Funano, K.
Osaka J. Math.
47 (2010), 911–920

RATE OF CONVERGENCE OF STOCHASTIC PROCESSES WITH VALUES IN $\mathbb{R}$-TREES AND HADAMARD MANIFOLDS

KEI FUNANO

(Received May 27, 2009)

Abstract

Under K.-T. Sturm's formulation, we obtain a Gaussian upper bound for tail probability of mean value of independent, identically distributed random variables with values in $\mathbb{R}$-trees and Hadamard manifolds.

1. Introduction and statement of the main result

The aim of this paper is to study the weak Law of Large Numbers for CAT(0)-space-valued stochastic processes (see Subsection 2.1 for the definition of CAT(0)-spaces).

Let N be a CAT(0)-space and $(\Omega, \Sigma, \mathbb{P})$ a probability space. Given a random variable $W\colon \Omega \to N$ such that the push-forward measure $W_*\mathbb{P}$ of $\mathbb{P}$ by W has the finite moment of order 2, we define its *expectation* $\mathbb{E}_{\mathbb{P}}(W)$ by the barycenter of the measure $W_*\mathbb{P}$ (the definition of the barycenter is in Subsection 2.1). In [8, Theorem 4.7], K.-T. Sturm introduced a natural definition of mean value of n-points $y_1, \dots, y_n$ in N, called *inductive mean value* and denoted by $(1/n)\sum_{i=1,\dots,n}^{\rightarrow} y_i$ (see Definition 2.5 for precise definition). For an independent, identically distributed N-valued random variables $(Y_i)_{i=1}^{\infty}$ on the probability space Ω, he obtained the weak Law of Large Numbers proving the following inequality

$$(1.1)\qquad \int_{\Omega} d_N\left(\frac{1}{n}\sum_{i=1,\dots,n}^{\rightarrow} Y_i(\omega), \mathbb{E}_{\mathbb{P}}(Y_1)\right)^2 d\mathbb{P}(\omega) \le \frac{1}{n}\int_{\Omega} d_N(Y_1(\omega), \mathbb{E}_{\mathbb{P}}(Y_1))^2\, d\mathbb{P}(\omega).$$

He also proved the strong Law of Large Numbers ([8, Theorem 4.7, Proposition 6.6]).

Motivated by Sturm's work, using the results of the theory of Lévy–Milman concentration of 1-Lipschitz maps obtained in [4, 5], we obtain the following Gaussian estimate.

Theorem 1.1. *Let $(Y_i)_{i=1}^{\infty}$ be a sequence of independent, identically distributed random variables on a probability space $(\Omega, \Sigma, \mathbb{P})$ with values in an $\mathbb{R}$-tree T. We*

2000 Mathematics Subject Classification. 53C21, 53C23.

assume that the support of the measure $(Y_1)_*\mathbb{P}$ *has bounded diameter* D. *Then, for any* $r > 0$, *we have*

$$\mathbb{P}\left(\left\{\omega \in \Omega \;\middle|\; d_T\left(\frac{1}{n}\overrightarrow{\sum_{i=1,\dots,n}} Y_i(\omega), \mathbb{E}_{\mathbb{P}}(Y_1)\right) \geq r\right\}\right) \leq 4e^{4/75}e^{-nr^2/150D^2}.$$

See Subsection 2.1 for definition of $\mathbb{R}$-trees.

In the case where N is an Hadamard manifold, we also obtain the following. For any $m \in \mathbb{N}$, we put

$$A_m := e^{1/(2m)}\left\{1 + \frac{\sqrt{\pi}e^{(m+1)/(4m-2)}e^{\pi^2}}{2}\right\}$$

and

$$\tilde{A}_m := e^{1/(4m)}\{1 + \sqrt{\pi}e^{(m+1)/(4m-2)}\}.$$

Note that both A_m and $\tilde{A}_m$ are bounded from above by universal constant $C > 0$.

