The Annals of Probability
2012, Vol. 40, No. 5, 1945–1979
DOI: 10.1214/11-AOP676

CONVERGENCE OF TIME-INHOMOGENEOUS GEODESIC RANDOM WALKS AND ITS APPLICATION TO COUPLING METHODS

BY KAZUMASA KUWADA[1]

Ochanomizu University

We study an approximation by time-discretized geodesic random walks of a diffusion process associated with a family of time-dependent metrics on manifolds. The condition we assume on the metrics is a natural time-inhomogeneous extension of lower Ricci curvature bounds. In particular, it includes the case of backward Ricci flow, and no further a priori curvature bound is required. As an application, we construct a coupling by reflection which yields a nice estimate of coupling time, and hence a gradient estimate for the associated semigroups.

1. Introduction. It has been well known that there is a strong connection between behavior of heat distributions or Brownian motions and geometry of their underlying space. Even on time-inhomogeneous spaces such as Ricci flow, this guiding principle has been confirmed through recent developments (see [1, 8, 17–19, 26, 32] and references therein). Some of them [1, 18] are based on coupling methods of stochastic processes. Given two stochastic processes $Y_1(t)$ and $Y_2(t)$ on a state space M, a coupling $\mathbf{X}(t) = (X_1(t), X_2(t))$ of $Y_1(t)$ and $Y_2(t)$ is a stochastic process on $M \times M$ such that X_i has the same law as Y_i for $i = 1, 2$. By constructing a suitable coupling which reflects the geometry of the underlying structure, one can obtain various estimates for heat kernels, harmonic maps, eigenvalues etc. under natural geometric assumptions (see [12, 15, 30], e.g.). Since coupling of random variables provides a coupling of their distributions, coupling methods are naturally connected with the theory of optimal transportation, which are used in some of aforementioned results [19, 26]. With further studies in this direction in mind, here we consider an approximation of diffusion processes associated with a family of time-dependent metrics by so-called geodesic random walks. Generally speaking, one of the major reasons that we establish approximation is to overcome technical difficulties in studying the object in the limit. This is also our case, and we will use the approximation in order to study a coupling of diffusion processes.

Let M be a smooth manifold with a family of complete Riemannian metrics $g(t)$ indexed by $t \in [T_1, T_2]$. By $(X(t))_{t\in[T_1,T_2]}$, we denote the $g(t)$-Brownian motion.

Received July 2010; revised March 2011.

[1]Supported in part by the JSPS fellowship for research abroad and by KAKENHI (22740083).
MSC2010 subject classifications. Primary 60F17, 53C21; secondary 58J65, 53C44, 58J35.
Key words and phrases. Geodesic random walk, Ricci flow, diffusion process, coupling.

It means that $X(t)$ is a time-inhomogeneous diffusion process on M associated with $\Delta_{g(t)}/2$, where $\Delta_{g(t)}$ is the Laplacian with respect to $g(t)$ (see [8] for a construction of $g(t)$-Brownian motion). A geodesic random walk $\tilde{X}^\alpha$ on M with a parameter α is a discrete time Markov chain whose one-step variation is given as follows: Given a position x at some time t, consider a random vector in $T_x M$. We map it to M by $g(t)$-exponential map to determine the next position. Here the parameter α is implemented as a (diffusive) scaling on time step and on the length of the random vector in $T_x M$; see Section 3 for more details. In this paper, we consider only the case that all the random vectors in tangent space is specified to the one having a uniform distribution on a $g(t)$-ball whose radius is comparable to α. A simplified version of our main theorem, the convergence of geodesic random walks, is stated as follows; see Theorem 3.1 and Section 3 for a more precise and general statement:

THEOREM 1.1. *Suppose*

$$\partial_t g(t) \le \mathrm{Ric}_{g(t)} \tag{1.1}$$

holds. Then a continuous time interpolation of $\tilde{X}^\alpha$ converges in law to X as $\alpha \to 0$.

