\Psi_\xi^{-1}(z)))\}\right)$$

holds, with $\Psi_\xi^{-1}(z) = H_{n_p,0}$, where $(H_{n_p,0})$ is a convergent subsequence sets of minimizers $(H_{n,0})$ with $|H_{n_p,0}| \leq n$.

Proof. From definition 4.3, our mean-set is $\mathbb{E}_c\xi = \Psi_{\mu_\xi}(H_{n,0}) = [p_i,\, \mathcal{K}h_i^0]_{i=1}^n$. Let $z \in \mathbb{E}_c\xi$, there exists a partition I_1 and I_2 of $\{1, \ldots, n\}$ such that

$$z = \left[m_1\left[\frac{p_j}{m_1}, \mathcal{K}h_j^0\right]_{j\in I_1}; m_2\left[\frac{p_j}{m_2}, \mathcal{K}h_j^0\right]_{j\in I_2}\right],$$

where $m_\ell = \sum\limits_{i\in I_\ell}\lambda_i$ for $\ell \in 1, 2$.

By a good adjustment of variables to make simply, write

$$z = \left[\left[\alpha_j, \mathcal{K}h_j^0 \right]_{j\in I_1} ; \left[\alpha_j, \mathcal{K}h_j^0 \right]_{j\in I_2} \right],$$

with $\sum_{j=1}^{n_p} \alpha_j = 1$, i.e. $z = [\alpha_i,\, \mathcal{K}h_i^0]_{i=1}^{n_p}$. Therefore, $z = \Psi_\mu(H_{n_p,0})$ for some measure μ with distribution α_i and where $(H_{n_p,0})$ is a convergent subsequence sets of minimizers $(H_{n,0})$ with $|H_{n_p,0}| \leq n$. Thank to Lember's results on the consistency of empirical k-centres with respect to Hausdorff topology τ, i.e. "every subsequence of empirical k-centres has a further subsequence almost surely converging to a theoretical k-centre" (see [24]).

Now, by fixing a vertex $u \in V\mathit{\Gamma}_X$ and setting $A_\ell = \overline{B}(u,\, r_\ell)$ with

$$r_\ell = (1 + \frac{1}{\ell})\theta_{\mathit{\Gamma}_X}(\Psi_\xi^{-1}(z)) \in [1,2]\theta_{\mathit{\Gamma}_X}(\mathbb{E}\xi)$$

for $\ell \geq 1$, we have that $A_\infty = \overline{B}(u, \theta_{\mathit{\Gamma}_X}(\Psi_\xi^{-1}(z)))$ and the balls (A_ℓ) form a decreasing sets sequence with limit A_∞.

Thus, for every $L \geq 1$, we have $\overline{B}(u, \theta_{\mathit{\Gamma}_X}(\Psi_\xi^{-1}(z))) \subseteq \bigcup_{\ell > L} A_\ell$. Therefore,

$$\overline{B}(u, \theta_{\mathit{\Gamma}_X}(\Psi_\xi^{-1}(z))) \subseteq \bigcap_{L\geq 1} \bigcup_{\ell > L} A_\ell.$$

Hence the theorem. □

We now formulate the SLLN (i.e. the almost sure consistency) with respect to a weighted mean-set for a wider class, i.e. a sample of mutually i.i.d. random $\mathit{\Gamma}_X$-variables in a finitely compact NCCCP metric spaces X.

5.2. **SLLN for NCCCP spaces.** A metric space X is called finitely compact if each closed and bounded subset of X is compact. A function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is called convex if $f\left((1-t)x + ty\right) \leq (1-t)f(x) + tf(y)$ for all $x, y \in X$
and $t \in [0;1]$.

Theorem 5.2. ***(SLLN).*** *Let X be a finitely comapct NCCCP space and $\mathbb{S}_\ell = (\xi_i)_{i=1,\dots\ell}$ a sample of random $\mathit{\Gamma}_X$-variables mutually i.i.d. on $(\Omega, \mathcal{F}, \mathbb{P})$, with the same distribution as ξ_1. If the space $X = \mathbb{R}^n$ or the function ϕ is strictly increasing or convex, then $\overline{\mathbb{S}_\ell} \longrightarrow \mathbb{E}_c\xi_1$ almost surely, when $\ell \to \infty$.*

The proof of this theorem is coming after the following remark and lemma whose proofs remain valid in separable Banach spaces.

