Convergence theorems for barycentric maps

Fumio Hiai

Tohoku University

2018, July (at Będlewo)

Joint work[1] with Yongdo Lim

[1] F.H. and Y. Lim, Convergence theorems for contractive barycentric maps, arXiv:1805.08558 [math.PR].

Idea

- When (M, d) is a global NPC = CAT(0) space, martingale convergence, strong law of large numbers and ergodic theorem were developed for M-valued random variables by Es-Sahib and Heinich, Sturm, Austin, Navas, ...
- By using the disintegration theorem, we develop those stochastic convergence theorems when (M, d) is a general complete metric space with a contractive barycentric map β.
- E.g., $M = \mathbb{P}(\mathcal{H})$ is the positive invertible operators on a Hilbert space $\mathcal{H}$, $d = d_T$ is the Thompson metric, and β is the Cartan barycenter (Karcher mean).

Plan

- Conditional expectations
- Martingale convergence theorem
- Ergodic theorem
- Large deviation principle

Preliminaries

- (M, d) is a complete metric space with the Borel σ-algebra $\mathcal{B}(M)$.
- $\mathcal{P}(M)$ is the set of probability measures on $\mathcal{B}(M)$ with full support.
- For $1 \le p < \infty$, $\mathcal{P}^p(M)$ is the set of $\mu \in \mathcal{P}(M)$ such that $\int_M d^p(x, y)\, d\mu(y) < \infty$ for some (hence, all) $x \in M$.

$$\mathcal{P}^1(M) \supset \mathcal{P}^p(M) \supset \mathcal{P}^q(M), \qquad 1 < p < q < \infty.$$

- For $1 \le p < \infty$, the p-Wasserstein distance is

$$d_p^W(\mu, \nu) := \left[\inf_{\pi \in \Pi(\mu,\nu)} \int_{M \times M} d^p(x, y)\, d\pi(x, y) \right]^{1/p}, \quad \mu, \nu \in \mathcal{P}(M),$$

where $\Pi(\mu, \nu)$ is the set of $\pi \in \mathcal{P}(M \times M)$ whose marginals are μ, ν.

$$d_1^W \le d_p^W \le d_q^W, \qquad 1 < p < q < \infty,$$

and $(\mathcal{P}^p(M), d_p^W)$ is a complete metric space.

- $(\Omega, \mathcal{A}, \mathbf{P})$ is a probability space.
- For $1 \le p < \infty$, $L^p(\Omega; M) = L^p(\Omega, \mathcal{A}, \mathbf{P}; M)$ is the set of strongly measurable functions $f : \Omega \to M$ such that $\int_\Omega d^p(x, f(\omega))\, d\mathbf{P}(\omega) < \infty$ for some (hence, all) $x \in M$.

$$L^1(\Omega; M) \supset L^p(\Omega; M) \supset L^q(\Omega; M) \qquad 1 < p < q < \infty.$$

Lemma

Let $1 \le p < \infty$.

- $L^p(\Omega; M)$ is a complete metric space with the L^p-distance

$$\mathbf{d}_p(\varphi, \psi) := \left[\int_\Omega d^p(\varphi(\omega), \psi(\omega))\, d\mathbf{P}(\omega)\right]^{1/p}.$$

- If $\varphi \in L^p(\Omega; M)$, then the push-forward measure $\varphi_*\mathbf{P} \in \mathcal{P}^p(M)$.
- If $\varphi, \psi \in L^p(\Omega; M)$, then $d_p^W(\varphi_*\mathbf{P}, \psi_*\mathbf{P}) \le \mathbf{d}_p(\varphi, \psi)$.

Conditional expectations

Let $1 \le p < \infty$ be fixed, and assume that $\beta : \mathcal{P}^p(M) \to M$ is a p-contractive barycentric map, i.e., $\beta(\delta_x) = x$ for all $x \in M$ and

$$d(\beta(\mu), \beta(\nu)) \le d_p^W(\mu, \nu), \qquad \mu, \nu \in \mathcal{P}^p(M).$$

Definition

The β-expectation $E^\beta(\varphi)$ of $\varphi \in L^p(\Omega; M)$ is defined by

$$E^\beta(\varphi) := \beta(\varphi_* \mathrm{P}) \quad \in M.$$

Proposition

- $d(E^\beta(\varphi), E^\beta(\psi)) \le \mathbf{d}_p(\varphi, \psi)$ for $\varphi, \psi \in L^p(\Omega; M)$.
- $E^\beta(\mathbf{1}_\Omega x) = x$ for $x \in M$.

