Convergence of Continuous Stochastic Processes on Compact Metric Spaces Converging in the Lipschitz Distance

Kohei Suzuki

Received: 12 November 2015 / Accepted: 5 January 2018 / Published online: 22 January 2018 © Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract We introduce a new distance, a Lipschitz–Prokhorov distance $d_{LP}$, on the set $\mathcal{P}M$ of isomorphism classes of pairs $(X, P)$ where $X$ is a compact metric space and $P$ is the law of a continuous stochastic process on $X$. We show that $(\mathcal{P}M, d_{LP})$ is a complete metric space. For Markov processes on Riemannian manifolds, we study relative compactness and convergence.

Keywords Weak convergence · Lipschitz convergence · Markov processes · Riemannian manifolds

Mathematics Subject Classification (2010) Primary 60F17 · Secondary 53C23

1 Introduction

The motivation of this paper is to study a convergence of continuous stochastic processes on varying compact metric spaces. One of our motivating examples is the following: Let $(M_n, g_n)$ be a sequence of compact Riemannian manifolds converging to a compact $C^{1,\alpha}$-Riemannian manifold $(M, g)$ in the Lipschitz distance (see e.g., [10] for Lipschitz distance and [4, Fig. 12.37] for an example of $C^{1,\alpha}$-Riemannian manifolds).

(Q) Do the Brownian motions $B_n$ on $M_n$ converge to the Brownian motion on the limit Riemannian manifold $M$?

Note that our state spaces are not fixed. If state spaces are fixed, we have the weak convergence of the laws of stochastic processes as a standard notion of convergences of stochastic processes.
processes. When state spaces are not fixed, however, we should take care about in which sense we consider a convergence of stochastic processes.

Our aim is to introduce a notion of convergence of stochastic processes on varying metric spaces and to answer (Q) with respect to this new notion. There has recently been a lot of activity in the broad area of convergence of measured metric spaces; however, for convergences of stochastic processes, we need to take care about path regularity of stochastic processes. This is because the spaces of continuous paths, or càdlàg paths (i.e., right continuous paths with left-limits) on Polish spaces become Polish with suitable distances such as the uniform distance, or the Skorokhod distance. This enables us to use standard arguments about probability measures on Polish spaces. However, if we consider varying state spaces $X_n$ to the limit space $X$ (e.g., Gromov–Hausdorff convergence), approximation maps $f_n : X_n \to X$ are not necessarily continuous (only measurable), and push-forwarded stochastic processes are not necessarily continuous, or right-continuous with left-limits. This makes it difficult for us to deal with convergences of processes because essentially we need to consider the space of measurable paths, which cannot come down to the usual frameworks of Polish spaces such as the space of continuous paths or the space of càdlàg paths.

We should remark that in the case of the Gromov–Hausdorff distance, although approximation maps are only measurable, we can embed all the spaces of sequences into one common metric space by isometric embeddings, and thus we can consider the weak convergence of the push-forwarded stochastic processes on this common metric space (see e.g., [3, 20]). However, this argument only gives a sequence-wise topology, which means that this common metric space depends on a given sequence of spaces and we can give a topology only for a given sequence $(X_i, P_i) (i \in \mathbb{N})$, but not for the set of all pairs $(X, P)$ where $X_i$ and $X$ are metric spaces as state spaces and $P_i$ and $P$ are probability measures on continuous path spaces on $X_i$ and $X$. To equip a topology or a metric for the set of all pairs $(X, P)$, we still have difficulty with respect to the regularity of approximation maps. See also Remark 2.8 on this point.

In this paper, we consider a convergence of continuous stochastic processes on compact metric spaces converging in the Lipschitz distance. When state spaces $X_n$ are convergent to $X$ in the Lipschitz distance, there exists a bi-Lipschitz approximation map $f_n : X_n \to X$. Therefore, under the circumstance where state spaces are varying, we can still consider the weak convergence of push-forwarded continuous stochastic processes on the limit space $X$ within the usual framework of the space of continuous path spaces on $X$. Moreover, we construct a complete metric Lipschitz–Prokhorov distance $d_{LP}$ which gives a topology to the whole space of pairs $(X, P)$. We also study several topological properties of this new metric such as compactness of some subsets. As an application, we study the convergence in $d_{LP}$ of the Brownian motions, or diffusions on Riemannian manifolds converging in the Lipschitz distance. We answer (Q) yes with respect to $d_{LP}$ under some geometric conditions of Riemannian manifolds.

To be more precise, the main object of this paper is a pair $(X, P)$ where $X$ is a compact metric space and $P$ is the law of a continuous stochastic process on $X$. Let $\mathcal{PM}$ be the set of all pairs $(X, P)$ modulo by an isomorphism relation (defined in Section 2). We will define (in Section 2) a new distance on $\mathcal{PM}$, which we will call the Lipschitz–Prokhorov distance $d_{LP}$, as a kind of mixture of the Lipschitz distance and the Prokhorov distance. The Lipschitz distance is a distance on the set of isometry classes of metric spaces, which was first introduced by Gromov (see e.g., [10]).

We summarize our results as follows:

(A) $(\mathcal{PM}, d_{LP})$ is a complete metric space (Theorems 2.6 and 2.9);
(B) Sufficient conditions for relative compactness in \((\mathcal{PM}, d_{LP})\) in the case of Riemannian manifolds (Theorem 3.9).

(C) Sequences in a relatively compact set are convergent if the corresponding Dirichlet forms of Markov processes are Mosco-convergent in the sense of Kuwae–Shioya [13] (Theorem 3.14);

(D) Examples for
- Brownian motions on Riemannian manifolds (Section 4.1);
- uniformly elliptic diffusions on Riemannian manifolds (Section 4.2).

Let us explain (A). Let \(C(X)\) be the set of continuous paths from \([0, T]\) to \(X\) equipped with the uniform metric where \(T > 0\) is a fixed positive real number. A map \(f: (X, P) \rightarrow (Y, Q)\) is called an \(\(\varepsilon, \delta\)\)-isomorphism if

- \(f: X \rightarrow Y\) is an \(\varepsilon\)-isometry (see Section 2);
- The following inequalities hold:
  \[
  \Phi_{f *} P(A) \leq Q(A^{\delta e^\varepsilon}) + \delta e^\varepsilon, \quad Q(A) \leq \Phi_{f *} P(A^{\delta e^\varepsilon}) + \delta e^\varepsilon, \\
  \Phi_{f^{-1} *} Q(B) \leq P(B^{\delta e^\varepsilon}) + \delta e^\varepsilon, \quad P(B) \leq \Phi_{f^{-1} *} Q(B^{\delta e^\varepsilon}) + \delta e^\varepsilon
  \]  
(1.1)

for any Borel sets \(A \subset C(Y)\) and \(B \subset C(X)\) and we mean that \(\Phi_f : C(X) \rightarrow C(Y)\) is defined by \(v \mapsto f(v(t))\).

We define \(d_{LP}\) as

\[
d_{LP}((X, P), (Y, Q)) = \inf\{\varepsilon + \delta \geq 0 : \exists (\varepsilon, \delta)-\text{isomorphism}\}.
\]  
(1.2)

The inequalities (1.1) indicate how to measure the distance between \(P\) and \(Q\), which live on different path spaces \(C(X)\) and \(C(Y)\): First we push-forward \(P\) by \(\Phi_f\), and then \(\Phi_{f *} P\) and \(Q\) live on the same path space \(C(Y)\). Second we measure \(\Phi_{f *} P\) and \(Q\) by a kind of a modified Prokhorov metric, which involves a space error \(e^\varepsilon\) due to an \(\varepsilon\)-isometry. Note that, if we replace \(e^\varepsilon\) to 1 in Eq. 1.1, then the triangle inequality for the Lipschitz–Prokhorov distance \(d_{LP}\) fails.

We explain (B). After we have the complete metric space \((\mathcal{PM}, d_{LP})\), one of the important questions is:

(Q1) Which subsets in \((\mathcal{PM}, d_{LP})\) are compact?

We restrict our interest to paris of Riemannian manifolds and Markov processes. We introduce a certain subset \(\mathcal{P}_\phi \mathcal{R} = \mathcal{P}_\phi \mathcal{R}(n, K, V, D) \subset \mathcal{PM}\) consisting of pairs \((M, P)\) where \(M\) is a Riemannian manifold with bounds for the sectional curvature, diameter and volume, and \(P = P^\mu\) be the law of a Markov process with an initial distribution \(\mu\) associated with a Dirichlet form whose heat kernel has a uniform bound by a given function \(\phi\) (see details in Section 3). In Theorem 3.9, we will show that \(\mathcal{P}_\phi \mathcal{R}\) is relatively compact in \((\mathcal{PM}, d_{LP})\).

We explain (C). Let \((M_i, P_i)\) be a sequence in the relatively compact subset \(\mathcal{P}_\phi \mathcal{R}\). By the relative compactness, we can take a converging subsequence from a sequence \((M_i, P_i) \in \mathcal{P}_\phi \mathcal{R}\). Then a question is:

(Q2) When does \((M_i, P_i)\) converge in \(d_{LP}\) without taking subsequences?

In Theorem 3.14, we will give a sufficient condition for such convergence in terms of the Mosco-convergence of the corresponding Dirichlet forms in the sense of Kuwae–Shioya [13]. The Mosco-convergence was first introduced by Mosco [14] (see also [15]). In [13], they generalized the Mosco-convergence to the case of varying state spaces.
Here we refer to some related topics. In Ogura [16], the author dealt with a convergence of discretized Brownian motions on compact Riemannian manifolds converging in the measured Gromov–Hausdorff sense. In [20], the author showed that the weak convergence of the Brownian motions in some common metric space is equivalent to the measured Gromov–Hausdorff convergence of the underlying metric measure spaces under Riemannian Curvature-Dimension (RCD) condition. In [21], the author studied a convergence of non-symmetric diffusion processes on RCD spaces. In [19], the author studied weak convergence of Markov processes on ultra-metric spaces. In [2, 3], they considered the weak convergence of strong Markov processeses on tree-like metric measure spaces. They studied relations between the Gromov vague convergence, or the Gromov-Hausdorff vague convergence of the underlying spaces, and the weak convergence of the law of strong Markov processes.

The present paper is organized as follows. In Section 2, we introduce the Lipschitz–Prokhorov distance $d_{LP}$ and show that $(PM, d_{LP})$ is a complete metric space. In Section 3, we give a sufficient condition for relative compactness and also give a sufficient condition for sequences in relatively compact sets to be convergent without taking subsequences. In Section 4, we give several examples. In Section 4.1, we consider the case of Brownian motion on Riemannian manifolds. In Section 4.2, we consider the case of diffusions associated with the uniformly elliptic second-order differential operators.

2 Lipschitz–Prokhorov Distance

In this section, we introduce a distance $d_{LP}$ on $PM$, called the Lipschitz–Prokhorov distance and show $(PM, d_{LP})$ is a complete metric space. We first recall the Lipschitz distance and the Prokhorov distance briefly.