As we will see in the sequel, there are several technical difficulties arising from the time-dependency on the metric. Nevertheless, the assumption of the full statement in Theorem 3.1 is much weaker in some respect than that in the classical time-homogeneous case. Thus this assertion itself would be of interest, independently of its application to coupling methods.

In the time-homogeneous case, the convergence in law of scaled geodesic random walks to the Brownian motion is used to study a coupling of Brownian motions $(X_1(t), X_2(t))$ by reflection; see [16, 27]. A coupling of this kind provides us a useful control of the coupling time τ^*, the first time when X_1 and X_2 meet. Even in our time-inhomogeneous case, Theorem 1.1 carries the same estimate in almost the same way. A simplified version of this assertion is as follows; for the complete statement of our main theorem, see Theorem 4.1.

THEOREM 1.2. *Suppose* (1.1). *Then, for each $x_1, x_2 \in M$, there exists a coupling $\mathbf{X}(t) := (X_1(t), X_2(t))$ of two $g(t)$-Brownian motions starting at (x_1, x_2) satisfying*

$$\mathbb{P}[\tau^* > t] \le \mathbb{P}\left[\inf_{T_1 \le s \le t} B(s) > -\frac{d_{g(T_1)}(x_1, x_2)}{2}\right] \tag{1.2}$$

for each t, where $d_{g(T_1)}$ is the distance function on M with respect to $g(T_1)$, and $B(t)$ is a one-dimensional standard Brownian motion starting at the time T_1.

Similarly to the time-homogeneous case, Theorem 1.2 yields a gradient estimate of the heat semigroup, which tells us a quantitative estimate on the smoothing effect of the heat semigroup; see Corollary 4.3. In addition, we can apply our method to construct different kinds of couplings. As one of them, coupling by space–time parallel transport is studied in [18] by using Theorem 1.1, and it sharpens the result by Topping [26], concerning the monotonicity of a transportation cost between the heat distributions whose cost is measured by Perelman's $\mathscr{L}$-distance.

Condition (1.1) is essentially the same as backward super Ricci flow in [19]; our condition is slightly different in constant since our $g(t)$-Brownian motion, and hence our heat equation corresponds to $\Delta_{g(t)}/2$ instead of $\Delta_{g(t)}$. Obviously, (1.1) is satisfied if $g(t)$ evolves according to the backward Ricci flow $\partial_t g(t) = \mathrm{Ric}_{g(t)}$. From a different point of view, condition (1.1) can be interpreted as a time-inhomogeneous analog of nonnegative Ricci curvature since $\partial_t g(t)$ vanishes if $g(t)$ is independent of t. Along this viewpoint, we can consider a time-inhomogeneous analog of more general lower Ricci curvature bounds, and we can obtain the conclusion under such generalized conditions in the sequel; see Assumption 1 and (4.1); cf. Remark 4.2. It should be remarked that, even in those cases, no uniform lower bound of $\mathrm{Ric}_{g(t)}$ only in terms of $g(t)$ without time derivative is assumed. In particular, no bounds of $g(t)$ or $g(t)$-curvature tensor being uniform in time are required. Since a Ricci flow will produce a singularity in a finite time, a time-uniform bound on $g(t)$ or $\mathrm{Ric}_{g(t)}$ seems to be restrictive. It might be possible to simplify the proof by supposing additional assumptions involving a time-uniform estimate; however, this is out of the scope of this paper.

In our argument, the distance function $d_{g(t)}$ with respect to the time-dependent metric $g(t)$ plays a prominent role. The first variation in t of $d_{g(t)}$ is described in terms of $\partial_t g(t)$, and the second variation of $d_{g(t)}$ in space variables involves a notion of curvature. Both of these variations appear in the bounded variation part of the radial process $d_{g(t)}(o, X(t))$ of the $g(t)$-Brownian motion via the Itô formula. Thus a relation between $\partial_t g(t)$ and $\mathrm{Ric}_{g(t)}$ [e.g., (1.1)], produces a nice control of the radial process. Although we will work on geodesic random walks instead of the $g(t)$-Brownian motion itself, such an observation is still efficient.