Theorem 1.2. *Let* $(Y_i)_{i=1}^{\infty}$ *be a sequence of independent, identically distributed random variables on a probability space* $(\Omega, \Sigma, \mathbb{P})$ *with values in an* m*-dimensional Hadamard manifold* N. *We assume that the support of the measure* $(Y_1)_*\mathbb{P}$ *has bounded diameter* D. *Then, for any* $r > 0$, *we have*

$$\begin{aligned}&\mathbb{P}\left(\left\{\omega \in \Omega \;\middle|\; d_N\left(\frac{1}{n}\overrightarrow{\sum_{i=1,\dots,n}} Y_i(\omega), \mathbb{E}_{\mathbb{P}}(Y_1)\right) \geq r\right\}\right)\\ &\leq \min\{A_m e^{-nr^2/16D^2m}, \tilde{A}_m e^{-nr^2/32D^2m}\}.\end{aligned}$$

There are many other way to define a mean value of points in a CAT(0)-space (see Remark 2.6). For example, in [2], A. Es-Sahib and H. Heinich introduced an another notion of mean value and expectation. They obtained the strong Law of Large Numbers under their definition. In this paper, we treat only Sturm's formulation.

2. Preliminaries

2.1. Basics of CAT(0)-spaces. In this subsection we explain several terminologies in geometry of CAT(0)-spaces. We refer to [8] for the details of the results on CAT(0)-spaces mentioned below.

Let (X, d_X) be a metric space. A rectifiable curve $\gamma\colon [0, 1] \to X$ is called a *geodesic* if its arclength coincides with the distance $d_X(\gamma(0), \gamma(1))$ and it has a constant speed, i.e., parameterized proportionally to the arclength. We say that a metric space is a *geodesic space* if any two points are joined by a geodesic between them. If any two points are joined by a unique geodesic, then the space is said to be *uniquely geodesic*. A complete geodesic space X is called a CAT(0)-*space* if we have

$$d_X\left(x,\, \gamma\left(\frac{1}{2}\right)\right)^2 \le \frac{1}{2} d_X(x,\, y)^2 + \frac{1}{2} d_X(x,\, z)^2 - \frac{1}{4} d_X(y,\, z)^2$$

for any $x, y, z \in X$ and any geodesic $\gamma\colon [0,1] \to X$ from y to z. For example, Hadamard manifolds, Hilbert spaces, and $\mathbb{R}$-trees are all CAT(0)-spaces. An $\mathbb{R}$-*tree* is a complete geodesic space such that the image of every simple path is the image of a geodesic.

It follows from the next theorem that CAT(0)-spaces are uniquely geodesic.

Theorem 2.1 (cf. [8, Corollary 2.5]). *Let N be a* CAT(0)-*space and $\gamma, \eta\colon [0, 1] \to N$ be two geodesics. Then, for any $t \in [0, 1]$, we have*

$$d_N(\gamma(t),\, \eta(t)) \le (1 - t)\, d_N(\gamma(0),\, \eta(0)) + t\, d_N(\gamma(1),\, \eta(1)).$$

Let N be a CAT(0)-space. We denote by $\mathcal{P}^2(N)$ the set of all Borel probability measure ν on N having the finite moment of order 2, i.e.,

$$\int_N d_N(x,\, y)^2\, d\nu(y) < +\infty$$

for some (hence all) $x \in N$. A point $x_0 \in N$ is called the *barycenter* of a measure $\nu \in \mathcal{P}^2(N)$ if x_0 is the unique minimizing point of the function

$$N \ni x \mapsto \int_N d_N(x,\, y)^2\, d\nu(y) \in \mathbb{R}.$$

We denote the point x_0 by $b(\nu)$. It is well-known that every $\nu \in \mathcal{P}^2(N)$ has the barycenter ([8, Proposition 4.3]).

A simple variational argument implies the following lemma.

Lemma 2.2 (cf. [8, Proposition 5.4]). *Let H be a Hilbert space. Then, for each $\nu \in \mathcal{P}^2(H)$, we have*

$$b(\nu) = \int_H y\, d\nu(y).$$

Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and N a CAT(0)-space. For an N-valued random variable $W\colon \Omega \to N$ satisfying $W_*\mathbb{P} \in \mathcal{P}^2(N)$, we define its *expectation* $\mathbb{E}_{\mathbb{P}}(f) \in N$ by the point $b(W_*\mathbb{P})$. By Lemma 2.2, in the case where N is a Hilbert space, this definition coincides with the classical one:

$$\mathbb{E}_{\mathbb{P}}(W) = \int_\Omega W(\omega)\, d\mathbb{P}(\omega).$$

The proof of the next lemma is easy, so we omit it.