Recall the Lipschitz distance. We say that a map $f : X \to Y$ between two metric spaces is an isometry if $f$ is surjective and distance preserving. Let $M$ denote the set of isometry classes of compact metric spaces. Let $X$ and $Y$ be in $M$. For a bi-Lipschitz homeomorphism $f : X \to Y$, the dilation of $f$ is defined to be the smallest Lipschitz constant of $f$:

$$\text{dil}(f) = \sup_{x \neq y} \frac{d(f(x), f(y))}{d(x, y)}.$$

For $\varepsilon \geq 0$, a bi-Lipschitz homeomorphism $f : X \to Y$ is said to be an $\varepsilon$-isometry if

$$|\log \text{dil}(f)| + |\log \text{dil}(f^{-1})| \leq \varepsilon.$$

By definition, 0-isometry is an isometry. The Lipschitz distance $d_L(X, Y)$ between $X$ and $Y$ is defined to be the infimum of $\varepsilon \geq 0$ such that an $\varepsilon$-isometry between $X$ and $Y$ exists:

$$d_L(X, Y) = \inf\{\varepsilon \geq 0 : \exists f : X \to Y \text{ $\varepsilon$-isometry}\}.$$

If no bi-Lipschitz homeomorphism exists between $X$ and $Y$, we define $d_L(X, Y) = \infty$. We say that a sequence $X_i$ in $M$ Lipschitz converges to $X$ if

$$d_L(X_i, X) \to 0 \quad (i \to \infty).$$

**Remark 2.1** Compared to the Gromov-Hausdorff distance, the Lipschitz distance can deal with smaller classes, which are bi-Lipschitz homeomorphic classes. For instance, collapsing sequences (i.e., the Hausdorff dimension of the limit space is less than that of the sequence) of Riemannian manifolds are excluded from the scope of the Lipschitz distance because the Hausdorff dimension is a bi-Lipschitz invariant. For example, let $S^n$ be a two dimensional
flat torus $S^1 \times \frac{1}{n} S^1$. Then as $n$ tends to infinity, $S_n$ converges to the circle $S^1$ in the Gromov-Hausdorff sense, but not in the Lipschitz sense.

We note that $(\mathcal{M}, d_L)$ is a complete metric space. This fact may be known, but we do not know any references and, for the readers’ convenience, we will prove the completeness of $(\mathcal{M}, d_L)$ in Proposition A.1 in Appendix. Note that $(\mathcal{M}, d_L)$ is not separable because the Hausdorff dimensions of $X$ and $Y$ must coincide if $d_L(X, Y) < \infty$ (see [7, Proposition 1.7.19] and Remark A.2 in the Appendix in this paper). We refer the reader to e.g., [7, 10] for details of the Lipschitz convergence.

Recall the Prokhorov distance. For $T > 0$, let $C_T(X)$ denote the space of continuous maps from $[0, T]$ to a compact metric space $X$ with the uniform metric

$$d_C(v, w) = \sup_{t \in [0, T]} d(v(t), w(t)) \quad \text{for} \quad v, w \in C_T(X).$$

We fix a constant $T > 0$ and we write $C(X)$ shortly for $C_T(X)$. Let $P(C(X))$ denote the set of probability measures on $C(X)$.

The Prokhorov distance between two probability measures $P$ and $Q$ on $C(X)$ is defined to be

$$d_P(P, Q) = \inf\{\delta \geq 0 : P(A) \leq Q(A^\delta) + \delta, Q(A) \leq P(A^\delta) + \delta \}$$

for any Borel set $A \subset C(X)$, where $A^\delta = \{x \in C(X) : d_C(x, A) = \inf_{y \in A} d_C(x, y) < \delta\}$. We know that $(P(C(X)), d_P)$ is a complete separable metric space (see [5, §6]). We refer the reader to e.g., [5, §6] for details of the Prokhorov distance.

Now we introduce the Lipschitz–Prokhorov distance. For a continuous map $f : X \to Y$, we define $\Phi_f : C(X) \to C(Y)$ by

$$\Phi_f(v)(t) = f(v(t)) \quad (v \in C(X), t \in [0, T]).$$

Let $(X, P)$ be a pair of a compact metric space $X$ and a probability measure $P$ on $C(X)$. Note that

$P$ is 
not
a probability measure on $X$, but on $C(X)$.

We say that two pairs of $(X, P)$ and $(Y, Q)$ are isomorphic if there is an isometry $f : X \to Y$ such that the push-forward measure $\Phi_f P$ is equal to $Q$. Note that $\Phi_f P = Q$ implies $\Phi_{f^{-1}} Q = P$ and thus the isomorphic relation becomes an equivalence relation. Let $\mathcal{PM}$ denote the set of isomorphism classes of pairs $(X, P)$. Let $(X, P)$ and $(Y, Q)$ be in $\mathcal{PM}$. Now we introduce a notion of an $(\varepsilon, \delta)$-isomorphism, which is a kind of generalization of $\varepsilon$-isometry. A map $f : (X, P) \to (Y, Q)$ is called an $(\varepsilon, \delta)$-isomorphism if the following hold:

(i) $f : X \to Y$ is an $\varepsilon$-isometry;

(ii) the following inequalities hold:

$$\Phi_f P(A) \leq Q(A^{\delta e^\varepsilon}) + \delta e^\varepsilon, \quad Q(A) \leq \Phi_f P(A^{\delta e^\varepsilon}) + \delta e^\varepsilon,$$

$$\Phi_{f^{-1}} Q(B) \leq P(B^{\delta e^\varepsilon}) + \delta e^\varepsilon, \quad P(B) \leq \Phi_{f^{-1}} Q(B^{\delta e^\varepsilon}) + \delta e^\varepsilon,$$

for any Borel sets $A \subset C(Y)$ and $B \subset C(X)$.

We now define a distance between $(X, P)$ and $(Y, Q)$ in $\mathcal{PM}$, which is called the Lipschitz–Prokhorov distance.
Definition 2.2 Let \((X, P)\) and \((Y, Q)\) be in \(\mathcal{PM}\). The Lipschitz–Prokhorov distance between \((X, P)\) and \((Y, Q)\) is defined to be the infimum of \(\epsilon + \delta \geq 0\) such that an \((\epsilon, \delta)\)-isomorphism \(f : (X, P) \to (Y, Q)\) exists:

\[
d_{LP}(X, P), (Y, Q) = \inf \{\epsilon + \delta \geq 0 : \exists f : (X, P) \to (Y, Q) \text{ \((\epsilon, \delta)\)-isomorphism}\}.
\]

If there is no \((\epsilon, \delta)\)-isomorphism between \((X, P)\) and \((Y, Q)\), we define

\[
d_{LP}(X, P), (Y, Q) = \infty.
\]

Remark 2.3 If we replace \(e^\epsilon\) to 1 in the inequalities (2.1), then the triangle inequality fails for \(d_{LP}\).

Remark 2.4 If we start at metric measure spaces, it may seem to be more natural than \((X, P)\) to consider triplets \((\mathcal{C}(X), d_{C}, P)\) where \(X \in \mathcal{M}\) and \(P\) is a probability measure on \(\mathcal{C}(X)\) because \(P\) is a probability measure not on \(X\) but on \(\mathcal{C}(X)\). However, I do not know whether the Lipschitz convergence of \(\mathcal{C}(X_i)\) implies the Lipschitz convergence of \(X_i\), or not. This makes it difficult for us to equip a good metric with the set of \((\mathcal{C}(X), d_{C}, P)\) enough for capturing the Lipschitz convergence of the underlying state spaces \(X\).

It is clear by definition that \(d_{LP}\) is well-defined in \(\mathcal{PM}\), that is, if \((X, P)\) is isomorphic to \((X', P')\) and \((Y, Q)\) is isomorphic to \((Y', Q')\), then

\[
d_{LP}(X, P), (Y, Q) = d_{LP}(X', P'), (Y', Q').
\]

It is also clear by definition that \(d_{LP}(X, P), (X, P) = 0\), and \(d_{LP}\) is non-negative and symmetric. To show that \(d_{LP}\) is a metric on \(\mathcal{PM}\), it is enough to show that \(d_{LP}\) satisfies the triangle inequality and that \((X, P)\) and \((Y, Q)\) are isomorphic if \(d_{LP}(X, P), (Y, Q) = 0\).

Before the proof, we utilize the following lemma:

Lemma 2.5 Let \(f : X \to Y\) be an \(\epsilon\)-isometry. Then \(\Phi_{f} : \mathcal{C}(X) \to \mathcal{C}(Y)\) is also an \(\epsilon\)-isometry with respect to the uniform metric \(d_{C}\). As a byproduct, for any \(a \geq 0\) and Borel set \(A \subset \mathcal{C}(Y)\), we have

\[
\Phi_{f}^{-1}(A^a) \subset \Phi_{f}^{-1}(A)^{\alpha \epsilon}\quad \text{and} \quad \Phi_{f}^{-1}(A)^{a} \subset \Phi_{f}^{-1}(A^{\alpha \epsilon}),
\]

and, for any Borel set \(B \subset \mathcal{C}(X)\), we have

\[
\Phi_{f}(B^a) \subset \Phi_{f}(B)^{\alpha \epsilon}\quad \text{and} \quad \Phi_{f}(B)^{a} \subset \Phi_{f}(B^{\alpha \epsilon}).
\]

Proof We first show that \(\Phi_{f} : \mathcal{C}(X) \to \mathcal{C}(Y)\) is an \(\epsilon\)-isometry. Since \(f\) is homeomorphic, it is clear that \(\Phi_{f}\) is a homeomorphism. It is enough to show that \(\text{dil}(\Phi_{f}) = \text{dil}(f)\) and \(\text{dil}(\Phi_{f}^{-1}) = \text{dil}(f^{-1})\). Let \(v, w \in \mathcal{C}(X)\). By the compactness of \([0, T]\) and the continuity of \(v, w\) and \(f\), there are \(t_0 \in [0, T]\) and \(s_0 \in [0, T]\) such that

\[
d_{C}(v, w) = d\left(v(t_0), w(t_0)\right)\quad \text{and} \quad d_{C}(\Phi_{f}(v), \Phi_{f}(w)) = d\left(f\left(v(s_0)\right), f\left(w(s_0)\right)\right).
\]

Then we have

\[
\frac{d_{C}(\Phi_{f}(v), \Phi_{f}(w))}{d_{C}(v, w)} = \frac{d\left(f\left(v(s_0)\right), f\left(w(s_0)\right)\right)}{d\left(v(t_0), w(t_0)\right)} \leq \text{dil}(f) \frac{d\left(v(s_0), w(s_0)\right)}{d\left(v(t_0), w(t_0)\right)} \leq \text{dil}(f).
\]
Thus we have $dil(\Phi_f) \leq dil(f)$. For $x \in X$, let $c_x \in C(X)$ denote the constant path on $x$, that is, $c_x(t) = x$ for all $t \in [0, T]$. Then we have

$$
\frac{d(f(x), f(y))}{d(x, y)} = \frac{d_c(\Phi_f(c_x), \Phi_f(c_y))}{d_c(c_x, c_y)} \leq dil(\Phi_f).
$$

Thus we have $dil(\Phi_f) = dil(f)$. By the same argument, we also have $dil(\Phi_f^{-1}) = dil(f^{-1})$. These imply that $\Phi_f$ is an $\varepsilon$-isometry.

We second show the inclusions in the statement. It is enough to show one of the inclusions, say, $\Phi_f(B') \subset \Phi_f(B)\delta\varepsilon^a$ for $a \geq 0$ and any Borel sets $B \subset C(X)$. Let $x \in \Phi_f(B')$ and $y \in B'$ such that $\Phi_f(y) = x$. Since $\Phi_f$ is an $\varepsilon$-isometry, we have

$$
d_c(x, \Phi_f(B)) \leq dil(\Phi_f)d_c(y, B) \leq e^\varepsilon d_c(y, B) \leq a\varepsilon^a.
$$

Thus we have $x \in \Phi_f(B)\delta\varepsilon^a$ and finish the proof.

Now we show that $d_{L,P}$ is a metric on $\mathcal{P}\mathcal{M}$.

**Theorem 2.6** $d_{L,P}$ is a metric on $\mathcal{P}\mathcal{M}$.

**Proof** It is enough to show the following two statements:

(i) $d_{L,P}$ satisfies the triangle inequality;

(ii) $d_{L,P}((X, P), (Y, Q)) = 0$ implies that $(X, P)$ and $(Y, Q)$ are isomorphic.