In the time-homogeneous case, the convergence of scaled geodesic random walks is first studied by Jørgensen [14] by using the convergence theory of semigroups; see [6, 22], also. However, in our framework, it is not clear whether we can apply a similar technique since the base measure, the Riemannian volume, depends on time, and hence we cannot expect that it makes the heat semigroup invariant. To avoid such a technical difficulty, we use the uniqueness of the martingale problem instead for identifying the limit. Another difficulty arises from the lack of time-uniform bounds of Riemannian metrics. It prevents us to expect a *global* comparison of geometric structures, such as $d_{g(t)}$, between different times. Thus we will make some efforts for localizing the problem by giving a uniform estimate of the first exit time of X^α from a large ball centered at a reference point. Note that our assumption admits lower unbounded Ricci curvatures even in the case

$\partial_t g(t) \equiv 0$ (see Assumption 1). Thus our assumption on the geometry of the underlying space is weaker than that in [14] (by considering Riemannian manifolds with lower unbounded curvature, we can easily find an example which does not satisfy the assumption in [14]). On the other hand, the assumption on the driving noises of the geodesic random walk in [14] is more general than our specified one. Though it might be possible to take a more general noise under our assumption, our result already works well for applying the approximation to coupling methods. As a related work, the theory of time-dependent Dirichlet forms has been developed for studying the time-inhomogeneous Markov processes in the literature [20]; see also [24]. Unfortunately, because of above-mentioned difficulties, our framework does not fall into the scope of those theories at this moment. It might be an important problem to extend those theories so that they includes our case.

The organization of this paper is as follows. In the rest of this section after this paragraph, we review existing approaches on the construction of couplings. By comparing those approaches with ours, we try to explain the reason why we choose our approach for constructing a coupling by reflection. In the next section, we show basic properties of a family of Riemannian manifolds $((M, g(t)))_t$. In particular, we prove that Riemannian metrics $(g(t))_t$ are *locally* comparable with each other. It will be used to give a uniform control of several error terms which appear as a result of our discrete approximation. In Section 3, we will study geodesic random walks in our time-inhomogeneous framework. There we introduce them and prove the convergence in law to a diffusion process. After a small discussion at the beginning of the section, the proof is divided into two main parts. In the first part, we will give a uniform estimate for the exit time of geodesic random walks from a big compact set. Our assumption here is almost the same as in [17] where nonexplosion of the diffusion process is studied; see Remark 3.3 for more details. In the second part, we prove the tightness of geodesic random walks on the basis of the result in the first part. In Section 4, we will construct a coupling by reflection and show an estimate of coupling time, which completes the proof of Theorem 1.2 as a special case. In Section 5, we will give a short remark about how our method is also applicable to study a coupling by parallel transport.

1.1. *Existing arguments on coupling methods.* As stated above, we compare our method of the proof with existing arguments in coupling methods from a technical point of view. We hope that the following observation will be helpful to extend coupling arguments other than our own in this time-inhomogeneous case.

In order to go into details, let us review a heuristic (and common) idea of the construction of a coupling by reflection as well as that of the derivation of (1.2). Given a Brownian particle X_1, we will construct X_2 by determining its infinitesimal motion $dX_2(t) \in T_{X_2(t)}M$ by using $dX_1(t) \in T_{X_1(t)}M$. First we take a minimal $g(t)$-geodesic γ joining $X_1(t)$ and $X_2(t)$. Next, by using the parallel transport along γ associated with the $g(t)$-Levi–Civita connection, we bring $dX_1(t)$ into