Remark 5.1.

(i) *Theorem 5.2 states that, if ℓ is large enough then the empirical mean-set of the random Γ_X-variables mutually i.i.d. coincides with the unique singleton mean-set. This result remains valid when (X, τ) is a normed topological space on which the function ϕ is strictly increasing or convex or if $x_n \xrightarrow{\tau} x$ and $\|x_n\| \longrightarrow \|x\|$ in X then $x_n \longrightarrow x$.*

(ii) *Given a sequence of random X-variables $(\xi_\ell)_{\ell \geq 1}$ mutually i.i.d with empirical distributions $p_\ell(\omega, \cdot)$ such that $\mathbb{P}_{\xi_\ell} = p \in]0, 1[$.*

This result on the convergence was also taken up by Cuestra and Matran [10], in the form

$$\mu(\{\omega \mid p_{\ell n}(\omega, \cdot) \underset{\ell \to \infty}{\longrightarrow} p\}) = 1$$

with respect to the distance d for almost every $\omega \in \Omega$.

Lemma 5.1. *With the same assumptions as in Theorem 5.2 and under one of the following four conditions, notably if (X, τ) is normed topological vector space or $X = \mathbb{R}^n$ or ϕ is strictly increasing or ϕ is convex, then $H_{n,0,\ell} \underset{\ell \to \infty}{\longrightarrow} H_{n,0,1}$ almost surely.*

Proof. Using the same approach as in [10, 24], let $(H_{n,0,\ell})_{\ell \geq 1}$ be a sequence of n elements subsets of X, where $H_{n,0,\ell} = \{h_j^{0,\ell} \mid j = 1, \ldots, n\}$. By Hahn-Banach theorem, there exists in $X^\star$, the dual space of X, a continuous sequence $(f_1, f_2, \ldots, f_n)$ of functions such that $f_i(h_i^{0,\ell}) \neq f_i(h_j^{0,\ell})$, for all $j \neq i$, with $i = 1, \ldots, n$.

For a good choice of indices in each subset $H_{n,0,\ell}$, take $\ell \geq 1$ and set $\ell(1)$ the smallest index i such that

$$|f_1(h_i^{0,\ell}) - f_1(h_1^{0,1})| \leq |f_1(h_j^{0,\ell}) - f_1(h_1^{0,1})|$$

holds, for all $j \neq i$. We claim that $h_{\ell(1)}^{0,\ell}$'s converge (weakly) to $h_1^{0,1}$.

Indeed, let's proceed by absurdity. Assume that $h_{\ell(1)}^{0,\ell}$ does not converge to $h_1^{0,1}$. Then, there exist a function $f \in X^\star$, a real $\epsilon > 0$ and a subsequence $(H_{n,0,\ell_p})$ of $(H_{n,0,\ell})$ such that, for all ℓ_p, the following inequality holds

$$|f(h_{\ell(1)}^{0,\ell_p}) - f(h_1^{0,1})| > \epsilon.$$

But according to the hypotheses, there exists a subsubsequence $(H_{n,0,\ell_{pq}})$ of the subsequence $(H_{n,0,\ell_p})$ such that $H_{n,0,\ell_{pq}} \longrightarrow H_{n,0,1}$. That is, such that the $h_{\ell'_{pq}(i)}^{0,\ell_{pq}}$'s converge to $h_i^{0,1}$, for $i = 1, \ldots, n$.

But, $\ell_{pq}(1)$ coincides several times with $\ell'_{pq}(i)$, for some indices i. Next, we can consider a new sub-subsequence, denote in the same way as the previous one, such that for indices i, the $h^{0,\ell_{pq}}_{\ell_{pq}(1)}$'s converge to $h^{0,1}_i$.

Furthermore, we obtain

$$|f_1(h^{0,\ell_{pq}}_{\ell_{pq}(1)}) - f_1(h^{0,1}_1)| \leq |f_1(h^{0,\ell_{pq}}_{\ell'_{pq}(1)}) - f_1(h^{0,1}_1)| \underset{\ell\to\infty}{\longrightarrow} 0.$$

Hence $i = 1$, according to the condition on the f_i's in $X^\star$ and their arguments. Therefore, $h^{0,\ell_{pq}}_{\ell_{pq}(1)}$'s converges to $h^{0,1}_1$. Which contradicts the hypothesis. So $h^{0,\ell}_{\ell(1)}$'s converges to $h^{0,1}_1$.