We first show the statement (i). Let $(X, P)$, $(Y, Q)$ and $(Z, R) \in \mathcal{P}\mathcal{M}$ such that there are $(\varepsilon_1, \delta_1)$-isomorphism $f_1 : X \to Y$ and $(\varepsilon_2, \delta_2)$-isomorphism $f_2 : Y \to Z$. It suffices to show that $f_2 \circ f_1 : X \to Z$ is an $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-isomorphism. In fact, this implies

$$
d_{L,P}((X, P), (Z, R)) < \varepsilon_1 + \varepsilon_2 + \delta_1 + \delta_2.
$$

By taking the infimum of $\varepsilon_1 + \delta_1$ and $\varepsilon_2 + \delta_2$, we have the triangle inequality.

Thus we now show that $f_2 \circ f_1 : X \to Z$ is an $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-isomorphism. We know that $f_2 \circ f_1$ is an $(\varepsilon_1 + \varepsilon_2)$-isometry (see e.g., [7, Theorem 7.2.4]). For any Borel set $A \subset C(Z)$, we have

$$
\Phi_{f_2 \circ f_1} P(A) = (\Phi_{f_1} P)(\Phi_{f_2}^{-1}(A))
\leq Q(\Phi_{f_2}^{-1}(A)^{\delta_1 e^{\varepsilon_1}}) + \delta_1 e^{\varepsilon_1}
\leq Q(\Phi_{f_2}^{-1}(A^{\delta_1 e^{\varepsilon_1+\varepsilon_2}})) + \delta_1 e^{\varepsilon_1}
\leq R(A^{\delta_1 e^{\varepsilon_1+\varepsilon_2} + \delta_2 e^{\varepsilon_2}}) + \delta_1 e^{\varepsilon_1} + \delta_2 e^{\varepsilon_2}
\leq R(A^{(\delta_1 + \delta_2) e^{\varepsilon_1+\varepsilon_2}}) + (\delta_1 + \delta_2) e^{\varepsilon_1+\varepsilon_2}.
$$

The inequality of the third line follows from Lemma 2.5. The other directions of Eq. 2.1 can be shown by the same argument. Thus we have that $f_2 \circ f_1$ is an $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-isomorphism. We have the triangle inequality.

We second show (ii). We show that there is an isometry $\tau : X \to Y$ such that $\Phi_1 : (C(X), P) \to (C(Y), Q)$ is a measure-preserving map, that is, for any real-valued uniformly continuous and bounded function $u$ on $C(Y)$, we have

$$
\int_{C(X)} u \circ \Phi_{\tau} dP = \int_{C(Y)} u \ dQ.
$$

\[\square\] Springer
Let \( f_i : X \to Y \) be an \((\epsilon_i, \delta_i)\)-isomorphism with \( \epsilon_i, \delta_i \to 0 \) as \( i \to \infty \). Since the dilations of \( \{f_i : i \in \mathbb{N}\} \) are uniformly bounded, by the Ascoli–Arzelà theorem, we can take a subsequence from \( \{f_i : i \in \mathbb{N}\} \) converging uniformly to a continuous function \( \iota \). Since \( \epsilon_i \to 0 \) as \( i \to \infty \), we have that \( \iota \) is an isometry from \( X \) to \( Y \), and, for any \( \varepsilon > 0 \), there is an \( i_0 \) such that, for any \( i_0 \leq i \), we have \( d(\iota(x), f_i(x)) < \varepsilon \) for all \( x \in X \). See e.g., [7, Theorem 7.2.4] for details. By this fact, we have

\[
d_C(\Phi_1(v), \Phi_{f_i}(v)) \leq \varepsilon \quad (\forall i \geq i_0, \forall v \in C(X)).
\]

By the uniform continuity of \( u \), we have

\[
\left| \int_{C(X)} u \circ \Phi_i \, dP - \int_{C(X)} u \circ \Phi_{f_i} \, dP \right| \leq \varepsilon' \, P(C(X)) \to 0 \quad (i \to \infty).
\]

(2.3)

Since \( d_P(\Phi_{f_i}^* P, Q) \to 0 \) as \( i \to \infty \), we know that \( \Phi_{f_i}^* P \) converges weakly to \( Q \) as \( i \to \infty \) (see e.g. [5, §6]):

\[
\int_{C(Y)} u \circ \Phi_i \, dP = \int_{C(Y)} u \, d(\Phi_{f_i}^* P) \to \int_{C(Y)} u \, dQ \quad (i \to \infty).
\]

(2.4)

By Eqs. 2.3 and 2.4, we have the equality (2.2) and we finish the proof.

Remark 2.7 When we take \( X = Y \), by definition, we have \( d_P(P, Q) \geq d_LP((X, P), (X, Q)) \). The relation between \( d_P \) and \( d_LP \) is as follows:

\[
d_LP((X, P), (X, Q)) = \inf_{\Phi:X \to X \text{isometry}} d_P(\Phi_{f_i}^* P, Q).
\]

(2.5)

The following example shows that \( d_LP((X, P), (X, Q)) = 0 \) does not imply \( d_P(P, Q) = 0 \): Let \( X = S^1 \) with the metric \( d \) where \( d \) is the restriction of the Euclidean metric in \( \mathbb{R}^2 \). Let \( x, y \in S^1 \) with \( x \neq y \). Let \( c_x, c_y \in C(S^1) \) denote the constant paths on \( x \) and \( y \), that is, \( c_x(t) = x \) and \( c_y(t) = y \) for all \( t \in [0, T] \). Let \( \delta_{c_x} \) and \( \delta_{c_y} \) be the Dirac measures on \( c_x \) and \( c_y \). Let \( f : S^1 \to S^1 \) be the rotation which rotate \( x \) to \( y \). Then, by Eq. 2.5, we have

\[
d_LP((S^1, \delta_{c_x}), (S^1, \delta_{c_y})) \leq d_P(\Phi_{f_i}^* \delta_{c_x}, \delta_{c_y}) = 0.
\]

We see, however, that \( d_P(\delta_{c_x}, \delta_{c_y}) = d(x, y) \neq 0 \). See also Fig. 1 below.

Remark 2.8 If we consider a sequence of metric spaces \( X_i \) converging to \( X \) in the sense of the Gromov–Hausdorff distance alternatively to the Lipschitz distance, we can embed \( X_i \)
and $X$ into one common metric space by isometric embeddings. By using these embeddings, we can consider the topology for each given sequence $(X_i, P_i)$ (see [3, 20]). However, we still have no good idea to give a metric on the space $\mathcal{P}\mathcal{M}$, or equip a topology in the whole space $\mathcal{P}\mathcal{M}$ in this case. This is because if we try to give a metric (say, $d_{GH}$) for this case with a similar idea to the Lipschitz-Prokhorov distance $d_{LP}$ in Definition 2.2, then as in the proof of Theorem 2.6, we need to push-forward a probability measure $P$ by a measurable approximation map $f$ in order to show that $d_{GH}((X, P), (Y, Q)) = 0$ implies that $(X, P)$ is isometric to $(Y, Q)$. See also [7, Theorem 7.3.30] for the part of $d_{GH}(X, Y) = 0$ implying that $X$ is isometric to $Y$. However, the push-forwarded measure $\Phi_{f, P}$ does not live on the space of continuous paths, but lives on the space of measurable paths, which makes it difficult for us to come down to the standard framework of convergences of stochastic processes.

Now we show that the metric space $(\mathcal{P}\mathcal{M}, d_{LP})$ is complete.

**Theorem 2.9** The metric space $(\mathcal{P}\mathcal{M}, d_{LP})$ is complete.

**Proof** Let $\{(X_i, P_i) : i \in \mathbb{N}\}$ be a $d_{LP}$-Cauchy sequence in $\mathcal{P}\mathcal{M}$. It is enough to show that there are a pair $(X, P) \in \mathcal{P}\mathcal{M}$ and a family of $(\epsilon_i, \delta_i)$-isomorphisms $f_i : (X_i, P_i) \to (X, P)$ with $\epsilon_i, \delta_i \to 0$ as $i \to \infty$.

The existence of $X$: The existence of $X$ follows directly from the completeness of $(\mathcal{M}, d_L)$. In fact, since $\{(X_i, P_i) : i \in \mathbb{N}\}$ is a $d_{LP}$-Cauchy sequence, the sequence $\{X_i : i \in \mathbb{N}\}$ is a $d_L$-Cauchy sequence in $\mathcal{M}$. By the completeness of $(\mathcal{M}, d_L)$ (see Proposition A.1 in Appendix), there is a compact metric space $X$ such that

$$d_L(X_i, X) \to 0 \quad (i \to \infty). \quad (2.6)$$

The existence of $P$ and $f_i$: Since $\{(X_i, P_i) : i \in \mathbb{N}\}$ is a $d_{LP}$-Cauchy sequence, there is a family of $(\epsilon_{ij}, \delta_{ij})$-isomorphisms $f_{ij} : X_i \to X_j$ for $i < j$ with $\epsilon_{ij} \to 0$ and $\delta_{ij} \to 0$ as $i, j \to \infty$. Take a subsequence such that $\epsilon_{i, i+1} + \delta_{i, i+1} < 1/2^i$. Let $\tilde{f}_{ij} : X_i \to X_j$ be defined by

$$\tilde{f}_{ij} = f_{j-1,j} \circ f_{j-2,j-1} \circ \cdots \circ f_{i,i+1} \quad (i < j), \quad (2.7)$$

and $\tilde{\epsilon}_{ij} = \sum_{l=i}^{j-1} \epsilon_{l,l+1}$ and $\tilde{\delta}_{ij} = \sum_{l=i}^{j-1} \delta_{l,l+1}$. By the proof of Theorem 2.6, we see that $\tilde{f}_{ij}$ is an $(\tilde{\epsilon}_{ij}, \tilde{\delta}_{ij})$-isomorphism and $\tilde{\epsilon}_{ij}, \tilde{\delta}_{ij} \to 0$ as $i, j \to \infty$. By the proof of Proposition A.1 in Appendix (see also the equality (6) in Appendix), there is a family of $\epsilon_i$-isometries $f_i : X_i \to X$ such that $\epsilon_i \to 0$ as $i \to \infty$ and

$$f_i \circ \tilde{f}_{ji} = f_j.$$

This implies that

$$\Phi_{f_i} \circ \Phi_{\tilde{f}_{ji}} = \Phi_{f_j}. \quad (2.8)$$

We now show that $\{\Phi_{f_{ij}} P_i : i \in \mathbb{N}\}$ is a $d_P$-Cauchy sequence in $\mathcal{P}(\overline{C}(X))$. Let us set

$$\epsilon = \epsilon(i, j) = (\tilde{\delta}_{ij} + \tilde{\epsilon}_{ji}) \epsilon_{ij} + \epsilon_{ii} + \epsilon_j.$$

Note that $\epsilon \to 0$ as $i, j \to \infty$. It suffices to show that, for any Borel set $A \subset \overline{C}(X)$,

$$\Phi_{f_{i}} P_i(A) - \Phi_{f_{j}} P_j(A^\epsilon) \leq \epsilon, \quad \Phi_{f_{i}} P_i(A^\epsilon) - \Phi_{f_{j}} P_j(A) \leq \epsilon.$$
We only show the left-hand side of the above inequalities (the right-hand side can be shown by the same argument). For any Borel set \( A \subset \mathcal{C}(X) \), we have

\[
\Phi_{f_i \circ f_j} P_i(A) - \Phi_{f_j} P_j(A^\delta)
\]

\[
= (\Phi_{f_i \circ f_j} P_i(A) - \Phi_{f_i \circ f_j \circ f_i} P_i(A^\delta)) + (\Phi_{f_i \circ f_j} P_j(A^\delta) - \Phi_{f_j \circ f_j} P_j(A))
\]

\[
= (I) + (II) + (III).
\]