Lemma 2.3. *Let N be a* CAT(0)*-space and $\nu \in \mathcal{P}^2(N)$. Then, we have*

$$d_N(b(\nu), \operatorname{Supp} \nu) \leq \operatorname{diam}(\operatorname{Supp} \nu).$$

Theorem 2.4 (Variance inequality, cf. [8, Proposition 4.4]). *Let N be a* CAT(0)*-space and $\nu \in \mathcal{P}^2(N)$. Then, for any $z \in N$, we have*

$$\int_N \{d_N(z, x)^2 - d_N(b(\nu), x)^2\} \, d\nu(x) \geq d_N(z, b(\nu))^2.$$

We now explain the inductive mean value introduced by Sturm in [8, Definition 4.6].

DEFINITION 2.5 (Inductive mean value). Given a sequence $(y_i)_{i=1}^{\mathbb{N}}$ of points in a uniquely geodesic space X, we define a new sequence of points $s_n \in X$, $n \in \mathbb{N}$, by induction as follows. We define $s_1 := y_1$ and $s_n := \gamma(1/n)$, where $\gamma\colon [0, 1] \to X$ is the geodesic connecting two points s_{n-1} and y_n. We denote the point s_n by $(1/n) \sum_{i=1,\dots,n}^{\to} y_i$ and call it the *inductive mean value* of the points $y_1, \dots, y_n$.

REMARK 2.6. (1) If the space X is a non-linear metric space, then the point $(1/n) \sum_{i=1,\dots,n}^{\to} y_i$ strongly depends on permutations of y_i as we see the following example. For $i = 1, 2, 3$, let $T_i := \{(i, r) \mid r \in [0, +\infty)\}$ be a copy of $[0, +\infty)$ equipped with the usual Euclidean distance function. The *tripod* T is the metric space obtained by gluing together all these spaces T_i, $i = 1, 2, 3$, at their origins with the intrinsic distance function. Let $y_1 := (1, 1)$, $y_2 := (2, 1)$, and $y_3 := (3, 1)$. Then, the inductive mean value of order y_1, y_2, y_3 is the point $(3, 1/3)$, whereas the one of order y_1, y_3, y_2 is the point $(2, 1/3)$.
(2) There are many other way to define a mean value of points $y_1, \dots, y_n$ in a CAT(0)-space (see [8, Remark 6.4]). For example, define a mean value as the barycenter of these points. Observe that this definition does not depend on order of the points (and so it is different from inductive mean value in general).

2.2. Invariants of mm-spaces and measures. In this subsection we define several invariants of mm-spaces and measures, which are needed for the proof of the main theorems.

An *mm-space* $X = (X, d_X, \mu_X)$ is a complete separable metric space (X, d_X) with a Borel probability measure μ_X. Let Y be a complete metric space and ν a finite Borel measure on Y having separable support with the total measure m. For any $\kappa > 0$, we define the *partial diameter* $\operatorname{diam}(\nu, m - \kappa)$ of ν as the infimum of the diameter of Y_0, where Y_0 runs over all Borel subsets of Y such that $\nu(Y_0) \geq m - \kappa$. Let X be an mm-space with $m_X := \mu_X(X)$ and Y a complete metric space. For any $\kappa > 0$, we define the *observable diameter* of X by

$$\operatorname{ObsDiam}_Y(X; -\kappa) := \sup\{\operatorname{diam}(f_*(\mu_X), m_X - \kappa) \mid f\colon X \to Y \text{ is a 1-Lipschitz map}\}.$$

The idea of the observable diameter comes from the quantum and statistical mechanics, i.e., we think of μ_X as a state on a configuration space X, and f is interpreted as an observable, i.e., an observation device giving us the visual (tomographic) image $f_*(\mu_X)$ on Y.

Let X be an mm-space. Given any two positive numbers κ_1 and κ_2, we define the *separation distance* $\mathrm{Sep}(X; \kappa_1, \kappa_2) = \mathrm{Sep}(\mu_X; \kappa_1, \kappa_2)$ of X as the supremum of the number $d_X(A_1, A_2)$, where A_1 and A_2 are Borel subsets of X such that $\mu_X(A_1) \geq \kappa_1$ and $\mu_X(A_2) \geq \kappa_2$, and we put

$$d_X(A_1, A_2) := \inf\{d_X(x_1, x_2) \mid x_1 \in A_1, x_2 \in A_2\}.$$

The next two lemmas are easy to prove.