We now define the integer $\ell(2)$ as the smallest index i such that

$$|f_2(h^{0,\ell}_i) - f_2(h^{0,1}_2)| \leq |f_2(h^{0,\ell}_j) - f_2(h^{0,1}_2)|$$

holds, for all $j \neq i$.

Proceeding by the same way, we say that $h^{0,\ell_{pq}}_{\ell_{pq}(2)}$'s converge to $h^{0,1}_2$. Under the same conditions on f_i's and their arguments, there exists a natural integer L such that $\ell(2) \neq \ell(1)$, for all $\ell \geq L$.

From now, proceeding step by step, we thus define the sequence of integers $\ell(1), \ell(2), \ldots, \ell(n)$, mutually distinct and whose respective existence implies pointwise convergence $h^{0,\ell}_j \underset{\ell\to\infty}{\longrightarrow} h^{0,1}_j$, for all $j = 1, \ldots, n$. Therefore, the set convergence $H_{n,0,\ell} \underset{\ell\to\infty}{\longrightarrow} H_{n,0,1}$ is established.

Using the measure probabilistic language, we say

$$\mu\left(\left\{\omega \mid \begin{array}{c} \forall (H_{n,0,\ell_p}) \subset \mathcal{E}_n,\ \exists (H_{n,0,\ell_{pq}}) \subset H_{n,0,\ell_p} \\ \text{such that}\, H_{n,0,\ell_{pq}} \underset{\ell\to\infty}{\longrightarrow} H_{n,0,1} \end{array}\right\}\right) = 1,$$

for almost every $\omega \in \Omega$. Hence the Lemma. □

We are now ready to give the proof of the SLLN.

Proof. of Theorem 5.2 (SLLN) The result to be proved can take the form:

$$\Psi_{\mu_\ell}(H_{n,0,\ell}) \underset{\ell\to\infty}{\longrightarrow} \Psi_{\mu_1}(H_{n,0,1})$$

almost surely.

From the definition of function Ψ_{μ_ℓ}, this convergence can be expressed in terms of convex combination convergence by

$$[p_i^{(\ell)},\, \mathcal{K}h_i^{0,\ell}]_{i=1}^n \underset{\ell\to\infty}{\longrightarrow} [p_i,\, \mathcal{K}h_i^{0,1}]_{i=1}^n.$$

But according to Remark 5.1, $p_i^{(\ell)} \underset{\ell\to\infty}{\longrightarrow} p_i$ and by Lemma 5.1, $H_{n,0,\ell} \underset{\ell\to\infty}{\longrightarrow} H_{n,0,1}$.

Therefore, linearity and continuity properties of the CC operation lead to the conclusion that

$$\mathbb{P}\left(\left\{\omega \mid \begin{array}{c}\text{Every subsequence } \Psi_{\mu_\ell}(H_{n,0,\ell_p}) \text{ of } \Psi_{\mu_\ell}(H_{n,0,\ell}) \text{ admits a sub}-\\ \text{subsequence } \Psi_{\mu_\ell}(H_{n,0,\ell_{pq}}) \text{ such that } \Psi_{\mu_\ell}(H_{n,0,\ell_{pq}}) \longrightarrow \Psi_{\mu_1}(H_{n,0,1})\end{array}\right\}\right) = 1.$$

Which traduces the almost surely convergence of empirical mean-set to theoretical mean-set.

Hence the theorem. □

The SLLN is often not directly usable as a probability/statistic tool. Nevertheless, it retains all its importance as it guarantees the consistency of estimators. Indeed, the problem of consistency of an estimator always arises when it is necessary to find conditions under which an empirical parameter can be considered as an estimator (generally unknown) of a theoretical parameter. The solution of such a problem is usually found, under certain conditions related to the topology of the space and the uniqueness of the theoretical parameter, in terms of weak or strong convergence accordingly.

5.2.1. *Consistency of empirical means-sets.* The consistency result is moreover justified by the SLLN we have just established.