Since \( \tilde{f}_{ji} : X_j \rightarrow X_i \) is the \((\tilde{\varepsilon}_{ji}, \tilde{\delta}_{ji})\)-isomorphism, we have

\[
(I) = (P_i \Phi_{f_i}^{-1}(A) - (\Phi_{f_j} \circ f_j) \Phi_{f_i} P_i(A^\delta)) \leq \tilde{\delta}_{ji} \tilde{\varepsilon}_{ji} \leq \varepsilon.
\]

By the same argument, we also have (III) \( \leq \varepsilon \). For the estimate of (II), by Lemma 2.5 and Eq. 2.8, we have

\[
\Phi_{f_j \circ f_i} P_i(A) = (\Phi_{f_j} \circ f_j)(\Phi_{f_i} P_i(A)) \leq P_j(\Phi_{f_j}^{-1}(A \tilde{\delta}_{ij} \tilde{\varepsilon}_{ij})) + \tilde{\delta}_{ij} \tilde{\varepsilon}_{ij}
\]

\[
\leq P_j(\Phi_{f_j}^{-1}(A \tilde{\delta}_{ij} \tilde{\varepsilon}_{ij} + \varepsilon)) + \tilde{\delta}_{ij} \tilde{\varepsilon}_{ij}
\]

\[
= P_j(\Phi_{f_j}^{-1}(A \tilde{\delta}_{ij} \tilde{\varepsilon}_{ij} + \varepsilon)) + \tilde{\delta}_{ij} \tilde{\varepsilon}_{ij}
\]

\[
= \Phi_{f_j \circ f_i} P_i(A^\delta) + \varepsilon.
\]

We used Lemma 2.5 in the second line and Eq. 2.8 in the third line. We have (II) \( \leq \varepsilon \). Thus \( \{\Phi_{f_i \circ f_i} P_i : i \in \mathbb{N}\} \) is a \( d_P \)-Cauchy sequence. By the completeness of \( (\mathcal{P}(\mathcal{C}(X)), d_P) \), there exists a probability measure \( P \) on \( \mathcal{C}(X) \) such that

\[
d_P(\Phi_{f_i \circ f_i} P_i, P) \rightarrow 0 \quad (i \rightarrow \infty).
\]  

(2.9)

We finally show that \( f_i : (X_i, P_i) \rightarrow (X, P) \) is an \((\varepsilon_i, \delta_i)\)-isomorphism for some sequence \( \delta_i \rightarrow 0 \) as \( i \rightarrow \infty \). By Eq. 2.9, there is a sequence \( \delta_i \rightarrow 0 \) as \( i \rightarrow \infty \) such that

\[
\Phi_{f_i \circ f_i} P_i(A) \leq P(A^\delta_i) + \delta_i \quad \text{and} \quad P(A) \leq \Phi_{f_i \circ f_i} P_i(A^\delta_i) + \delta_i,
\]  

(2.10)

for any Borel set \( A \subset \mathcal{C}(X) \). By the inequalities (2.10) and Lemma 2.5, we have

\[
\Phi_{f_i^{-1}} P_i(B) = \lim_{f_i^{-1}} P_i\big(\Phi_{f_i}^{-1}(B)\big) \leq \Phi_{f_i} P_i\big(\Phi_{f_i}^{-1}(B)\big) + \delta_i
\]

\[
\leq P_i\big(\Phi_{f_i}^{-1}(B^\delta_i)\big) + \delta_i e^{\delta_i}
\]

\[
\leq P_i\big(B^\delta_i\big) + \delta_i e^{\delta_i} \quad (\forall B \subset \mathcal{C}(X_i) : \text{Borel}).
\]  

(2.11)

By the same argument, we also have

\[
P_i(B) \leq \Phi_{f_i^{-1}} P_i(B^\delta_i e^{\delta_i}) + \delta_i e^{\delta_i} \quad (\forall B \subset \mathcal{C}(X_i) : \text{Borel}).
\]  

(2.12)

Note in Eq. 2.10 that

\[
P(A^\delta_i) + \delta_i \leq P(A^\delta_i e^{\delta_i}) + \delta_i e^{\delta_i} \quad \text{and} \quad \Phi_{f_i \circ f_i} P_i(A^\delta_i) + \delta_i \leq \Phi_{f_i \circ f_i} P_i(A^\delta_i e^{\delta_i}) + \delta_i e^{\delta_i}.
\]  

(2.13)

Thus, by Eqs. 2.10, 2.11, 2.12 and 2.13, we have

\[
\Phi_{f_i \circ f_i} P_i(A) \leq P(A^\delta_i e^{\delta_i}) + \delta_i e^{\delta_i}, \quad P(A) \leq \Phi_{f_i \circ f_i} P_i(A^\delta_i e^{\delta_i}) + \delta_i e^{\delta_i},
\]

\[
\Phi_{f_i^{-1}} P_i(B) \leq P_i(B^\delta_i e^{\delta_i}) + \delta_i e^{\delta_i}, \quad P_i(B) \leq \Phi_{f_i^{-1}} P_i(B^\delta_i e^{\delta_i}) + \delta_i e^{\delta_i},
\]  

(2.14)
for any Borel sets $A \subset C(X)$ and $B \subset C(X_i)$. Since $f_i : X_i \rightarrow X$ is an $\varepsilon_i$-isometry with $\varepsilon_i \rightarrow 0$ as $i \rightarrow 0$, the inequalities (2.14) means that $f_i : (X_i, P_i) \rightarrow (X, P)$ is an $(\varepsilon_i, \delta_i)$-isomorphism with $\varepsilon_i, \delta_i \rightarrow 0$ as $i \rightarrow \infty$. We therefore have the desired result.

Remark 2.10 The metric space $(\mathcal{P}M, d_{LP})$ is not separable. This is because $(\mathcal{M}, d_L)$ is not separable as in Remark A.2. Let $X \in \mathcal{M}$ and $\mathcal{M}_X = \{Y \in \mathcal{M} : d_L(X, Y) < \infty\}$. Let $\mathcal{P} \mathcal{M}_X$ be the set of isomorphism classes of pairs $(Y, P)$ where $Y \in \mathcal{M}_X$ and $P \in \mathcal{P}(C(Y))$. By [22], there is an $X \in \mathcal{M}$ such that $(\mathcal{M}_X, d_L)$ is not separable. Thus, for such $X$, $(\mathcal{P} \mathcal{M}_X, d_{LP})$ is not separable.

By the proof of Theorem 2.9, we have the following:

**Corollary 2.11** Let $(X_i, P_i), (X, P) \in \mathcal{P} \mathcal{M}$ for all $i \in \mathbb{N}$. The sequence $(X_i, P_i)$ converges to $(X, P)$ in $d_{LP}$ as $i \rightarrow \infty$ if and only if there is a family of $\varepsilon_i$-isometries $f_i : X_i \rightarrow X$ with $\varepsilon_i \rightarrow 0$ as $i \rightarrow \infty$ such that

$$\Phi_{f_i \#} P_i \rightharpoonup P \quad \text{weakly} \quad (i \rightarrow \infty).$$

### 3 Relative Compactness and Uniqueness of Limit Points

In this section, we focus on conditions for relative compactness and uniqueness of limit points with respect to the topology of the Lipschitz-Prokhorov distance $d_{LP}$. In the first section, we give sufficient conditions for relative compactness in terms of Aldous’s tightness criterion. In the second section, we give sufficient conditions for uniqueness of limit points by the weak convergence of the finite-dimensional distributions. In the third section, we consider the particular case that state spaces are Riemannian manifolds and probability measures are the laws of Markov processes associated with strongly local regular symmetric Dirichlet forms. We first give sufficient conditions for relative compactness in terms of heat kernel estimates. Secondly, we give sufficient conditions for the weak convergence of the finite-dimensional distributions in terms of the Mosco convergence of Dirichlet forms.

#### 3.1 Relative Compactness

In this subsection, we study a relative compactness criterion for sequences in $\mathcal{P} \mathcal{M}$.

**Assumption 3.1** Let $(M_i, P_i) \in \mathcal{P} \mathcal{M}$ be a sequence of pairs satisfying the following conditions:

(i) $\{M_i\}$ is a relatively compact with respect to the Lipschitz topology;

(ii) $P_i$ is the law of a strong Markov process $X_i$ with continuous sample paths and having a common probability space $(\Omega, \mathcal{G}, P)$. Moreover, we suppose that the following (varying state space version of) *Aldous’s tightness criterion* holds: for any bounded sequence of stopping times $\tau_i \in [0, T]$ and any $\delta_i > 0$ with $\delta_i \rightarrow 0$ and $\tau_i + \delta_i \in [0, T]$, it holds that

$$d_i(X_i(\tau_i), X_i(\tau_i + \delta_i)) \rightarrow 0 \quad \text{in prob. w.r.t. } P. \quad (3.1)$$

**Theorem 3.2** Under Assumption 3.1, the sequence $\{(M_i, P_i)\}_{i \in \mathbb{N}} \subset \mathcal{P} \mathcal{M}$ is relatively compact with respect to the Lipschitz-Prokhorov distance $d_{LP}$. 

\[ \text{ Springer} \]
Remark 3.3 Aldous’s tightness criterion consists of a tightness of one-dimensional marginals \( \{X(t)\}_{i \in \mathbb{N}} \) (\( \forall t \geq 0 \)) and the condition (3.1) ([1]). However, since our state spaces are always compact, only the condition (3.1) is enough.

Proof of Theorem 3.2 Since \( \{M_i\}_{i \in \mathbb{N}} \) is relatively compact with respect to the topology induced by the Lipschitz distance, there exist a limit metric space \( X \) with \( X_i \rightarrow X \) in \( d_L \) after taking a converging sequence (we use the same subscript \( i \) for this subsequence for simplicity of notation) and \( \varepsilon_i \)-isometries \( f_i : X_i \rightarrow X \) with \( \varepsilon_i \rightarrow 0 \). Thus we have

\[
d(f_i(X_i(t)), f_i(X_i(s))) < \varepsilon_i d_i(X_i(t), X_i(s)) \quad \forall t, s \geq 0.
\]

By Eq. 3.1 in Assumption 3.1 and Aldous’s tightness criterion [1], we conclude that \( \{\Phi_1 f_i^* P_i\}_{i \in \mathbb{N}} \) is tight. Therefore by Corollary 2.11, we have shown that \( \{(X_i, P_i)\}_{i \in \mathbb{N}} \) is relatively compact.

3.2 Uniqueness of Limit Points

In this subsection, we give sufficient conditions for uniqueness of limit points by the weak convergence of the finite-dimensional distributions.

Assumption 3.4 Let \( (M_i, P_i), (M, P) \in \mathcal{PM}(i \in \mathbb{N}) \). Assume the following:

(i) Let \( M_i \) converge to \( M \) in the Lipschitz distance with an \( \varepsilon_i \)-isometry \( f_i : M_i \rightarrow M \) satisfying \( \varepsilon_i \rightarrow 0 \) as \( i \rightarrow \infty \). Let \( \Phi_1 f_i^* P_i \) be a relatively compact.