Lemma 2.7 (cf. [6, Section $3\frac{1}{2}$.30]). *Let X and Y be two mm-spaces and $f : X \to Y$ be an α-Lipschitz map such that $f_*(\mu_X) = \mu_Y$. Then, for any $\kappa_1, \kappa_2 > 0$, we have*

$$\mathrm{Sep}(Y; \kappa_1, \kappa_2) \leq \alpha\, \mathrm{Sep}(X; \kappa_1, \kappa_2).$$

Lemma 2.8. *Given two positive numbers κ_1 and κ_2 such that $\kappa_1 \geq 1/2$ and $\kappa_2 > 1/2$, we have*

$$\mathrm{Sep}(\nu; \kappa_1, \kappa_2) = 0.$$

Lemma 2.9 (cf. [6, Section $3\frac{1}{2}$.33]). *Let X be an mm-space. Then, for any $\kappa, \kappa' > 0$ with $\kappa > \kappa'$, we have*

$$\mathrm{ObsDiam}_{\mathbb{R}}(X; -\kappa') \geq \mathrm{Sep}(X; \kappa, \kappa).$$

See also [5, Lemma 2.5] for the proof of the above lemma.

Let N be a CAT(0)-space and $\nu \in \mathcal{P}^2(N)$. Given any $\kappa > 0$, we define the *central radius* $\mathrm{CRad}(\nu, 1-\kappa)$ as the infimum of $\rho > 0$ such that $\nu(B_N(b(\nu), \rho)) \geq 1-\kappa$. Let X be an mm-space and N a CAT(0)-space such that $f_*(\mu_X) \in \mathcal{P}^2(N)$ for any 1-Lipschitz map $f : X \to N$. For any $\kappa > 0$, we define

$$\mathrm{ObsCRad}_N(X; -\kappa) := \sup\{\mathrm{CRad}(f_*(\mu_X), 1-\kappa) \mid f : X \to N \text{ is a 1-Lipschitz map}\},$$

and call it the *observable central radius* of X.

From the definition, we immediately obtain the following lemma.

Lemma 2.10 (cf. [6, Section $3\frac{1}{2}$.31]). *For any $\kappa > 0$, we have*

$$\mathrm{ObsDiam}_{\mathbb{R}}(X; -\kappa) \leq 2\, \mathrm{ObsCRad}_{\mathbb{R}}(X; -\kappa).$$

Observable diameters, separation distances, observable central radii are introduced by Gromov in [6, Chapter $3\frac{1}{2}$] to capture the theory of the Lévy–Milman concentration of 1-Lipschitz maps visually.

Given an mm-space X, we define the concentration function $\alpha_X\colon (0,+\infty)\to\mathbb{R}$ of X as the supremum of $\mu_X(X\setminus A_{+r})$, where A runs over all Borel subsets of X such that $\mu_X(A)\geq 1/2$ and A_{+r} is an open r-neighborhood of A. Concentration functions were introduced by D. Amir and V. Milman in [1].

3. Proof of the main theorem

Lemma 3.1. *Let N be a* CAT(0)*-space. Then, for any $n\in\mathbb{N}$, the map*

$$s_n\colon N^{\otimes n}\ni(x_1,x_2,\dots,x_n)\mapsto\frac{1}{n}\overset{\rightarrow}{\sum_{i=1,\dots,n}}x_i\in N$$

is $(1/n)$-Lipschitz with respect to the l^1-distance function on the product space $N^{\otimes n}$.

Proof. Assuming that the map s_{n-1} is $1/(n-1)$-Lipschitz, by Theorem 2.1, we have

$$\begin{aligned}
&d_N(s_n((x_i)_{i=1}^n),s_n((y_i)_{i=1}^n))\\
&\leq\left(1-\frac{1}{n}\right)d_N(s_{n-1}((x_i)_{i=1}^{n-1}),s_{n-1}((y_i)_{i=1}^{n-1}))+\frac{1}{n}\,d_N(x_n,y_n)\\
&\leq\left(1-\frac{1}{n}\right)\frac{1}{n-1}\sum_{i=1}^{n-1}d_N(x_i,y_i)+\frac{1}{n}d_N(x_n,y_n)\\
&=\frac{1}{n}\sum_{i=1}^{n}d_N(x_i,y_i).
\end{aligned}$$

This completes the proof. □

To prove Theorem 1.1, we need the following two theorems.