Next, assume that $(\Omega, \mathcal{A})$ is a standard Borel space, i.e., isomorphic to $(X, \mathcal{B}(X))$ of a Polish space X. Let $\mathcal{B}$ be a sub-σ-algebra of $\mathcal{A}$. Then there exists a disintegration $(\mathbf{P}_\omega)_{\omega\in\Omega}$ with respect to $\mathcal{B}$, a family of probability measures on $(\Omega, \mathcal{A})$, such that for every $A \in \mathcal{A}$,

(i) $\omega \in \Omega \mapsto \mathbf{P}_\omega(A)$ is $\mathcal{B}$-measurable,

(ii) $\omega \mapsto \mathbf{P}_\omega(A)$ is a conditional expectation $E_{\mathcal{B}}(\mathbf{1}_A)$ of $\mathbf{1}_A$ with respect to $\mathcal{B}$,

Such a family $(\mathbf{P}_\omega)_{\omega\in\Omega}$ is unique up to a $\mathbf{P}$-null set, and moreover

(iii) for every $f \in L^1(\Omega; \mathbb{R})$, $f \in L^1(\Omega, \mathcal{A}, \mathbf{P}_\omega; \mathbb{R})$ for $\mathbf{P}$-a.e. ω and $\omega \mapsto \int_\Omega f(\tau)\,d\mathbf{P}_\omega(\tau)$ is a conditional expectation $E_{\mathcal{B}}(f)$ of f with respect to $\mathcal{B}$. In particular,

$$\int_\Omega f\,d\mathbf{P} = \int_\Omega \left[\int_\Omega f(\tau)\,d\mathbf{P}_\omega(\tau)\right] d\mathbf{P}(\omega).$$

Definition

The β-conditional expectation $E_{\mathcal{B}}^{\beta}(\varphi)$ of $\varphi \in L^p(\Omega; M)$ is defined by

$$E_{\mathcal{B}}^{\beta}(\varphi) := \beta(\varphi_* \mathbf{P}_\omega), \qquad \omega \in \Omega.$$

Theorem

Let $\varphi, \psi \in L^p(\Omega; M)$.

(1) $E_{\mathcal{B}}^{\beta}(\varphi) \in L^p(\Omega, \mathcal{B}, \mathbf{P}; M)$.

(2) $\mathbf{d}_p(E_{\mathcal{B}}^{\beta}(\varphi), E_{\mathcal{B}}^{\beta}(\psi)) \le \mathbf{d}_p(\varphi, \psi)$.

(3) $\varphi \in L^p(\Omega, \mathcal{B}, \mathbf{P}; M)$ if and only if $E_{\mathcal{B}}^{\beta}(\varphi) = \varphi$. Hence $E_{\mathcal{B}}^{\beta}(E_{\mathcal{B}}^{\beta}(\varphi)) = E_{\mathcal{B}}^{\beta}(\varphi)$.

(4) When $\mathcal{B} = \{\emptyset, \Omega\}$, $E_{\mathcal{B}}^{\beta}(\varphi) = E^{\beta}(\varphi)$.

When (M, d) is a global NPC space or CAT(0) space, (i.e., for any $x_0, x_1 \in M$ there exists a $y \in M$ such that

$$d^2(y, z) \le \frac{d^2(x_0, z) + d^2(x_1, z)}{2} - \frac{d^2(x_0, x_1)}{4} \quad \text{for all } z \in M),$$

the canonical barycentric map λ on $\mathcal{P}^1(M)$ is

$$\lambda(\mu) := \operatorname*{arg\,min}_{z \in M} \int_M [d^2(z, x) - d^2(y, x)]\, d\mu(x), \quad \mu \in \mathcal{P}^1(M),$$

independently of the choice of $y \in M$.

Sturm's[2] definition in the case of a global NPC space is

$$\mathbf{E}_{\mathcal{B}}(\varphi) := \operatorname*{arg\,min}_{\psi \in L^2(\Omega, \mathcal{B}, \mathrm{P}; M)} \mathrm{d}_2(\varphi, \psi)$$

for $\varphi \in L^2(\Omega; M)$, and $\mathbf{E}_{\mathcal{B}}$ extends continuously to $L^1(\Omega; M)$.