(ii) \( P_i \) and \( P \) are the laws of a stochastic processes \( X_i \) and \( X \) respectively with continuous sample paths. The weak convergence of the finite-dimensional distributions holds: for any \( k \in \mathbb{N} \), any \( 0 = t_0 < t_1 < t_2 < \ldots < t_k \leq T \) and any bounded continuous functions \( g_1, g_2, \ldots, g_k \in C_b(M) \),

\[
E_i \left( f_i^* g_1 \circ X_i(t_1) f_i^* g_2 \circ X_i(t_2) \cdots f_i^* g_k \circ X_i(t_k) \right) \rightarrow E \left( g_1 \circ X(t_1) g_2 \circ X(t_2) \cdots g_k \circ X(t_k) \right) \quad (i \rightarrow \infty).
\]

Theorem 3.5 Under Assumption 3.4, \( (M_i, P_i) \) converges to \( (M, P) \) in the Lipschitz-Prokhorov distance \( d_{LP} \).

Proof By the standard argument that the uniqueness of limit points follows from the weak convergence of the finite-dimensional distribution (e.g., [5]), Assumption 3.4 implies the weak convergence of \( \Phi_1 f_i^* P_i \) to \( P \). By Corollary 2.11, we have the desired result.

3.3 The Case of Riemannian Manifolds

In this subsection, we consider the particular case that state spaces are Riemannian manifolds and probability measures are the laws of Markov processes associated with strongly local regular symmetric Dirichlet forms.

We first give a sufficient condition for relative compactness of sets of Riemannian manifolds with respect to the topology of the Lipschitz distance \( d_L \). For a positive integer \( n \) and positive constants \( K, V, D > 0 \), let \( R(n, K, V, D) \) denote the set of isometry classes of \( n \)-dimensional connected compact Riemannian manifolds \( M \) satisfying

\[
|\text{sec}(M)| \leq K, \quad \text{Vol}(M) \geq V, \quad \text{and} \quad \text{Diam}(M) \leq D.
\]

\[208\]
Here $\sec(M)$, Vol$(M)$ and Diam$(M)$ denote the sectional curvature (see [4, Section 4.4.2] for the definition), the Riemannian volume and the diameter of $M$, respectively. We write $R(n, K, V, D)$.

Note that, by the Bishop inequality, there is a constant $V'$ such that

$$\text{Vol}(M) < V', \quad \text{for all } M \in R.$$  (3.4)

By [10, Theorem 8.19] and [12] (see also [4, Theorem 383]), we know that $R$ is relatively compact in $(\mathcal{M}, d_{L})$.  (3.5)

We explain Markov processes considered in this section. Let $(\mathcal{E}, \mathcal{F})$ be a regular symmetric Dirichlet form on $L^{2}(M, \text{Vol})$ and $\{T_{t}\}_{t \in (0, \infty)}$ be the corresponding semigroup on $L^{2}(M, \text{Vol})$. We say that $\{p(t, x, y) : x, y \in M, t \in (0, \infty)\}$ is a heat kernel of $\{T_{t}\}_{t \in (0, \infty)}$ if $p(t, x, y)$ becomes an integral kernel of $T_{t}$: for $f \in L^{2}(M; \text{Vol})$

$$T_{t}f(x) = \int_{M} f(y) p(t, x, y) \text{Vol}(dy) \quad (\forall x \in M).$$

**Assumption 3.6** The following conditions hold:

(i) $(\mathcal{E}, \mathcal{F})$ is a strongly local regular symmetric Dirichlet form on $L^{2}(M, \text{Vol})$;

(ii) the corresponding semigroup $\{T_{t}\}_{t \in (0, \infty)}$ is a Feller semigroup and has a jointly continuous heat kernel $p(t, x, y)$ on $(0, T] \times M \times M$.

See, e.g., [6, Theorem I.9.4] for the Feller property.

**Remark 3.7** We assumed the joint-continuity and the Feller property of the heat kernel only for simplicity. The following arguments can be modified by excluding null-capacity sets with respect to $(\mathcal{E}, \mathcal{F})$ if we need to remove the Feller property assumption. The joint-continuity of a function $\phi$ which will be taken in Eq. 3.8 is assumed for the same reason.

By [6, Theorem I.9.4], there is a Hunt process

$$\left(\Omega, \mathcal{M}, \{\mathcal{M}(t)\}_{t \in [0, \infty)}, \{P^{x}\}_{x \in M}, \{X(t)\}_{t \in [0, \infty)}\right)$$

such that, for bounded Borel function $f$ in $L^{2}(M, \text{Vol})$, we have

$$E^{x}(f(X(t))) = T_{t}f(x) \quad (\forall t \in (0, \infty), \forall x \in M).$$

By the locality of $(\mathcal{E}, \mathcal{F})$, we know that $X(\cdot)$ has continuous paths almost surely. By the strong locality of $(\mathcal{E}, \mathcal{F})$ and the compactness of $M$, we know that $X(t) \in M$ for all $t \in [0, \infty)$ almost surely. Thus we see that $X : \Omega \to \mathcal{C}(M)$ almost surely and the law of $X$ lives on $\mathcal{C}(M)$. We refer the reader to e.g., [9] for details of Dirichlet forms and Hunt processes.

Let $\mu$ be a probability measure on $M$. Let $P^{\mu}$ denote the probability measure with the initial distribution $\mu$:

$$P^{\mu}(A) = \int_{A} P^{x} \mu(dx) \quad (\forall A \subset M : \text{Borel}).$$  (3.7)

Now we introduce a main object in this section, a subset $\mathcal{P}_{\phi}R(n, K, V, D)$ of $\mathcal{P}M$ determined by a certain function $\phi$. Let $\phi : (0, T] \times [0, D] \to [0, \infty)$ be a jointly continuous function satisfying that, for any $\varepsilon > 0$,

$$\lim_{\lambda \to 0} \sup_{r > \varepsilon, \xi \in (0, \lambda]} \phi(\xi, r) = 0,$$  (3.8)

where $D > 0$ is the uniform bound of diameters of elements in $R = R(n, K, V, D)$.  © Springer
Definition 3.8 For a function $\phi$ satisfying the above conditions, the set $\mathcal{P}_\phi \mathcal{R}(n, K, V, D)$ is defined to be the set of isomorphism classes of pairs $(M, P)$ where $M \in \mathcal{R}$ and $P$ is the law of $P^\mu$ for an initial distribution $\mu$ and a Hunt process on $M$ associated with $(\mathcal{E}, \mathcal{F})$ satisfying Assumption 3.6 and that the heat kernel $p(t, x, y)$ is dominated by $\phi$ in the following sense: there exists a $\tau > 0$ such that for all $t \in (0, \tau \wedge T]$, all $x, y \in M$ and all $M \in \mathcal{R}$,

$$p(t, x, y) \leq \phi(t, d_M(x, y)).$$

(3.9)

We also write $\mathcal{P}_\phi \mathcal{R}$ shortly for $\mathcal{P}_\phi \mathcal{R}(n, K, V, D)$.

Then we have the main theorem of this section:

**Theorem 3.9** The set $\mathcal{P}_\phi \mathcal{R}$ is relatively compact in $(\mathcal{PM}, d_{LP})$.

**Remark 3.10** Let $\overline{\mathcal{P}_\phi \mathcal{R}}$ be the completion of $\mathcal{P}_\phi \mathcal{R}$ with respect to $d_{LP}$. As a byproduct of Theorem 3.9, we see $(\overline{\mathcal{P}_\phi \mathcal{R}}, d_{LP})$ is a compact metric space.

Let $\overline{\mathcal{R}}$ be the completion of $\mathcal{R}$ with respect to $d_L$. In general, $M \in \overline{\mathcal{R}}$ has a $C^{1, \alpha}$-Riemannian structure for any $0 < \alpha < 1$. See, e.g., [4, Theorem 384].

**Proof of Theorem 3.9** Since any metric spaces satisfy the first axiom of countability, that is, every point has a countable neighborhood basis, it is enough to show that any sequence $(M_i, P_i) \in \mathcal{P}_\phi \mathcal{R}$ has a converging subsequence $(M_i') \subset \mathcal{R}$. Thus there is a family of $\varepsilon_i$-isometries $f_i : M_i \to M$ with $\varepsilon_i \to 0$ as $i \to \infty$.

By Corollary 2.11, the proof is completed if we show $\{\Phi_{f_i*} P_i : i \in \mathbb{N}\}$ is relatively compact in $(\mathcal{PM}(C(M)), d_P)$. By Theorem 3.2, we only have to check (3.1). By the strong Markov property, Eqs. 3.4 and 3.8, we have that, for any $\eta > 0$,

$$P^x \left( d_i \left( X_i(\tau_i + \delta_i), X_i(\tau_i) \right) > \eta \right) = E^x \left( P^{X_i(\tau_i)} \left( d_i \left( X_i(\delta_i), X_i(0) \right) > \eta \right) \right)$$

$$< \sup_{x \in M_i} P^x \left( d_i \left( X_i(\delta_i), X_i(0) \right) > \eta \right)$$

$$\leq \sup_{x \in M_i, \xi \in (0, \delta_i]} \int_{B(x, \eta)^c} p_i(\xi, x, y) m_i(dy).$$

$$\leq \sup_{x, y \in M_i, \xi \in (0, \delta_i]} \int_{B(x, \eta)^c} \phi(\xi, d_i(x, y)) m_i(dy)$$

$$\leq V' \sup_{x, y \in M_i, \xi \in (0, \delta_i]} \phi(\xi, d_i(x, y))$$

$$i \to \infty \rightarrow 0.$$  

(3.10)

Thus we finish the proof.

We start to consider the second objective in this section, that is, a sufficient condition for sequences in $\mathcal{P}_\phi \mathcal{R}$ to be convergent. Let $(M_i, P_i) \in \mathcal{P}_\phi \mathcal{R}$. By Theorem 3.9, we know that there is a subsequence $(M_i', P_i')$ converging to some $(M, P)$ in the completion $\overline{\mathcal{P}_\phi \mathcal{R}}^{d_{LP}}$ with respect to $d_{LP}$. Hereafter we consider under what conditions, the whole sequence $(M_i, P_i)$ converges to $(M, P)$.
Let \((M_i, g_i)\) be a sequence of Riemannian manifolds with Riemannian metrics \(g_i\). Assume that \(M_i\) converges to some \(M \in \mathcal{M}\) in the Lipschitz distance \(d_L\) with \(\varepsilon_i\)-isometries \(f_i : M_i \to M\). We know that the limit space \(M\) has a structure of the \(n\)-dimensional \(C^{1,\alpha}\) Riemannian manifold for any \(0 < \alpha < 1\). That is, \(M\) is a \(n\)-dimensional \(C^\infty\)-manifold with a \(C^{1,\alpha}\)-Riemannian metric \(g\). See, e.g., [4, Theorem 384 & Fig. 12.37]. Let \(\text{Vol}\), and \(\text{Vol}_i\) be Riemannian volumes induced by \(g_i\) and \(g\).

Let \((\mathcal{E}_i, \mathcal{F}_i)\) be a sequence of Dirichlet forms on \(L^2(M_i; \text{Vol}_i)\) and \((\mathcal{E}, \mathcal{F})\) is a Dirichlet form on \(L^2(M; \text{Vol})\) both satisfying Assumption 3.6. We consider a convergence of Dirichlet forms \((\mathcal{E}_i, \mathcal{F}_i)\) to \((\mathcal{E}, \mathcal{F})\), which is a special case of [13, Definition 2.11]. Let us set \(\mathcal{E}_i(u) := \mathcal{E}_i(u, u)\) for \(u \in \mathcal{F}_i\) and \(\mathcal{E}_i(u) := \infty\) if \(u \in L^2(M_i; \text{Vol}) \setminus \mathcal{F}_i\) (we also treat \((\mathcal{E}, \mathcal{F})\) in the same manner). For \(u \in L^2(M_i)\), we define the push-forward \(f_i^* u \in L^2(M)\) by \(f_i^* u(x) = u \circ f_i^{-1}(x)\) for \(x \in M\). We can check \(f_i^* u \in L^2(M)\) because of the following inequality:

\[
e^{-n \varepsilon_i} \text{Vol} \leq f_i^* \text{Vol}_i \leq e^{n \varepsilon_i} \text{Vol},
\]  

(3.11)

The above inequality follows from the definition of an \(\varepsilon_i\)-isometry. By this inequality, we have that \(f_i^* \text{Vol}_i\) are absolutely continuous with respect to \(\text{Vol}\). Similarly, for \(u \in L^2(M)\), we define the pull-back \(f_i^* u\) by \(u \circ f_i(x)\) for any \(x \in M_i\). We can also check \(f_i^* u \in L^2(M_i)\) by the same argument of \(f_i^* u \in L^2(M)\).