Theorem 3.2 (cf. [3, Lemma 5.5]). *Let ν be a Borel probability measure on an $\mathbb{R}$-tree such that $\nu\in\mathcal{P}^2(T)$. Then, there exists a 1-Lipschitz function $\varphi_\nu\colon T\to\mathbb{R}$ such that*

$$\begin{aligned}
\mathrm{CRad}(\nu,1-\kappa)\leq{}&\mathrm{CRad}((\varphi_\nu)_*(\nu),1-\kappa)+\mathrm{Sep}\left(\nu;\frac{1}{3},\frac{\kappa}{2}\right)\\
&+\mathrm{Sep}\left((\varphi_\nu)_*(\nu);\frac{1}{3},\frac{\kappa}{2}\right)+\mathrm{Sep}((\varphi_\nu)_*(\nu);1-\kappa,1-\kappa)
\end{aligned}$$

for any $\kappa>0$.

Theorem 3.3 (cf. [7, Corollary 1.17]). *Let $X = X_1 \otimes \cdots \otimes X_n$ be a product mm-space of mm-spaces X_i with finite diameter D_i, $i = 1, \ldots, n$, equipped with the product probability measure $\mu_X := \mu_{X_1} \otimes \cdots \otimes \mu_{X_n}$ and the l^1-distance function $d_{l^1} := \sum_{i=1}^n d_{X_i}$. Then, for any 1-Lipschitz function $f : X \to \mathbb{R}$ and any $r > 0$, we have*

$$\mu_X(\{x \in X \mid |f(x) - \mathbb{E}_{\mu_X}(f)| \geq r\}) \leq 2e^{-r^2/2D^2}, \tag{3.1}$$

where $D^2 := \sum_{i=1}^n D_i^2$. Moreover, we have

$$\alpha_X(r) \leq e^{-r^2/8D^2}. \tag{3.2}$$

Proof of Theorem 1.1. Let $s_n : T^{\otimes n} \to T$ be a map which sends every point in $T^{\otimes n}$ to its inductive mean value. Putting $\nu := (Y_1)_* \mathbb{P}$, we first prove the following.

Claim 3.4. *We have*

$$\nu^{\otimes n}(\{x \in T^{\otimes n} \mid d_T(s_n(x), \mathbb{E}_{\nu^{\otimes n}}(s_n)) \geq r\}) \leq 4e^{-nr^2/75D^2}.$$

Proof. Since the metric space (T, nd_T) is an $\mathbb{R}$-tree, by virtue of Theorem 3.2, there exists a 1-Lipschitz function $\varphi_n : (T, nd_T) \to \mathbb{R}$ such that

$$\begin{aligned}
&n \operatorname{CRad}((s_n)_*(\nu^{\otimes n}), 1-\kappa) \\
&\leq \operatorname{CRad}((\varphi_n \circ s_n)_*(\nu^{\otimes n}), 1-\kappa) + n \operatorname{Sep}\left((s_n)_*(\nu^{\otimes n}); \frac{1}{3}, \frac{\kappa}{2}\right) \\
&\quad + \operatorname{Sep}\left((\varphi_n \circ s_n)_*(\nu^{\otimes n}); \frac{1}{3}, \frac{\kappa}{2}\right) + \operatorname{Sep}((\varphi_n \circ s_n)_*(\nu^{\otimes n}); 1-\kappa, 1-\kappa)
\end{aligned}$$

for any $\kappa > 0$. By Lemma 3.1, the function $\varphi_n \circ s_n : (T^{\otimes n}, d_{l^1}) \to \mathbb{R}$ is 1-Lipschitz. Combining Lemma 2.7 with Lemmas 2.8, 2.9, and 2.10, for any $\kappa, \kappa' > 0$ such that $\kappa' < \kappa < 1/2$, we hence have