Theorem 5.3. ***(Consistency)*** *Let $\Gamma_X = (X_{Haus}, E_{Haus}, \theta_\Gamma)$ be a chained VWM graph attached to X and $\mathbb{S}_\ell = (\xi_i)_{i=1,\dots\ell}$ a sample of random Γ_X-variables mutually i.i.d. on $(\Omega, \mathcal{F}, \mathbb{P})$, with the same distribution as ξ_1. Let μ_ℓ be an empirical measure constructed from a stationary and ergodic sequence of ℓ observations of a μ distribution on a complete space $(\Omega, \mathcal{F}, \mathbb{P})$. If $\mathbb{E}_c\xi_1$ is a unique singleton, then $\overline{\mathbb{S}_\ell}$ almost surely converges to $\mathbb{E}_c\xi_1$.*

Proof. In the case in question, the consistency of empirical mean-sets

$$\overline{\mathbb{S}_\ell} = \left[p_i^{(\ell)},\, \mathcal{K}h_i^{0,\ell}\right]_{i=1}^n$$

as an estimator of the theoretical mean-set

$$\mathbb{E}_c\xi_1 = \left[p_i,\, \mathcal{K}h_i^{0,1}\right]_{i=1}^n$$

is deduced from the consistency results of the sampled means of order $k = n$, available in [24,29,30,35] and thanks to continuity of the convexification function Ψ_{μ_ℓ} of X associated to the distribution μ_ℓ. $\square$

5.3. **Chebyshev's inequality for NCCCP spaces.** Given that the mean-set of a random variable ξ with values in a NCCCP space X is always reduced to a singleton and the variance is finite, one may be interested, as in the classical general real space, to the boundness of the probability that a value of such a random variable will differ from the mean-set by more than a fixed real number ϵ.

In the next theorem, we propose and prove an analogue of Chebyshev's inequality for a random variable ξ with values in a NCCCP space X.

Theorem 5.4. ***(Chebyshev's inequality)*** *Let $\Gamma_X = (V\Gamma_X, E\Gamma_X, \theta_{\Gamma_X})$ be a chained VWM graph attached to X and $\mathbb{S}_\ell = (\xi_i)_{i=1,\ldots\ell}$ a sample of random Γ_X-variables mutually i.i.d. on $(\Omega, \mathcal{F}, \mathbb{P})$, with the same distribution law as ξ_1. If $\mathbb{E}_c\xi_1 = \Psi_{\mu_1}(H_{n,0,1})$ is the mean of ξ_1, then for $X = \mathbb{R}^n$ or ϕ strictly increasing, there exists a constant $\lambda = \lambda_{(\Gamma_X,\xi_1)} > 0$ such that $\mathbb{P}(\{\overline{\mathbb{S}_n} \neq \mathbb{E}_c\xi_1\}) \leq n^{-1}\lambda$.*

Proof. Let $\mathbb{E}_c\xi_1 = \Psi_{\mu_1}(H_{n,0,1}) = \{e\}$ for some $e \in V\Gamma_X$. As we are in the case of a single mean-set, it suffices to show that

$$\mathbb{P}(\{\exists u \in V\Gamma_X \setminus \{e\} \text{ s.t. } \theta_{\Gamma_X,\ell}(u) \leq \theta_{\Gamma_X,\ell}(e)\}) \leq \ell^{-1}\lambda.$$

$\square$

6. Conclusive Remarks

In this paper, the question of extending the mean-set probabilistic approach to metric spaces is investigated. Some geometric tools capable of supporting, from an algebraic point of view, a mathematical basis for computing some probabilities have been proposed. We focused on finitely compact NCCCP spaces (X, d_X) and considered a metric graph Γ_X that we provided with a compatible metric which extends the classic words distance. We defined a weight-function θ_Γ on the set of vertices of this graph which achieves its minimums on $\mathcal{P}_{(c,E)}(X)$, the set of the ends of the c-good E-finite paths in X.

We defined a generalised mean-set $\mathbb{E}_c\xi$ of a given Γ_X-random variable ξ as the image, by a convexification function Ψ_{μ_ξ} of order n of X, of the compact subset that minimizes the weight-function attached to the measure μ_ξ.

As results, we generalized several concepts constructed about the expectation and the SLLN. We observe that:

(i) When $n = 1$, any one-step path joining two points x and y in space X is reduced to a usual continuous morphism between the two points. As the ends of such paths are all reduced to singletons, we find here in the general context of the definition of classic mean or barycenter in metric spaces.

(ii) Our definition generalizes that of Terán and Molchanov [34] when the space is convexifiable and the discrepancy function is $\phi(x) = x$.

(iii) In discrete cases, for $\phi(x) = x^2$ and $c = 1$, we find out in the context of Mosina [25].