Now we define a convergence of Dirichlet forms, which is a special case of [13, Definition 2.11] (see also [8, Definition 8.1]).

**Definition 3.11** We say that \((\mathcal{E}_i, \mathcal{F}_i)\) converges in the Mosco sense to \((\mathcal{E}, \mathcal{F})\) if the following statement holds: there is a family of \(\varepsilon_i\)-isometries \(f_i : M_i \to M\) with \(\varepsilon_i \to 0\) as \(i \to \infty\) satisfying

(i) for any \(u_i \in L^2(M_i)\) and \(u \in L^2(M)\) satisfying \(f_i^* u_i\) converges weakly to \(u\) in \(L^2(M)\), we have

\[
\liminf_{i \to \infty} \mathcal{E}_i(u_i) \geq \mathcal{E}(u);
\]

(ii) for any \(u \in L^2(M)\), there exists a sequence \(u_i \in L^2(M_i)\) satisfying that \(f_i^* u_i\) converges to \(u\) in \(L^2(M)\) and

\[
\limsup_{i \to \infty} \mathcal{E}_i(u_i) \leq \mathcal{E}(u).
\]

Note that the notion of the Mosco-convergence does not depend on a specific family of \(\varepsilon_i\)-isometries \(f_i\) in the following sense: if \((\mathcal{E}_i, \mathcal{F}_i)\) converges in the Mosco sense to another Dirichlet form \((\mathcal{E}', \mathcal{F}')\) with respect to another family of \(\varepsilon_i\)-isometries \(g_i : M_i \to M'\), then there is an isometry \(\iota : M \to M'\) satisfying

\[
\mathcal{E}'(u, v) = \mathcal{E}(\iota^* u, \iota^* v) \quad (\forall u, v \in \mathcal{D}(\mathcal{E}')).
\]

Let \(\{G_i(\alpha)\}_{\alpha > 0}\) and \(\{G(\alpha)\}_{\alpha > 0}\) be the resolvents corresponding to \((\mathcal{E}_i, \mathcal{F}_i)\) and \((\mathcal{E}, \mathcal{F})\), respectively. We have the following statement, which is a special case of [13, Theorem 2.4] (see also [8, Theorem 8.3]):

**Proposition 3.12** The following statements are equivalent:

(i) \((\mathcal{E}_i, \mathcal{F}_i)\) converges in the Mosco sense to \((\mathcal{E}, \mathcal{F})\);
(ii) there is a family of $\varepsilon_i$-isometries $f_i : M_i \to M$ satisfying
\[ f_i \ast T_i(t) f_i^\ast u \to T(t) u \quad \text{in} \quad L^2(M), \]
for any $u \in L^2(M)$ and the convergence is uniformly in $t \in [0, T]$.

(iii) there is a family of $\varepsilon_i$-isometries $f_i : M_i \to M$ satisfying
\[ f_i \ast G_i(\alpha) f_i^\ast u \to G(\alpha) u \quad \text{in} \quad L^2(M), \]
for any $\alpha > 0$ and any $u \in L^2(M)$.

Proof Modify the proof of [8, Theorem 8.3] as $\pi_i u = f_i^\ast u$ for $u \in L^2(M)$ and $E_i u = f_i^\ast u$ for $u \in L^2(M_i)$. Noting the inequality (3.11), we can do the same argument in [8, Theorem 8.3] and have the desired result. \qed

Assumption 3.13 For $i \in \mathbb{N}$, let $(M_i, P_i) \in \mathcal{P}_0\mathcal{R}$ for $(\mathcal{E}_i, \mathcal{F}_i)$ with an initial distribution $\mu_i = \varphi_i \text{Vol}$, where $\varphi_i \in L^2(M_i)$. Let $(M, P)$ be an element in the completion $\overline{\mathcal{P}_0\mathcal{R}}$ with respect to $d_{L^p}$ and assume $P$ to be the law of $P^\mu$ for a Hunt process associated with $(\mathcal{E}, \mathcal{F})$ satisfying Assumption 3.6 with an initial distribution $\mu = \varphi \text{Vol}$ with $\varphi \in L^2(M)$. Assume that there is a family of maps $f_i : M_i \to M$ satisfying the following conditions:

(i) $f_i$ is an $\varepsilon_i$-isometry with $\varepsilon_i \to 0$ as $i \to \infty$;

(ii) $f_i$ satisfies (i) and (ii) of Definition 3.11;

(iii) $f_i \ast \varphi_i$ converges to $\varphi$ in $L^2(M)$.

Theorem 3.14 If Assumption 3.13 holds, then $(M_i, P_i)$ converges to $(M, P)$ as $i \to \infty$ in the sense of $d_{L^p}$.

Proof of Theorem 3.14 By Theorem 3.9 and Theorem 3.5, it suffices to show that the finite-dimensional distributions of $\Phi_{f_i \ast P_i}$ converge weakly to those of $P$.

Since $M$ is compact, any bounded continuous functions on $M$ are square-integrable. Thus it suffices to show that, for any $k \in \mathbb{N}$, any $0 = t_0 < t_1 < t_2 < \cdots < t_k \leq T$ and any bounded Borel measurable functions $g_1, g_2, \ldots, g_k$ in $L^2(M)$,
\[ E_i(f_i^\ast g_1 \circ X_i(t_1) f_i^\ast g_2 \circ X_i(t_2) \cdots f_i^\ast g_k \circ X_i(t_k)) \to E(g_1 \circ X(t_1)g_2 \circ X(t_2) \cdots g_k \circ X(t_k)) \quad (i \to \infty). \quad (3.12) \]

Let us set inductively
\[ h^k_i = g_k, \quad h^{k-1}_i(\cdot) = g_{k-1}(\cdot) f_i \ast T_i(t_k - t_{k-1}) f_i^\ast h^k_i(\cdot), \]
\[ h^1_i(\cdot) = g_1(\cdot) f_i \ast T_i(t_2 - t_1) f_i^\ast h^2_i(\cdot), \]
and
\[ h^k(\cdot) = g_k(\cdot) \quad h^{k-1}(\cdot) = g_{k-1}(\cdot) T(t_k - t_{k-1}) h^k(\cdot), \]
\[ h^1(\cdot) = g_1(\cdot) T(t_2 - t_1) h^2(\cdot). \]

By Proposition 3.12 and boundedness of $g_k$, we have
\[ \| h^{k-1}_i - h^{k-1} \|^2_{L^2(M)} = \| g_{k-1}(\cdot) f_i \ast T_i(t_k - t_{k-1}) f_i^\ast h^k_i(\cdot) - g_{k-1}(\cdot) T(t_k - t_{k-1}) h^k(\cdot) \|^2_{L^2(M)} \to 0 \quad (i \to \infty). \quad (3.13) \]
Inductively, we have

\[
\|h_i^{k-2} - h^{k-2}\|_{L^2(M)} \\
= \|g_{k-2}(\cdot) f_i^* T_i(t_{k-1} - t_{k-2}) f_i^* h_i^{k-1}(\cdot) - g_{k-2}(\cdot) T(t_{k-1} - t_{k-2}) h^{k-1}(\cdot)\|_{L^2(M)} \\
\leq \|g_{k-2} f_i^* T_i(t_{k-1} - t_{k-2}) f_i^* h_i^{k-1} - g_{k-2} f_i^* T_i(t_{k-1} - t_{k-2}) f_i^* h_i^{k-1}\|_{L^2(M)} \\
+ \|g_{k-2} f_i^* T_i(t_{k-1} - t_{k-2}) f_i^* h_i^{k-1} - g_{k-2} T(t_{k-1} - t_{k-2}) h^{k-1}\|_{L^2(M)} \\
\leq \|g_{k-2}\|_{\infty} \|f_i^* T_i(t_{k-1} - t_{k-2})\|_{op} \|f_i^* h_i^{k-1} - f_i^* h^{k-1}\|_{L^2(M)} \\
+ \|g_{k-2}\|_{\infty} \|f_i^* T_i(t_{k-1} - t_{k-2}) f_i^* h_i^{k-1} - T(t_{k-1} - t_{k-2}) h^{k-1}\|_{L^2(M)} \\
=: (I)_i + (II)_i,
\]

where \(\|f_i^* T_i(t_{k-1} - t_{k-2})\|_{op}\) means the operator norm of \(f_i^* T_i(t_{k-1} - t_{k-2}) : L^2(M_i) \rightarrow L^2(M)\).

The quantity \((II)_i\) converges to 0 as \(i \rightarrow \infty\) by (ii) of Assumption 3.13 and Proposition 3.12.

We estimate \((I)_i\). By the inequality (3.11) and the contraction property of the semigroup \(\{T(t)\}_{t>0}\), we can check easily that there is a constant \(C\) independent to \(i\) satisfying

\[
\|f_i^* T_i(t_{k-1} - t_{k-2})\|_{op} \leq C. \tag{3.14}
\]

By Eq. 3.13 and the inequality (3.11), we have

\[
\|f_i^* h_i^{k-1} - f_i^* h^{k-1}\|_{L^2(M)} \rightarrow 0 \quad (i \rightarrow \infty). \tag{3.15}
\]

Thus we have \((I)_i \rightarrow 0\) as \(i \rightarrow \infty\).

By using the above argument inductively and the Markov property, we have

\[
\|h_i^1 - h^1\|_{L^2(M)} \\
= \left\| E_i^{T_i^{-1}(x)} \left( f_i^* g_1 \circ X_i(t_1) f_i^* g_2 \circ X_i(t_2) \cdots f_i^* g_k \circ X_i(t_k) \right) - E_i^x (g_1 \circ X(t_1) g_2 \circ X(t_2) \cdots g_k \circ X(t_k)) \right\|_{L^2(M)} \\
\rightarrow 0 \quad (i \rightarrow \infty). \tag{3.16}
\]

On the other hand, by the inequality (3.11), we have that

\[
\frac{d (f_i^* Vol_i)}{d Vol} \rightarrow 1_M \quad \text{uniformly}, \tag{3.17}
\]

where \(1_M\) means the indicator function on \(M\). By the fact (3.17) and (iii) of Assumption 3.13, we have

\[
\| \frac{d (f_i^* (\varphi_i Vol_i))}{d Vol} - \varphi \|_{L^2(M)} \rightarrow 0 \quad (i \rightarrow \infty). \tag{3.18}
\]
Thus, by Eqs. 3.16 and 3.18, using the Schwarz inequality, we have

$$
\left| \left( f_1^* g_1 \circ X_1(t_1) f_2^* g_2 \circ X_1(t_2) \cdots f_k^* g_k \circ X_1(t_k) \right) 
- \left( g_1 \circ X(t_1) g_2 \circ X(t_2) \cdots g_k \circ X(t_k) \right) \right| 
= \int_M E_{t}^{L^2(c)} \left( f_1^* g_1 \circ X_1(t_1) f_2^* g_2 \circ X_1(t_2) \cdots f_k^* g_k \circ X_1(t_k) \right) f_i^*(\psi_i \text{Vol})(dx)
- \int_M E_{t}^{L^2(c)} \left( g_1 \circ X(t_1) g_2 \circ X(t_2) \cdots g_k \circ X(t_k) \right) \psi(x) \text{Vol}(dx)
\to 0 \quad (i \to \infty).
$$

We therefore have shown (3.12) and we have completed the proof.

\[\square\]

4 Examples

4.1 Brownian Motions on Riemannian Manifolds

In this subsection, we consider the case when a state space $M$ is in $\mathcal{R} = \mathcal{R}(n, K, V, D)$ and a probability measure $P$ is the law of the Brownian motion, which is the Markov process induced by the Laplacian on $M$. In this case, the convergence of processes should follow only from the convergence of state spaces. We show that the convergence in $d_{L^P}$ follows only from the convergence in $d_L$.