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

Theorem

Assume that $(\Omega, \mathcal{A})$ is a standard Borel space and (M, d) is a global NPC space. Then for every $p \in [1, \infty)$ and $\varphi \in L^p(\Omega; M)$,

$$\mathbf{E}_{\mathcal{B}}(\varphi) = E_{\mathcal{B}}^{\lambda}(\varphi).$$

Remark

Unlike the usual conditional expectation, the β-conditional expectation is not associative in general, that is, for sub-σ-algebras $\mathcal{C} \subset \mathcal{B} \subset \mathcal{A}$,

$$E_{\mathcal{C}}^{\beta}(E_{\mathcal{B}}^{\beta}(\varphi)) \neq E_{\mathcal{C}}^{\beta}(\varphi).$$

Martingale convergence theorem

Let $(\Omega, \mathcal{A}, P)$ be a standard Borel probability space, and $\{\mathcal{B}_n\}_{n=1}^{\infty}$ be a sequence of sub-σ-algebras of $\mathcal{A}$ such that

$$\mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots \quad \text{or} \quad \mathcal{B}_1 \supset \mathcal{B}_2 \supset \cdots .$$

Let $\mathcal{B}_\infty$ be the sub-σ-algebra generated by $\bigcup_{n=1}^{\infty} \mathcal{B}_n$ or $\mathcal{B}_\infty := \bigcap_{n=1}^{\infty} \mathcal{B}_n$.

Theorem

Assume that $(\Omega, \mathcal{A}, P)$ and $\{\mathcal{B}_n\}_{n=1}^{\infty}$ are as stated above. Let $\beta : \mathcal{P}^p(M) \to M$ be as before. Then for every $\varphi \in L^p(\Omega; M)$, as $n \to \infty$,

$$\mathrm{d}_p(E^{\beta}_{\mathcal{B}_n}(\varphi), E^{\beta}_{\mathcal{B}_\infty}(\varphi)) \longrightarrow 0,$$

$$d(E^{\beta}_{\mathcal{B}_n}(\varphi)(\omega), E^{\beta}_{\mathcal{B}_\infty}(\varphi)(\omega)) \longrightarrow 0 \ \text{a.e.}$$

Assume that $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots$. Since $E^\beta_{\mathcal{B}_m}(E^\beta_{\mathcal{B}_n}(\varphi)) = E^\beta_{\mathcal{B}_m}(\varphi)$ $(m < n)$ does not hold, we follow Sturm's[2] idea to define martingales of M-valued random variables.

Definition

For $\varphi \in L^p(\Omega; M)$ and $k \geq 1$, we can define

$$\begin{aligned} E^\beta[\varphi \| (\mathcal{B}_n)_{n \geq k}] &:= \lim_{m \to \infty} E^\beta_{\mathcal{B}_k} \circ \cdots \circ E^\beta_{\mathcal{B}_m}(\varphi) \\ &= \lim_{m \to \infty} E^\beta_{\mathcal{B}_k} \circ \cdots \circ E^\beta_{\mathcal{B}_m}(E^\beta_{\mathcal{B}_\infty} \varphi)) \end{aligned}$$

in metric $\mathbf{d}_p$. Call $E^\beta[\varphi \| (\mathcal{B}_n)_{n \geq k}]$ the filtered β-conditional expectation with respect to $(\mathcal{B}_n)_{n \geq k}$.

Proposition

Let $\varphi, \psi \in L^p(\Omega; M)$.

(1) $E^\beta[\varphi \| (\mathcal{B}_n)_{n\ge k}] \in L^p(\Omega, \mathcal{B}_k, \mathbf{P}; M)$ for all $k \ge 1$.

(2) For every $k \ge 1$, $\varphi \in L^p(\Omega, \mathcal{B}_k, \mathbf{P}; M)$ if and only if $E^\beta[\varphi \| (\mathcal{B}_n)_{n\ge k}] = \varphi$.

(3) $\mathbf{d}_p(E^\beta[\varphi \| (\mathcal{B}_n)_{n\ge k}], E^\beta[\psi \| (\mathcal{B}_n)_{n\ge k}]) \le \mathbf{d}_p(\varphi, \psi)$ for all $k \ge 1$.