(iv) Our definition also generalizes the Fréchet's mean concept [17] and coincide with it after some adjustement when $\phi(x) = x^2$.

(v) If $p = 1$ in $\phi(x) = x^p$, we have an analogue of the median of the distribution μ and our central parameter of class p is just the classic mean value.

(vi) Our definition naturally generalizes (within one parameter) the notion of moving average (expectation) of order $k = n$ and $E = X \times X$.

(vii) In a separable Banach space X equipped with a convex combination operation, by considering the simple function η_{H_n} as we definied (see 4.2.1) and under the same conditions on the weights λ_i's, our approach comes out with results similar to those of Bator and Ziéba [4] with the metric combination operation or Terán and Molchanov with the convex combination operation.

We also characterized mean-sets for random variables and proposed a version of the strong law of large numbers, for finitely compact NCCCP spaces.

Our approach is open to some applications in cluttered environments that need to be well checked.

ACKNOWLEDGMENT

The authors are grateful to the reviewers of AMSJ.

References

[1] A. Ahidar-Coutrix, T. Le Gouic, Q. Paris: *On the rate of convergence of empirical barycentres in metric spaces: curvature, convexity and extendible geodesics*, HAL Id: hal-01810530, (2019). Available at https://hal.archives-ouvertes.fr/hal-01810530v3.

[2] A.S. Árnadóttir, W. Lederle, R.G. Möller: *Cayley–Abels graphs and invariants of totally disconnected, locally compact groups*, Journal of the Australian Mathematical Society (2022), 1–33.

[3] D. Balding, P.A. Ferrari, R. Fraiman, M. Sued: *Limit theorems for sequences of random trees*, Test, **18**(2) (2009), 302–315.

[4] A. Bator, W. Ziéba: *Expectation in metric spaces and characterizations of Banach spaces*, Demonstratio Mathematica, **42**(4) (2009), 901–908.

[5] A. Bator, W. Ziéba: *On some definition of expectation of random element in metric space*, Annales Universitatis Mariae Curie-Sklodowska, **63**(1) (2009), 39.

[6] N.P. Brown: *Topological dynamical systems associated to II1-factors*, Advances in Mathematics, **227**(4) (2011), 1665–1699.

[7] V. Capraro, T. Fritz: *On the axiomatization of convex subsets of Banach spaces*, Proceedings of the American Mathematical Society, **141**(6) (2013), 2127–2135.

[8] C.-S. Chang, W. Liao, Y.-S. Chen, L.-H. Liou: *A mathematical theory for clustering in metric spaces*, IEEE Transactions on Network Science and Engineering, **3**(1) (2016), 2–16.

[9] A. Colubi, M. López-Diaz, J. S. Dominguez-Menchero, M.A. Gil: *A generalized strong law of large numbers*, Probability Theory and Related Fields, **114**(3) (1999), 401–417.

[10] J.A. Cuesta, C. Matrán: *The strong law of large numbers for k-means and best possible nets of Banach valued random variables*, Probability theory and related fields, **78**(4) (1988), 523–534.

[11] F.S. De Blasi, L. Tomassini: *On the Strong Law of Large Numbers in Spaces of Compacts Sets*, Journal of Convex Analysis, **18**(1), (2011), 285–300.

[12] Y. De Cornulier, P. De La Harpe: *Metric geometry of locally compact groups*, EMS Tracts in Mathematics, **25**, 2016.

[13] P.R. De Fitte: *Théorème ergodique ponctuel et lois fortes des grands nombres pour des points aléatoires d'un espace métrique à courbure négative*, The Annals of Probability, **25**(2) (1997), 738–766.

[14] P.R. De Fitte: *Autour de la notion de mesure*, HDR thesis, Université de Rouen, 2004.

[15] H.W. De Waard: *A new approach to distributed data fusion*, Thesis, fully internal, Universiteit van Amsterdam, 2008.

[16] S. Djimnaibeye, D. Tieudjo, N. Youmbi: *Probability on groups and an application to cryptography*, Computer Science Journal of Moldova, **69**(3) (2015), 360–375.

[17] M. Fréchet: *Les éléments aléatoires de nature quelconque dans un espace distancié*, Annales de l'institut Henri Poincaré, **10**(4) (1948), 215–310.

[18] C.E. GINESTET, A. SIMMONS, E.D. KOLACZYK: *Weighted Frechet means as convex combinations in metric spaces: properties and generalized median inequalities*, Statistics & Probability Letters, **82**(10) (2012), 1859–1863.