$T_{X_2(t)}M$. Finally we define $dX_2(t)$ as a reflection of it with respect to a hyperplane being $g(t)$-perpendicular to $\dot{\gamma}$ in $T_{X_2(t)}M$. From this construction, the Itô formula implies that $d_{g(t)}(X_1(t), X_2(t))$ should become a semimartingale at least until $(X_1(t), X_2(t))$ hits the $g(t)$-cutlocus $\mathrm{Cut}_{g(t)}$. The semimartingale decomposition is given by variational formulas of arc length. On the bounded variation part, there appear the time-derivative of $d_{g(t)}$ and (a trace of) the second variation of $d_{g(t)}$, which is dominated in terms of the Ricci curvature. With the aid of our condition (1.1), these two terms are compensated and a nice domination of the bounded variation part follows. Thus the hitting time to 0 of $d_{g(t)}(X_1(t), X_2(t))$, which is the same as τ^*, can be estimated by that of the dominating semimartingale. Indeed, we can regard $2B(t) + d_{g(T_1)}(x_1, x_2)$ which appeared in the right-hand side of (1.2) as the dominating semimartingale. The effect of our reflection appears in the martingale part $2B(t)$ which makes it possible for the dominating semimartingale to hit 0. This construction seems to work as long as $(X_1(t), X_2(t))$ is not in the cutlocus. Moreover, if we succeed in constructing it beyond the cutlocus, then the same domination should hold. Indeed, the effect of singularity at the cutlocus should decrease $d_{g(t)}(X_1(t), X_2(t))$. Thus a "local time at the cutlocus" will be nonpositive, and hence negligible.

After this observation, we can conclude that almost all technical difficulties are concentrated on the treatment of singularity at the cutlocus in order to make the heuristic argument rigorous. In fact, Theorem 1.2 is shown in [21] by using SDE methods under the assumption that the $g(t)$-cutlocus is empty for every $t \in [T_1, T_2]$. It should be remarked that the joint distribution of the coupled particle $(X_1(t), X_2(t))$ could be singular to the Riemannian measure on $M \times M$ (at least it is the case when M is a flat Euclidean space). Thus it is not clear that the cutlocus is really "small" for the coupled particle despite the fact that the cutlocus (as a subset of $M \times M$) has null $g(t)$-Riemannian measure.

In our approach, we first construct a coupling of geodesic random walks and then take a limit to obtain the desired coupling. Since we first derive a dominating semimartingale for coupled geodesic random walks, we need only a *difference inequality* instead of the Itô formula. By virtue of this difference, we can obtain a desired estimate beyond the cutlocus by dividing a minimal geodesic joining particles into small pieces so that the endpoints of each piece are uniformly away from the cutlocus; see Lemma 4.4. As a result, we can avoid extracting a local time at the cutlocus and directly obtain a dominating process which does not involve such a term. Moreover, the dependency on time parameter of the cutlocus does not cause much difficulty in our approach.

In the time-homogeneous case, there are several arguments [9, 12, 28–30] to construct a coupling by reflection by approximating it with ones which move as mentioned above, if they are distant from the cutlocus and move independently if they are close to the cutlocus. In some of those arguments, we need to estimate the size of the total time when particles are close to the cutlocus. In such a case, an extension of these arguments to the time-inhomogeneous case does not seem straightforward since the $g(t)$-cutlocus depends on time and estimates should be more complicated. The argument in [30] uses supermartingales to extract the local time at the cutlocus in an implicit way, and no estimate of times spent around the cutlocus is necessary. Thus it seems possible to extend his argument in the time-inhomogeneous case. Since his argument relies on some detailed properties of parabolic PDEs, we need to develop time-inhomogeneous analogs of them to complete this plan. The fact that our assumption (1.1) [or (4.1)] does *not* imply any time-uniform lower bound of the Ricci curvature by a constant might be an obstacle.

If we employ the theory of optimal transportation, we will work on couplings of heat distributions instead of coupling of Brownian motions. Once we move to the world of heat distributions, we can expect that the cutlocus is treated more easily since they are of measure zero with respect to the Riemannian measure. However, at this moment, the theory of optimal transport is not so strong a tool in this context for the following two reasons. First, the range of the theory is restrictive in the sense that it only deals with couplings corresponding to the coupling by parallel transport. Second, the theory of optimal transportation provides a weaker result than a probabilistic approach does, even in studying couplings by parallel transport; for instance, see [19] and compare it with [1]. It should be remarked that such a difference between these two approaches exists even in the time-homogeneous case.