$$\begin{aligned}
&n \operatorname{CRad}((s_n)_*(\nu^{\otimes n}), 1-\kappa) \\
&\leq \operatorname{CRad}((\varphi_n \circ s_n)_*(\nu^{\otimes n}), 1-\kappa) + n \operatorname{Sep}\left((s_n)_*(\nu^{\otimes n}); \frac{1}{3}, \frac{\kappa}{2}\right) \\
&\quad + \operatorname{Sep}\left((\varphi_n \circ s_n)_*(\nu^{\otimes n}); \frac{1}{3}, \frac{\kappa}{2}\right) \\
&\leq \operatorname{ObsCRad}_{\mathbb{R}}((T^{\otimes n}, d_{l^1}, \nu^{\otimes n}); -\kappa) + 2 \operatorname{Sep}\left(\nu^{\otimes n}; \frac{\kappa}{2}, \frac{\kappa}{2}\right) \\
&\leq \operatorname{ObsCRad}_{\mathbb{R}}((T^{\otimes n}, d_{l^1}, \nu^{\otimes n}); -\kappa) \\
&\quad + 2 \operatorname{ObsDiam}_{\mathbb{R}}\left((T^{\otimes n}, d_{l^1}, \nu^{\otimes n}); -\frac{\kappa'}{2}\right) \\
&\leq 5 \operatorname{ObsCRad}_{\mathbb{R}}\left((T^{\otimes n}, d_{l^1}, \nu^{\otimes n}); -\frac{\kappa'}{2}\right).
\end{aligned}$$

According to the inequality (3.1), we thus get

$$n \,\mathrm{CRad}((s_n)_*(\nu^{\otimes n}), 1-\kappa) \leq 5D\sqrt{2n \log \frac{4}{\kappa'}}.$$

Letting $\kappa' \to \kappa$ yields that

$$\mathrm{CRad}((s_n)_*(\nu^{\otimes n}), 1-\kappa) \leq 5D\sqrt{\frac{2}{n} \log \frac{4}{\kappa}} \tag{3.3}$$

for any $\kappa \in (0, 1/2)$. Given $\kappa \geq 1/2$, taking an arbitrary $\kappa' \in (0, 1/2)$, we also estimate

$$\begin{aligned}\mathrm{CRad}((s_n)_*(\nu^{\otimes n}), 1-\kappa) &\leq \mathrm{CRad}((s_n)_*(\nu^{\otimes n}), 1-\kappa') \\ &\leq 5D\sqrt{\frac{2}{n} \log \frac{4}{\kappa'}} \\ &= 5D\frac{\sqrt{\log(4/\kappa')}}{\sqrt{\log(4/\kappa)}}\sqrt{\frac{2}{n} \log \frac{4}{\kappa}} \\ &\leq 5D\frac{\sqrt{\log(4/\kappa')}}{\sqrt{\log 4}}\sqrt{\frac{2}{n} \log \frac{4}{\kappa}}.\end{aligned}$$

Letting $\kappa' \to 1/2$, we hence get

$$\mathrm{CRad}((s_n)_*(\nu^{\otimes n}), 1-\kappa) \leq 5D\sqrt{\frac{3}{n} \log \frac{4}{\kappa}}. \tag{3.4}$$

The above two inequalities (3.3) and (3.4) imply the claim. □

Put $a_n := d_T(\mathbb{E}_{\nu^{\otimes n}}(s_n), b(\nu))$. By Sturm's inequality (1.1), we have

$$\int_{T^{\otimes n}} d_T(s_n(x), b(\nu))^2 \, d\nu^{\otimes n}(x) \leq \frac{1}{n}\int_T d_T(x, b(\nu))^2 \, d\nu(x).$$

Lemma 2.3 together with Theorem 2.4 thus implies that

$$a_n^2 \leq \int_{T^{\otimes n}} d_T(s_n(x), b(\nu))^2 d\nu^{\otimes n}(x) \leq \frac{1}{n}\int_T d_T(x, b(\nu))^2 d\nu(x) \leq \frac{4D^2}{n}.$$

For any $r > a_n$, by using Claim 3.4, we therefore obtain

$$\begin{aligned}&\mathbb{P}\left(\left\{\omega \in \Omega \;\middle|\; d_T\left(\frac{1}{n}\overrightarrow{\sum_{i=1,\dots,n}} Y_i(\omega), \mathbb{E}_{\mathbb{P}}(Y_1)\right) \geq r\right\}\right) \\ &= \nu^{\otimes n}(\{x \in T^{\otimes n} \mid d_T(s_n(x), b(\nu)) \geq r\}) \\ &\leq \nu^{\otimes n}(\{x \in T^{\otimes n} \mid d_T(s_n(x), \mathbb{E}_{\nu^{\otimes n}}(s_n)) \geq r - a_n\}) \\ &\leq 4e^{-n(r-a_n)^2/75D^2} \\ &\leq 4e^{na_n^2/75D^2} e^{-nr^2/150D^2} \\ &\leq 4e^{4/75} e^{-nr^2/150D^2}.\end{aligned}$$