[19] R.V. HELGASON: *Convex combinations of two points: What are these, anyway?* , EMIS 8394, July 8, 2004 [Convex Combinations 1].

[20] R.V. HELGASON: *Convex combinations of more than two points*, EMIS 8394, July 8, 2004 [Convex Combinations 2].

[21] W. HERER: *Espérance mathématique d'une variable aléatoire à valeurs dans un espace métrique à courbure négative*, Comptes rendus de l'Académie des sciences. Série 1, Mathématique, **306**(15) (1988), 681–684.

[22] W. HERER: *Mathematical expectation and strong law of large numbers for random variables with values in a metric space of negative curvature*, Probab. Math. Statist, **13**(1) (1992), 59–70.

[23] A. KARLSSON, F. LEDRAPPIER: *On laws of large numbers for random walks*, The Annals of Probability, **34**(5) (2006), 1693–1706.

[24] J. LEMBER: *On minimizing sequences for k-centres*, Journal of Approximation Theory, **120**(1) (2003), 20–35.

[25] N. MOSINA: *Probability on graphs and groups: theory and applications*, Ph D. Thesis, Columbia University, (2009). Available at http://www.math.columbia.edu/ thaddeus/theses/2009/mosina.pdf

[26] N. MOSINA, A. USHAKOV: *Mean-set attack: cryptanalysis of Sibert et al. authentication protocol*, Journal of Mathematical Cryptology, **4**(2) (2010), 149–174.

[27] N. MOSINA, A. USHAKOV: *Strong law of large numbers on graphs and groups*, Groups Complexity Cryptology, **3**(1) (2011), 67–103.

[28] E. MOURIER: *Eléments aléatoires dans un espace de Banach*, Annales de l'institut Henri Poincaré, **13**(3) (1953), 161–244.

[29] D. POLLARD: *Strong consistency of k-means clustering*, The Annals of Statistics, (1981), 135–140.

[30] D. POLLARD: *Quantization and the method of k-means*, IEEE Transactions on Information theory, **28**(2) (1982), 199–205.

[31] K.-T. STURM: *Nonlinear martingale theory for processes with values in metric spaces of nonpositive curvature*, The Annals of Probability, **30**(3) (2002), 1195–1222.

[32] K.-T. STURM: *Probability measures on metric spaces of nonpositive curvature, Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces*, Lecture Notes from a Quarter Program on Heat Kernels, Random Walks, and Analysis on Manifolds and Graphs: April 16-July 13, 2002, Emile Borel Centre of the Henri Poincaré Institute, (2003), 338-357.

[33] H. SVERDRUP-THYGESON: *Strong law of large numbers for measures of central tendency and dispersion of random variables in compact metric spaces*, The Annals of Statistics, (1981), 141–145.

[34] P. TERÁN, I. MOLCHANOV: *The law of large numbers in a metric space with a convex combination operation*, Journal of theoretical probability, **19**(4) (2006), 875–898.

[35] M. THORPE, F. THEIL, A.M. JOHANSEN, N. CADE: *Convergence of the k-means minimization problem using* Γ*-convergence*, SIAM Journal on Applied Mathematics, **75**(6) (2015), 2444–2474.

[36] A. TORTRAT: *Lois de probabilité sur un espace topologique complètement régulier et produits infinis à termes indépendants dans un groupe topologique*, Annales de l'IHP Probabilités et statistiques, **1**(3) (1965), 217–237.

[37] N. TRAN THUAN: *An approach for metric space with a convex combination operation and applications*, arXiv e-prints, arXiv–1501, (2015).

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE OF THE FACULTY OF SCIENCE
THE UNIVERSITY OF NGAOUNDERE
P.O. BOX 454, NGAOUNDERE,
CAMEROON.
Email address: fotch2016@gmail.com

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE OF THE FACULTY OF SCIENCE
THE UNIVERSITY OF NGAOUNDERE
P.O. BOX 454, NGAOUNDERE,
CAMEROON.
Email address: elvishoupa@gmail.com

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE OF THE NATIONAL SCHOOL OF AGRO-INDUSTRIAL SCIENCE
THE UNIVERSITY OF NGAOUNDERE
P.O. BOX 455, NGAOUNDERE,
CAMEROON.
Email address: tieudjo@gmail.com