Let $(M, g)$ be in $\mathcal{R}$. Let $\nabla$ denote the gradient operator induced by $g$. Let $(\mathcal{E}, \mathcal{F})$ be the smallest closed extension of the following bilinear form on $L^2(M; \text{Vol})$: (see e.g., [9, Example 5.7.2.])

$$
\mathcal{E}(u, v) = \frac{1}{2} \int_M g_x(\nabla u, \nabla v) \text{Vol}(dx) \quad (u, v \in C^\infty(M)).
$$

(4.1)

We write $\overline{\text{Vol}}(dx) = \text{Vol}(dx)/\text{Vol}(M)$.

**Definition 4.1** The set $\mathcal{L}\mathcal{R}(n, K, V, D)$ is defined to be the set of isomorphism classes of pairs $(M, P)$ where $M \in \mathcal{R}$ and $P$ is the law of $P^\mu$ for a Markov process on a time interval $[0, T]$ associated with $(\mathcal{E}, \mathcal{F})$ defined in Eq. 4.1 with an initial probability measure $\mu = \overline{\text{Vol}}_M$. We denote $\mathcal{L}\mathcal{R}$ shortly for $\mathcal{L}\mathcal{R}(n, K, V, D)$.

We show the relative compactness of $\mathcal{L}\mathcal{R}$.

**Proposition 4.2** The set $\mathcal{L}\mathcal{R}$ is relatively compact in $(\mathcal{P}M, d_{L^P})$.

*Proof of Proposition 4.2* Let $p_M(t, x, y)$ be the heat kernel of the standard energy form (4.1). It suffices to show that there is a jointly continuous function $\phi : (0, T] \times [0, D] \to [0, \infty)$ satisfying (3.8) and dominating $p_M(t, x, y)$ as Eq. 3.9. In fact, if we show this, we have

$$
\mathcal{L}\mathcal{R} \subset \mathcal{P}_\phi \mathcal{R}.
$$
Convergence of Continuous Stochastic Processes on Compact Metric

Since $P_\phi \mathcal{R}$ is relatively compact by Theorem 3.9, we obtain the desired result.

By e.g., [18, Theorem 7.4 & Proposition 7.5], we have that there exist positive constants, $C_1 = C_1(n, K, D)$ and $C_2 = C_2(n, K, D)$ depending only on $n, K, D$ such that
\[
p(t, x, y) \leq \frac{C_1}{m(B_\sqrt{t}(x))} \exp\left\{-C_2 \frac{d(x, y)^2}{t}\right\},
\]
for all $x, y \in X$ and $0 < t \leq D^2$. Here $B_r(x)$ means the metric ball centered at $x$ with radius $r$. By the Bishop–Gromov inequality (see, e.g., [4, Theorem 107]), we have the following volume growth estimate: there exist positive constants $\nu = \nu(n, K, D) > 0$ and $c = c(n, K, D) > 0$ depending only on $n, K, D$ such that, for all $n \in \mathbb{N}$
\[
m_n(B_r(x)) \geq cr^{2\nu}(0 \leq r \leq 1 \wedge D).
\]
Thus we have the following upper heat kernel estimate:
\[
p(t, x, y) \leq \frac{C_1}{ct^{\nu}} \exp\left\{-C_2 \frac{d(x, y)^2}{t}\right\},
\]
for all $x, y \in X$ and $0 < t \leq D^2$.

Thus setting
\[
\phi(\xi, r) = \frac{C_1}{c^{\xi\nu}} \exp\left(-C_2 \frac{r^2}{\xi}\right),
\]
we can check easily that $\phi$ satisfies (3.8).

Thus we have completed the proof.

Let $(M_i, \mathcal{P}_i) \in \mathcal{LR}$ and assume $M_i$ converges to some $M \in \overline{\mathcal{R}}^d_L$, where $\overline{\mathcal{R}}^d_L$ denotes the completion of $\mathcal{R}$ with respect to $d_L$. As stated in Section 3, the limit space $M$ has a $C^{1,\alpha}$-Riemannian structure. Such manifolds are in the class of Lipschitz–Riemannian manifolds (see, e.g., [13, §3]). In this framework, we have a Riemannian volume $\text{Vol}_M$ induced by the $C^{1,\alpha}$-Riemannian metric $g$ and the standard energy form $(\mathcal{E}, \mathcal{F})$ in Eq. 4.1. See the detail in [13, §3] and references therein.

Let $\mathcal{P}$ be the law of the Markov process on $M$ associated with the above $(\mathcal{E}, \mathcal{F})$ whose initial distribution is the Riemannian volume $\text{Vol}_M$. Then we have the following:

**Proposition 4.3** If $M_i$ converges to $M$ in $d_L$, then $(M_i, \mathcal{P}_i)$ converges to $(M, \mathcal{P})$ in $d_{LP}$.

**Proof** By Theorem 3.14, it is sufficient to check that Assumption 3.13 are satisfied. Since we assume that $M_i \to M$ in $d_L$, there is a family of $\varepsilon_i$-isometries $f_i : M_i \to M$ with $\varepsilon_i \to 0$ as $i \to \infty$. By the inequality (3.11), we can easily check that there is a function $h : [0, \infty) \to [0, \infty)$ with $\lim_{r \to 0} h(r) = 0$ satisfying the following inequalities (see [13, §3]): for any $u \in L^2(M)$ and $u_i \in L^2(M_i)$,
\[
\|f_i^* u\|_{L^2(M_i)} - \|u\|_{L^2(M)} \leq h(\varepsilon_i)\|u\|_{L^2(M)},
\]
\[
|\mathcal{E}_i(f_i^* u) - \mathcal{E}(u)| \leq h(\varepsilon_i)\mathcal{E}(u),
\]
\[
|\mathcal{E}_i(u_i) - \mathcal{E}(f_i^* u_i)| \leq h(\varepsilon_i)\mathcal{E}_i(u_i).
\]
By these inequalities, we can check easily that the conditions of Definition 3.11 are satisfied with $f_i$ (see [13, Proposition 3.1]). Since we take $\mu_i = \text{Vol}_{M_i}$ and $\mu = \text{Vol}_M$ in this section,
there is nothing to check about (iii) of Assumption 3.13. Thus the conditions in Assumption 3.13 are satisfied and we finish the proof.

By Proposition 4.3, we know what is the completion $\mathcal{L}^{dL}_{\mathcal{R}}$ of $\mathcal{LR}$ with respect to $d_{LP}$. We define the subset $\mathcal{L}^{dL}_{\mathcal{R}}$ consisting of pairs $(M, P)$ where $M \in \mathcal{R}^{dL}$ and $P$ is the law of $P^\mu$ where $P$ is the Brownian motion associated with the standard energy form $(\mathcal{E}, \mathcal{F})$ defined in Eq. 4.1 with the initial distribution $\mu = \text{Vol}_M$.

**Corollary 4.4** We have the following:

$$\mathcal{L}^{dL}_{\mathcal{R}} = \mathcal{L}^{dL}_{\mathcal{R}}.$$

**Proof** We first show $\mathcal{L}^{dL}_{\mathcal{R}} \subset \mathcal{L}^{dL}_{\mathcal{R}}$. Let $(M, P) \in \mathcal{L}^{dL}_{\mathcal{R}}$. Then we have a sequence $(M_i, P_i) \in \mathcal{L}^{dL}_{\mathcal{R}}$ such that $(M_i, P_i) \to (M, P)$ in $d_{LP}$. Thus $M_i \in \mathcal{R}$ converges to $M$ in $d_{L}$, and $M \in \mathcal{R}^{dL}$. By Proposition 4.3, we have that $P$ is the law of the Brownian motion associated with the standard energy form on $M$ with the initial distribution $\mu = \text{Vol}_M$. Thus $(M, P) \in \mathcal{L}^{dL}_{\mathcal{R}}$.

We second show $\mathcal{L}^{dL}_{\mathcal{R}} \supset \mathcal{L}^{dL}_{\mathcal{R}}$. Let $(M, P) \in \mathcal{L}^{dL}_{\mathcal{R}}$. Then we have a sequence $M_i \to M$ in $d_{L}$. Let $P_i$ be the law of the Brownian motion on $M_i$ associated with the standard energy form with the initial distribution $\mu = \text{Vol}_M$. Thus, by Proposition 4.3, we have that $(M_i, P_i) \to (M, P)$ in $d_{LP}$ and thus $(M, P) \in \mathcal{L}^{dL}_{\mathcal{R}}$.

### 4.2 Uniformly Elliptic Diffusions on Riemannian Manifolds

In this subsection, we consider $(M, P)$ where $(M, g) \in \mathcal{R}$ and $P$ is a law of a Markov process associated with another smooth Riemannian metric $h$ comparable to the given Riemannian metric $g$, that is, there is a constant $\Lambda > 1$ satisfying

$$\Lambda^{-1} g \leq h \leq \Lambda g. \quad (4.5)$$

The generator associated with $h$ is a second order differential operator having smooth coefficients with the uniform elliptic condition in local coordinates.

To be precise, let $\nabla_h$ denote the gradient operator induced by $h$ satisfying (4.5). Let $\text{Vol}_h$ be the volume measure associated with $h$. Let $(\mathcal{E}^h, \mathcal{F}^h)$ be the smallest closed extension of the following bilinear form on $L^2(M; \text{Vol}_h)$:

$$\mathcal{E}^h(u, v) = \frac{1}{2} \int_M h_x(\nabla_h u, \nabla_h v) \text{Vol}_h(dx) \quad (u, v \in C^\infty(M)). \quad (4.6)$$

We write $\text{Vol}_h(dx) = \text{Vol}_h(dx)/\text{Vol}_h(M)$.

**Definition 4.5** For $\Lambda > 1$, the set $\mathcal{L}_\Lambda \mathcal{R}(n, K, V, D)$ is defined to be the set of isomorphism classes of pairs $(M, P)$ where $(M, g) \in \mathcal{R}$ and $P$ is the law of $P^\mu$ for a Markov process on $[0, T]$ associated with $(\mathcal{E}^h, \mathcal{F}^h)$ defined in Eq. 4.6 for $h$ satisfying Eq. 4.5 with an initial probability measure $\mu = \text{Vol}_h$. We denote $\mathcal{L}_\Lambda \mathcal{R}$ shortly for $\mathcal{L}_\Lambda \mathcal{R}(n, K, V, D)$.

We show the relative compactness of $\mathcal{L}_\Lambda \mathcal{R}$.

**Proposition 4.6** The set $\mathcal{L}_\Lambda \mathcal{R}$ is relatively compact in $(\mathcal{P}, \mathcal{M}, d_{LP})$. 
Proof By e.g., [11, §4], we have the same heat kernel estimate as Eq. 4.4 for \((\mathcal{E}^h, \mathcal{F}^h)\). Note that, of course, the positive constant \(C_1, C_2, c\) in Eq. 4.4 depends also on \(\Lambda\) in this case. Thus the proof follows from the same argument of Proposition 4.2.