(4) Associativity: For every $l \ge k \ge 1$,

$$E^\beta[E^\beta[\varphi \| (\mathcal{B}_n)_{n\ge l}] \| (\mathcal{B}_n)_{n\ge k}] = E^\beta[\varphi \| (\mathcal{B}_n)_{n\ge k}].$$

Definition

A sequence $\{\varphi_k\}_{k=1}^\infty$ in $L^p(\Omega; M)$ is called a filtered β-martingale with respect to $\{\mathcal{B}_n\}_{n=1}^\infty$ if $\varphi_k \in L^p(\Omega, \mathcal{B}_k, \mathbf{P}; M)$ for every $k \ge 1$ and

$$E^\beta[\varphi_{k+1} \| (\mathcal{B}_n)_{n\ge k}] = \varphi_k, \qquad k \ge 1,$$

equivalently, $E^\beta[\varphi_l \| (\mathcal{B}_n)_{n\ge k}] = \varphi_k$ for all $l \ge k \ge 1$.

Theorem

Let $\{\varphi_k\}_{k=1}^{\infty}$ be a filtered β-martingale with respect to $\{\mathcal{B}_n\}$. Then the following are equivalent:

(i) there exists a $\varphi \in L^p(\Omega; M)$ such that $\varphi_k = E^\beta[\varphi \| (\mathcal{B}_n)_{n \geq k}]$ for all $k \geq 1$;

(ii) φ_k converges to some $\varphi_\infty \in L^p(\Omega, \mathcal{B}_\infty, \mathbf{P}; M)$ in metric $\mathbf{d}_p$ as $k \to \infty$.

Remark

Assume that (M, d) is a global NPC space (or more generally, a complete length space) and it is locally compact. It is known [2] that if $\{\varphi_k\}$ in $L^p(\Omega; M)$ is a filtered martingale and $\sup_k \mathbf{d}_p(z, \varphi_k) < \infty$ for some $z \in M$, then there exists a $\mathcal{B}_\infty$-measurable function $\varphi_\infty : \Omega \to M$ such that $\varphi_k(\omega) \to \varphi_\infty(\omega)$ $\mathbf{P}$-a.e. But it is unknown that this holds in our general setting.

Ergodic theorem

Let T be a $\mathbf{P}$-preserving measurable transformation on $(\Omega, \mathcal{A}, \mathbf{P})$. Let $\beta : \mathcal{P}^p(M) \to M$ be as before. For each $\varphi \in L^p(\Omega; M)$, consider the empirical measures (random probability measures) of φ

$$\mu_n^\varphi(\omega) := \frac{1}{n} \sum_{k=0}^{n-1} \delta_{\varphi(T^k\omega)}, \qquad n \in \mathbb{N},$$

i.e., for Borel sets $B \subset M$,

$$\mu_n^\varphi(\omega)(B) = \frac{\#\{k \in \{0, 1, \dots, n-1\} : \varphi(T^k\omega) \in B\}}{n},$$

and consider the sequence of M-valued functions $\beta(\mu_n^\varphi) : \omega \in \Omega \mapsto \beta(\mu_n^\varphi(\omega)) \in M$ for $n \in \mathbb{N}$.

Lemma

For every $\varphi, \psi \in L^p(\Omega; M)$, $\beta(\mu_n^\varphi) \in L^p(\Omega; M)$ and

$$\mathbf{d}_p(\beta(\mu_n^\varphi), \beta(\mu_n^\psi)) \le \mathbf{d}_p(\varphi, \psi), \qquad n \in \mathbb{N}.$$

Extending the ergodic theorems in [3] [4],

---

[3] T. Austin, A $CAT(0)$-valued pointwise ergodic theorem, *J. Topol. Anal.* **3** (2011), 145–152.

[4] A. Navas, An L^1 ergodic theorem with values in a non-positively curved space via a canonical barycenter map, *Ergod. Th. Dynam. Sys.*, **33** (2013), 609–623.