Arnaudon, Coulibaly and Thalmaier [1] recently developed a new method to construct a coupling, which works even in the time-inhomogeneous case. They consider a one-parameter family of coupled particles along a curve. Intuitively speaking, they concatenate coupled particles along a curve by iteration of making a coupling by parallel transport. Since “adjacent” particles are infinitesimally close to each other, we can ignore singularities on the cutlocus when we construct a coupled particle from an “adjacent” one. It should be noted that their method does not seem to be able to be applied directly in order to construct a coupling by reflection. Indeed, their construction of a chain of coupled particles heavily relies on a multiplicative (or semigroup) property of the parallel transport. However, our reflection operation obviously fails to possess such a multiplicative property. Since our reflection map changes orientation, there is no chance to interpolate it with a continuous family of isometries.

2. Properties on time-dependent metric. As in Section 1, let M be a m-dimensional manifold and $(g(t))_{t\in[T_1,T_2]}$ a family of complete Riemannian metrics on M which smoothly depends on t, for $-\infty < T_1 < T_2 < \infty$.

REMARK 2.1. It seems to be restrictive that our time parameter only runs over the compact interval $[T_1, T_2]$. An example of $g(t)$ we have in mind is a solution to the backward Ricci flow equation. In this case, we can work on a semi-infinite interval $[T_1, \infty)$ only when we study an ancient solution of the Ricci flow. Thus

$T_2 < \infty$ is not so restrictive. In addition, we could extend our results to the case on $[T_1, \infty)$ with a small modification of our arguments. It would be helpful to study an ancient solution. To deal with a singularity of Ricci flow, it could be nice to work on a semi-open interval $(T_1, T_2]$, where T_1 is the first time when a singularity emerges. In that case, we should be more careful since we cannot give "an initial condition at T_1" to define a $g(t)$-Brownian motion on M.

We collect some notation which will be used in the sequel. Throughout this paper, we fix a reference point $o \in M$. Let $\mathbb{N}_0$ be nonnegative integers. For $a, b \in \mathbb{R}$, $a \wedge b$ and $a \vee b$ stand for $\min\{a, b\}$ and $\max\{a, b\}$, respectively. Let $\mathrm{Cut}_{g(t)}(x)$ be the set of the $g(t)$-cutlocus of x on M. Similarly, the $g(t)$-cutlocus $\mathrm{Cut}_{g(t)}$ and the space–time cutlocus $\mathrm{Cut}_{\mathrm{ST}}$ are defined by

$$\mathrm{Cut}_{g(t)} := \{(x, y) \in M \times M \mid y \in \mathrm{Cut}_{g(t)}(x)\},$$
$$\mathrm{Cut}_{\mathrm{ST}} := \{(t, x, y) \in [T_1, T_2] \times M \times M \mid (x, y) \in \mathrm{Cut}_{g(t)}\}.$$

Set $D(M) := \{(x, x) | x \in M\}$. The distance function with respect to $g(t)$ is denoted by $d_{g(t)}(x, y)$. Note that $\mathrm{Cut}_{\mathrm{ST}}$ is closed and that $d_{g(\cdot)}(\cdot, \cdot)$ is smooth on $[T_1, T_2] \times M \times M \setminus (\mathrm{Cut}_{\mathrm{ST}} \cup [T_1, T_2] \times D(M))$; see [19]; cf. [17]. We denote an open $g(s)$-ball of radius R centered at $x \in M$ by $B_R^{(s)}(x)$. Some additional notation will be given at the beginning of the next section.

In the following three lemmas (Lemmas 2.2–2.4), we discuss a local comparison between $d_{g(t)}$ and $d_{g(s)}$ for $s \neq t$. Those will be a geometric basis of further arguments.