If $r \leq a_n$, then we have

$$\mathbb{P}\left(\left\{\omega \in \Omega \;\middle|\; d_T\left(\frac{1}{n}\overrightarrow{\sum_{i=1,\dots,n}} Y_i(\omega),\, \mathbb{E}_{\mathbb{P}}(Y_1)\right) \geq r\right\}\right)$$
$$\leq e^{na_n^2/150D^2}e^{-na_n^2/150D^2} < e^{2/75}e^{-nr^2/150D^2} < 4e^{4/75}e^{-nr^2/150D^2}.$$

Combining these two inequalities completes the proof of the theorem. $\square$

Theorem 1.2 follows from the same proof of Theorem 1.1 together with the inequality (3.2) and the following theorem. We shall consider an mm-space satisfying

$$\alpha_X(r) \leq C_X e^{-c_X r^2} \tag{3.5}$$

for some positive constants $c_X, C_X > 0$ and any $r > 0$. For such an mm-space X and $m \in \mathbb{N}$, we put

$$A_{m,X} := 1 + \frac{\sqrt{\pi}e^{(m+1)/(4m-2)}}{2}\max\{e^{(\pi C_X)^2/2}, 2C_X e^{(\pi C_X)^2}\}$$

and

$$\tilde{A}_{m,X} := 1 + \sqrt{\pi}C_X e^{(m+1)/(4m-2)}.$$

Theorem 3.5 (cf. [4, Theorem 1.1]). *Let an mm-space X satisfies* (3.5), *N be an m-dimensional Hadamard manifold, and $f\colon X \to N$ a* 1*-Lipschitz map. Then, for any $r > 0$, we have*

$$\mu_X(\{x \in X \mid d_N(f(x), \mathbb{E}_{\mu_X}(f)) \geq r\}) \leq \min\{A_{m,X}e^{-(c_X/(8m))r^2}, \tilde{A}_{m,X}e^{-(c_X/(16m))r^2}\}.$$

ACKNOWLEDGEMENTS. The author would like to thank to Professors Kazuhiro Kuwae and Daehong Kim for motivating this work. This work was partially supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists.

References

[1] D. Amir and V.D. Milman: *Unconditional and symmetric sets in n-dimensional normed spaces*, Israel J. Math. **37** (1980), 3–20.
[2] A. Es-Sahib and H. Heinich: *Barycentre canonique pour un espace métrique à courbure négative*; in Séminaire de Probabilités XXXIII, Lecture Notes in Math. **1709**, Springer, Berlin, 1999, 355–370.
[3] K. Funano: *Central and L^p-concentration of* 1*-Lipschitz maps into $\mathbb{R}$-trees*, J. Math. Soc. Japan **61** (2009), 483–506.

[4] K. Funano: *Exponential and Gaussian concentration of* 1-*Lipschitz maps*, to appear in Manuscripta Math.

[5] K. Funano: *Observable concentration of mm-spaces into spaces with doubling measures*, Geom. Dedicata **127** (2007), 49–56.

[6] M. Gromov: Metric Structures for Riemannian and Non-Riemannian Spaces, based on the 1981 French original, with appendices by M. Katz, P. Pansu and S. Semmes, translated from the French by Sean Michael Bates, Progress in Mathematics **152**, Birkhäuser Boston, Boston, MA, 1999.

[7] M. Ledoux: The Concentration of Measure Phenomenon, Mathematical Surveys and Monographs **89**, Amer. Math. Soc., Providence, RI, 2001.

[8] K.-T. Sturm: *Probability measures on metric spaces of nonpositive curvature*; in Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces (Paris, 2002), Contemp. Math. **338**, Amer. Math. Soc., Providence, RI, 2003, 357–390.

Department of Mathematics and Engineering
Graduate School of Science and Technology
Kumamoto university
Kumamoto 860–8500
Japan
e-mail: yahoonitaikou@gmail.com