Acknowledgments The author thanks Prof. Kouji Yano for careful reading of his manuscript and successive encouragement. He thanks Prof. Takashi Kumagai for giving comments and detailed references in related fields. He thanks to Yohei Yamazaki for a lot of valuable and constructive comments and useful discussions. He also thanks to an anonymous referee for useful suggestions and references. This work was supported by Grant-in-Aid for JSPS Fellows Number 261798 and DAAD PAJAKO Number 57059240.

Appendix A

Recall that \(\mathcal{M}\) is the set of isometry classes of compact metric spaces and \(d_L\) is the Lipschitz distance on \(\mathcal{M}\) (see Section 2). We show the completeness of the metric space \((\mathcal{M}, d_L)\).

Proposition A.1 \((\mathcal{M}, d_L)\) is a complete metric space.

Proof Let \(\{X_i : i \in \mathbb{N}\}\) be a \(d_L\)-Cauchy sequence in \(\mathcal{M}\). It suffices to show that there are a compact metric space \(X \in \mathcal{M}\) and \(\varepsilon_i\)-isometries \(f_i : X_i \rightarrow X\) with \(\varepsilon_i \rightarrow 0\) as \(i \rightarrow \infty\).

The construction of \(X\): Let \(f_{ij} : X_i \rightarrow X_j\) be an \(\varepsilon_{ij}\)-isometry for \(i < j\) where \(\varepsilon_{ij} \rightarrow 0\) as \(i, j \rightarrow \infty\). Take a subsequence such that \(\varepsilon_{i,i+1} < 1/2^i\). Let \(\tilde{f}_{ij} : X_i \rightarrow X_j\) be defined by

\[
\tilde{f}_{ij} = f_{j-1,j} \circ f_{j-2,j-1} \circ \cdots \circ f_{i,i+1} \quad (i < j),
\]

and \(\tilde{\varepsilon}_{ij} = \sum_{l=i}^{j-1} \varepsilon_{l,l+1}\). Then \(\tilde{f}_{ij}\) is an \(\tilde{\varepsilon}_{ij}\)-isometry and \(\tilde{\varepsilon}_{ij} \rightarrow 0\) as \(i, j \rightarrow \infty\). Since every compact metric space is separable, there is a countable dense subset \(\{x^i_\alpha : \alpha \in \mathbb{N}\} \subset X_1\).

We define, for any \(i > 1,\)

\[
x^i_\alpha = \tilde{f}_{ii}(x^1_\alpha).
\]

Since \(\tilde{f}_{ii}\) is a homeomorphism, the subset \(\{x^i_\alpha : i \in \mathbb{N}\}\) is dense in \(X_i\) for each \(i\). Fix \(\alpha, \beta \in \mathbb{N}\), and consider the sequence of the real numbers

\[
\{d(x^i_\alpha, x^i_\beta) : i \in \mathbb{N}\}.
\]

Since \(\{\tilde{f}_{ii} : i \in \mathbb{N}\}\) has a bounded Lipschitz constant and the compact metric space \(X_1\) is bounded, we have that Eq. 2 is a bounded sequence:

\[
d(x^i_\alpha, x^i_\beta) \leq \sup_i d_{\text{il}}(\tilde{f}_{ii})d(x^1_\alpha, x^1_\beta) < \infty.
\]

Thus we can take a subsequence of Eq. 2 converging to some real number, write \(r(\alpha, \beta)\). We can check that \(r\) becomes a metric on \(\{\alpha : \alpha \in \mathbb{N}\}\). In fact, if \(r(\alpha, \beta) = 0\), we have \(\alpha = \beta\) because

\[
0 < \frac{1}{\sup_i d_{\text{il}}(\tilde{f}_{ii})} d(x^1_\alpha, x^1_\beta) \leq d(x^i_\alpha, x^i_\beta).
\]

By definition, \(r(\alpha, \alpha) = 0\) and \(r\) is symmetric and non-negative. It is easy to see the triangle inequality. Let \((X, d)\) be the completion of the metric space \((\{\alpha : \alpha \in \mathbb{N}\}, r)\). The compactness of \((X, d)\) will be shown later in this proof.

The construction of \(\varepsilon_i\)-isometries \(f_i\): We define a map \(f_i : \{x^i_\alpha : \alpha \in \mathbb{N}\} \rightarrow X\) by

\[
f_i(x^i_\alpha) = \alpha.
\]
Now we extend the map \( f_i \) to the whole space \( X_i \). Since \( \text{dil}(\tilde{f}_{ij}) \) is bounded, we have
\[
d(f_i(x^i_\alpha), f_i(x^i_\beta)) = d(\alpha, \beta) = \lim_{j \to \infty} d(x^i_\alpha, x^i_\beta) = \lim_{j \to \infty} d(\tilde{f}_{ij}(x^i_\alpha), \tilde{f}_{ij}(x^i_\beta)) \\
\leq \sup_j \text{dil}(\tilde{f}_{ij}) d(x^i_\alpha, x^i_\beta) < \infty. \tag{4}
\]
Let \( x^i_{\alpha(n)} \to x^i \) as \( n \to \infty \). By the inequality (4), we have that \( \lim_{n \to \infty} f_i(x^i_{\alpha(n)}) \) exists. This limit does not depend on the way of taking sequences converging to \( x^i \) (use the triangle inequality to check it). Thus we define
\[
f_i(x^i) = \lim_{n \to \infty} f_i(x^i_{\alpha(n)}),
\]
and this is well-defined. Thus we have extended the map \( f_i \) to the whole space \( X_i \).

Now we check that \( f_i \) is bi-Lipschitz. We have
\[
d(f_i(x^i_\alpha), f_i(x^i_\beta)) = d(\alpha, \beta) = \lim_{j \to \infty} d(x^i_\alpha, x^i_\beta) = \lim_{j \to \infty} d(\tilde{f}_{ij}(x^i_\alpha), \tilde{f}_{ij}(x^i_\beta)) \\
\geq \frac{1}{\sup_j \text{dil}(\tilde{f}_{ij}^{-1})} d(x^i_\alpha, x^i_\beta). \tag{5}
\]
Note that \( 0 < \sup_j \text{dil}(\tilde{f}_{ij}^{-1}) < \infty \). By the inequality (5), we see that \( f_i \) is bijective. By the inequality (4) and (5), we have \( f_i \) is bi-Lipschitz. Since \( f_i \) is a homeomorphism and \( X_i \) is compact, we see that \( X = f_i(X_i) \) is compact. Thus \( X \in \mathcal{M} \).

Finally we check that \( f_i \) is an \( \varepsilon_i \)-isometry for some \( \varepsilon_i \to 0 \) as \( i \to \infty \). We set
\[
\varepsilon_i = \max\{\log(\sup_j \text{dil}(\tilde{f}_{ij}^{-1})), \log(\sup_j \text{dil}(\tilde{f}_{ij}))\}.
\]
Then, by the inequality (4) and (5), we can see that \( f_i : X_i \to X \) is an \( \varepsilon_i \)-isometry with \( \varepsilon_i \to 0 \) as \( i \to \infty \). Thus we have shown that \( X \) is the \( d_L \)-limit of \( X_i \). We have completed the proof.

Note that, in the above proof, we have that
\[
f_j \circ \tilde{f}_{ij} = f_i. \tag{6}
\]
We use Eq. 6 in the proof of Theorem 2.9.

Remark A.2 Note that \((\mathcal{M}, d_L)\) is not separable. This is because of the following two facts:
(a) if \( d_L(X, Y) < \infty \), the Hausdorff dimensions of \( X \) and \( Y \) must coincide;
(b) for any non-negative real number \( d \), there is a compact metric space \( X \) whose Hausdorff dimension is equal to \( d \).

See, e.g., [7, Proposition 1.7.19] for (a) and [17] for (b). Let \( X \in \mathcal{M} \) and \( \mathcal{M}_X = \{Y \in \mathcal{M} : d_L(X, Y) < \infty\} \). We also note that there is a \( X \in \mathcal{M} \) such that even when we restrict \( d_L \) to \( \mathcal{M}_X \), the metric space \((\mathcal{M}_X, d_L)\) is not separable. See [22].

References
1. Aldous, D.: Stopping times and tightness. Ann. Probab. 6(2), 335–340 (1978)
2. Athreya, S., Lühr, W., Winter, A.: The gap between Gromov–vague and Gromov–Hausdorff–vague topology. Stochastic Process. Appl. 126(9), 2527–2553 (2016)
3. Athreya, S., Löhr, W., Winter, A.: Invariance principle for variable speed random walks on trees. Ann. Probab. 45(2), 625–667 (2017)
4. Berger, M.: A Panoramic View of Riemannian Geometry, 2nd edn. Springer-Verlag, Berlin (2007)
5. Billingsley, P.: Convergence of Probability Measures, vol. 493, 2nd edn. Wiley, New York (1999)
6. Blumenthal, R.M., Getoor, R.K.: Markov processes and Potential theory. Academic Press, New York (1968)
7. Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry, Graduate Studies in Mathematics, vol. 33. American Mathematical Society, Providence, RI, 2001. Errata list available online at http://www.pdmi.ras.ru/svivanov/papers/bbi-errata.pdf
8. Chen, Z.Q., Kim, P., Kumagai, T.: Discrete approximation of symmetric jump processes on metric measure spaces. Probab. Theory Relat. Fields 155, 703–749 (2013)
9. Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes, 2nd revised and extended edn. de Gruyter (2011)
10. Gromov, M.: Metric Structures for Riemannian and Non-Riemannian Spaces. Birkhäuser Boston Inc, Boston. Based on the 1981 French original [MR 85e:53051], With appendices by M. Katz, P. Pansu, and S. Semmes, Translated from the French by Sean Michael Bates (1999)
11. Kasue, A., Kumura, H.: Spectral convergence of Riemannian manifolds II. Tôhoku Math. J. 48, 71–120 (1996)
12. Katsuda, A.: Gromov’s convergence theorem and its application. Nagoya Math. J. 100, 11–48 (1985)
13. Kuwae, K., Shioya, T.: Convergence of spectral structures: a functional analytic theory and its applications to spectral geometry. Commun. Anal. Geom. 11, 599–673 (2003)
14. Mosco, U.: Approximation of the solutions of some variational inequalities. Ann. Scuola Normale Sup., Pisa 21, 373–394 (1967)
15. Mosco, U.: Composite media and asymptotic Dirichlet forms. J. Funct. Anal. 123, 368–421 (1994)
16. Ogura, Y.: Weak convergence of laws of stochastic processes on Riemannian manifolds. Probab. Theory Relat. Fields 119, 529–557 (2001)
17. Sharapov, S., Sharapov, V.: Dimensions of some generalized Cantor sets, preprint: http://classes.yale.edu/fracAndDim/cantorDims/CantorDims.html
18. Sturm, K.-T.: Diffusion processes and heat kernels on metric spaces. Ann. Probab. 26, 1–55 (1998)
19. Suzuki, K.: Convergences and projection Markov property of Markov processes on ultrametric spaces. RIMS J. 50(3), 569–588 (2014)
20. K. Suzuki: Convergence of brownian motion on metric measure spaces under riemannian curvature-dimension conditions, Preprint Available at: arXiv:1703.07234
21. Suzuki, K.: Convergence of non-symmetric diffusion processes on RCD spaces, Preprint Available at: arXiv:1709.09536
22. Suzuki, K., Yamazaki, Y.: Non-separability of the Lipschitz distance. Pacific J. Math. Ind. 7(1), 1–7 (2015)