LEMMA 2.2. *Let M_0 be a compact subset of M. Then there exists $\kappa = \kappa(M_0)$ such that*

$$\mathrm{e}^{-2\kappa|t-s|} g(s) \le g(t) \le \mathrm{e}^{2\kappa|t-s|} g(s)$$

holds on M_0 for $t, s \in [T_1, T_2]$. In particular, if a minimal $g(s)$-geodesic γ joining $x, y \in M_0$ is included in M_0, then, for $t \in [T_1, T_2]$,

$$d_{g(t)}(x, y) \le \mathrm{e}^{\kappa|t-s|} d_{g(s)}(x, y).$$

PROOF. Let $\pi : TM \to M$ be a canonical projection. Let us define $\hat{M}_0$ by

$$\hat{M}_0 := \{(t, v) \in [T_1, T_2] \times TM \mid \pi(v) \in M_0, |v|_{g(t)} \le 1\}.$$

Note that $\hat{M}_0$ is closed since $g(\cdot)$ is continuous. We claim that $\hat{M}_0$ is sequentially compact. Let us take a sequence $((t_n, v_n))_{n\in\mathbb{N}} \subset \hat{M}_0$. We may assume $t_n \to t \in [T_1, T_2]$ and $\pi(v_n) \to p \in M_0$ as $n \to \infty$ by taking a subsequence if necessary. Let U be a neighborhood of p such that $\{v \in TM | \pi(v) \in U\} \simeq U \times \mathbb{R}^m$. For sufficiently large n, v_n is in $U \times \mathbb{R}^m$ and we write $v_n = (p_n, \tilde{v}_n)$. If we cannot take any convergent subsequence of $(v_n)_{n\in\mathbb{N}}$, then $|\tilde{v}_n| \to \infty$ as $n \to \infty$, where $|\cdot|$

stands for the standard Euclidean norm on $\mathbb{R}^m$ [irrelevant to $(g(t))_{t\in[T_1,T_2]}$]. Set $v'_n = (p_n, |\tilde{v}_n|^{-1}\tilde{v}_n)$. Then, there exists a subsequence $(v'_{n_k})_{k\in\mathbb{N}} \subset (v'_n)_{n\in\mathbb{N}}$ such that $v'_{n_k} \to v'_\infty = (p, \bar{v}')$ as $n \to \infty$ for some $\bar{v}' \in \mathbb{R}^m$ with $|\bar{v}'| = 1$. Since $g(\cdot)$ is continuous, $g(t_{n_k})(v'_{n_k}, v'_{n_k}) \to g(t)(v'_\infty, v'_\infty)$ as $k \to \infty$. On the other hand, $g(t_{n_k})(v'_{n_k}, v'_{n_k}) \le |\tilde{v}_{n_k}|^{-2} \to 0$ since $g(t_n)(v_n, v_n) \le 1$. Thus $\bar{v}'$ must be 0. It contradicts with $|\bar{v}'| = 1$. Hence $\hat{M}_0$ is sequentially compact.

Since $\hat{M}_0 \ni (t, v) \mapsto \partial_t g(t)(v, v)$ is continuous, there exists a constant $\kappa = \kappa(M_0) > 0$ such that $|\partial_t g(t)(v, v)| \le 2\kappa$ for every $(t, v) \in \hat{M}_0$. Take $v \in \pi^{-1}(M_0)$, $v \neq 0_{\pi(v)}$. Then

$$\partial_t g(t)(v, v) = |v|^2_{g(t)} \partial_t g(t)\big(|v|^{-1}_{g(t)} v, |v|^{-1}_{g(t)} v\big) \le 2\kappa |v|^2_{g(t)}.$$

Thus $\partial_t \log g(t)(v, v) \le 2\kappa$ holds. By integrating it from s to t with $s < t$, we obtain $g(t)(v, v) \le e^{2\kappa(t-s)} g(s)(v, v)$. We can obtain the other inequality similarly.