Theorem

There exists a map

$$\Gamma : L^p(\Omega; M) \longrightarrow \{\varphi \in L^p(\Omega; M) : \varphi \circ T = \varphi\}$$

such that for every $\varphi, \psi \in L^p(\Omega; M)$,

(i) $d(\beta(\mu_n^\varphi(\omega)), \Gamma(\varphi)(\omega)) \to 0$ a.e. as $n \to \infty$,

(ii) $\mathbf{d}_p(\beta(\mu_n^\varphi), \Gamma(\varphi)) \to 0$ as $n \to \infty$,

(iii) $\mathbf{d}_p(\Gamma(\varphi), \Gamma(\psi)) \le \mathbf{d}_p(\varphi, \psi)$.

Furthermore, if T is ergodic, then $\Gamma(\varphi)$ is a constant $E^\beta(\varphi)$, the β-expectation of φ.

Theorem

Assume that $(\Omega, \mathcal{A})$ is a standard Borel space, and let $\mathcal{I} := \{A \in \mathcal{A} : T^{-1}A = A\}$, the sub-$\sigma$-algebra consisting of T-invariant sets. Then for every $\varphi \in L^p(\Omega; M)$,

$$\Gamma(\varphi) = E_{\mathcal{I}}^{\beta}(\varphi),$$

the β-conditional expectation of φ with respect to $\mathcal{I}$.

Example

Let $\mathbb{P} = \mathbb{P}(\mathcal{H})$ be the set of positive invertible operators on a Hilbert space $\mathcal{H}$, with the Thompson metric $d_T(A, B) := \|\log A^{-1/2}BA^{-1/2}\|$. The Karcher barycenter $G : \mathcal{P}^1(\mathbb{P}) \to \mathbb{P}$ determined by

$$X = G\left(\frac{1}{n}\sum_{i=1}^{n}\delta_{A_i}\right) \iff \sum_{i=1}^{n}\log(X^{-1/2}A_iX^{-1/2}) = 0$$

is a contractive barycentric map and monotone for the Löwner order $A \le B$. For $\varphi \in L^p(\Omega; \mathbb{P})$ with $1 \le p < \infty$, note that

$$G(\mu_n^\varphi(\omega)) = G\left(\frac{1}{n}\sum_{k=0}^{n-1}\delta_{\varphi(T^k\omega)}\right) = G(\varphi(\omega), \varphi(T\omega), \dots, \varphi(T^{n-1}\omega)),$$

which is the Karcher mean of $\varphi(T^k\omega)$ $(0 \le k \le n-1)$. We have

$$\lim_{n\to\infty} G(\varphi, \varphi \circ T, \dots, \varphi \circ T^{n-1}) = \Gamma(\varphi) \ \text{ a.e. and in metric } \mathbf{d}_p.$$

When $(\Omega, \mathcal{A})$ is a standard Borel space, $\Gamma(\varphi) = E_{\mathcal{I}}^G(\varphi)$. Moreover, Γ is monotone and $\Gamma(\varphi^{-1}) = \Gamma(\varphi)^{-1}$ follows from $G(\mu^{-1}) = G(\mu)^{-1}$.

Large deviation principle

A sequence (μ_n) of Borel probability measures on a metric space $\mathcal{X}$ is said to satisfy the LDP with a rate function I if for every $\Gamma \in \mathcal{B}(\mathcal{X})$,

$$-\inf_{x\in\Gamma^\circ} I(x) \le \liminf_{n\to\infty} \frac{1}{n} \log \mu_n(\Gamma) \le \limsup_{n\to\infty} \frac{1}{n} \log \mu_n(\Gamma) \le -\inf_{x\in\overline{\Gamma}} I(x),$$

where Γ° and $\overline{\Gamma}$ denote the interior and the closure of Γ.

- $(\Omega, \mathcal{A}, \mathbf{P})$ is a probability space.
- Σ is a Polish space.
- $\mathcal{P}(\Sigma)$ becomes a Polish space with the weak topology.
- $\mathbf{X} = (X_1, X_2, \dots)$ is a sequence of i.i.d. random variables $X_n : \Omega \to \Sigma$, with distribution $\mu_0 \in \mathcal{P}(\Sigma)$.
- The empirical measure of $\mathbf{X}$ is

$$\mu_n^{\mathbf{X}}(\omega) := \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i(\omega)}, \qquad n \in \mathbb{N}.$$

- The distribution $\widehat{\mu}_n$ of $\mu_n^X : \Omega \to \mathcal{P}(\Sigma)$ is

$$\widehat{\mu}_n(\Gamma) := P(\mu_n^X \in \Gamma) = \mu_0^{\times n}\left(\left\{(x_1,\dots,x_n) \in \Sigma^n : \frac{1}{n}\sum_{i=1}^n \delta_{x_i} \in \Gamma\right\}\right)$$

for Borel sets $\Gamma \subset \mathcal{P}(\Sigma)$.

Sanov theorem

A sequence of the distributions $\widehat{\mu}_n$ of the empirical measures μ_n^X satisfies the LDP with the relative entropy functional $S(\cdot\|\mu_0)$ as the good rate function, where the relative entropy is

$$S(\mu\|\mu_0) := \begin{cases} \int_\Sigma \log \frac{d\mu}{d\mu_0}\, d\mu & \text{if } \mu \ll \mu_0 \text{ (absolutely continuous)}, \\ \infty & \text{otherwise.} \end{cases}$$

- Let $\mathbf{X} = (X_1, X_2, \dots)$ be a sequence of i.i.d. M-valued random variables such that the distribution μ_0 of X_n is in $\mathcal{P}^\infty(M)$, i.e., $X_n \in L^\infty(\Omega; M)$. Then there is a bounded Polish subset Σ of M such that X_n's are Σ-valued random variables.
- Let $\beta : \mathcal{P}^p(M) \to M$ be as before. Then $\mathcal{P}(\Sigma) \subset \mathcal{P}^p(M)$ and $\beta|_{\mathcal{P}(\Sigma)} : \mathcal{P}(\Sigma) \to M$ is continuous in the weak topology.
- The push-forward of $\widehat{\mu}_n$ by $\beta|_{\mathcal{P}(\Sigma)}$ is the distribution of $\beta(\mu_n^{\mathbf{X}})$, i.e., for every $\Gamma \in \mathcal{B}(M)$,

$$\begin{aligned}\widehat{\mu}_n(\{\mu \in \mathcal{P}(\Sigma) : \beta(\mu) \in \Gamma\}) &= \mathbf{P}(\beta(\mu_n^{\mathbf{X}}) \in \Gamma) \\ &= \mathbf{P}\Bigl(\Bigl\{\omega \in \Omega : \beta\Bigl(\frac{1}{n}\sum_{i=1}^n \delta_{X_i(\omega)}\Bigr) \in \Gamma\Bigr\}\Bigr).\end{aligned}$$

Applying the contraction principle for LDP to the Sanov theorem and the continuous map $\beta : \mathcal{P}(\Sigma) \to M$,

Theorem

Let $X_1, X_2, \dots$ be a sequence of i.i.d. M-valued random variables having the distribution $\mu_0 \in \mathcal{P}^\infty(M)$. Then a sequence of the distributions of the β-values $\beta(\mu_n^{\mathrm{X}}) = \beta(\frac{1}{n}\sum_{i=1}^n \delta_{X_i})$ satisfies the LDP with the good rate function

$$I(x) := \inf\{S(\mu\|\mu_0) : \mu \in \mathcal{P}(\Sigma),\ x = \beta(\mu)\}, \qquad x \in M.$$

That is, for every $\Gamma \in \mathcal{B}(M)$,

$$\begin{aligned} -\inf_{x\in\Gamma^\circ} I(x) &\le \liminf_{n\to\infty} \frac{1}{n} \log \mathrm{P}(\beta(\mu_n^{\mathrm{X}}) \in \Gamma) \\ &\le \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{P}(\beta(\mu_n^{\mathrm{X}}) \in \Gamma) \le -\inf_{x\in\overline{\Gamma}} I(x). \end{aligned}$$

The above LDP implies the strong law of large numbers for X_n.[5]

Corollary

Let $X_1, X_2, \dots$ be a sequence of i.i.d. M-valued random variables having the distribution $\mu_0 \in \mathcal{P}^\infty(M)$. Then

$$\beta\left(\frac{1}{n}\sum_{i=1}^{n} \delta_{X_i(\omega)}\right) \longrightarrow \beta(\mu_0) \ \text{a.e.} \quad \text{as } n \to \infty.$$

[5] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature, Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), 357–390, *Contemporary Mathematics* **338**, Amer. Math. Soc., Providence, RI, 2003.

Thank you for your attention!