For the latter assertion, for a, b with $\gamma(a) = x$ and $\gamma(b) = y$,

$$\begin{aligned} d_{g(t)}(x, y) &\le \int_a^b |\dot{\gamma}(u)|_{g(t)}\, du \le e^{\kappa|t-s|} \int_a^b |\dot{\gamma}(u)|_{g(s)}\, du \\ &= e^{\kappa|t-s|} d_{g(s)}(x, y). \end{aligned}$$ □

LEMMA 2.3. *For $R > 0$, $x \in M$ and $t \in [T_1, T_2]$, there exists $\delta = \delta(x, t, R) > 0$ such that $\bar{B}^{(s)}_r(x) \subset \bar{B}^{(t)}_{3r}(x)$ for $r \le R$ and $s \in [T_1, T_2]$ with $|s - t| \le \delta$.*

PROOF. Set $\kappa := \kappa(\bar{B}^{(t)}_{3R}(x))$ as in Lemma 2.2 and $\delta := \kappa^{-1} \log 2$. Take $p \in \bar{B}^{(s)}_r(x)$ and a minimal $g(s)$-geodesic $\gamma : [a, b] \to M$ joining x and p. Suppose that there exists $u_0 \in [a, b]$ such that $\gamma(u_0) \in \bar{B}^{(t)}_{3r}(x)^c$. Let $\bar{u}_0 := \inf\{u \in [a, b] | \gamma(u) \in \bar{B}^{(t)}_{3r}(x)^c\}$. Since $\gamma([a, \bar{u}_0]) \subset \bar{B}^{(t)}_{3r}(x) \subset \bar{B}^{(t)}_{3R}(x)$ and $d_{g(t)}(x, \gamma(\bar{u}_0)) = 3r$, Lemma 2.2 yields

$$d_{g(s)}(x, p) \ge \int_a^{\bar{u}_0} |\dot{\gamma}(u)|_{g(s)}\, du \ge e^{-\kappa\delta} \int_a^{\bar{u}_0} |\dot{\gamma}(u)|_{g(t)}\, du \ge \frac{3r}{2}.$$

This is absurd. Hence $\gamma([a, b]) \in \bar{B}^{(t)}_{3r}(x)$. In particular, $\gamma(b) = p \in \bar{B}^{(t)}_{3r}(x)$. □

LEMMA 2.4. *For $R > 0$, there exists a compact subset $M_0 = M_0(R)$ of M such that*

$$\Big\{p \in M \Big| \inf_{t\in[T_1,T_2]} d_{g(t)}(o, p) \le R\Big\} \subset M_0. \tag{2.1}$$

PROOF. For each $t \in [T_1, T_2]$, take $\delta(o, t, R + 1) > 0$ according to Lemma 2.3. Take $\{t_i\}_{i=1}^n \subset [T_1, T_2]$ such that

$$[T_1, T_2] \subset \bigcup_{i=1}^n \big(t_i - \delta(o, t_i, R + 1), t_i + \delta(o, t_i, R + 1)\big).$$

Let us define a compact set $M_0 \subset M$ by $M_0 := \bigcup_{i=1}^n \bar{B}_{3R}^{(t_i)}(o)$. Take $p \in M$ such that $\inf_{T_1 \le t \le T_2} d_{g(t)}(o, p) \le R$. For $\varepsilon \in (0, 1)$, take $s \in [T_1, T_2]$ such that $d_{g(s)}(o, p) \le R + \varepsilon$. Then there exists $j \in \{1, \dots, N\}$ such that $|s - t_j| < \delta(o, t_j, R + 1)$. By Lemma 2.3, it implies $p \in \bar{B}_{R+\varepsilon}^{(s)}(o) \subset \bar{B}_{3(R+\varepsilon)}^{(t_j)}(o) \subset \bigcup_{i=1}^n \bar{B}_{3(R+\varepsilon)}^{(t_i)}(o)$. Hence the conclusion follows by letting $\varepsilon \downarrow 0$. $\square$

Another useful consequence of Lemmas 2.2 and 2.3 is the following:

LEMMA 2.5. *$d_{g(\cdot)}(\cdot, \cdot)$ is continuous on $[T_1, T_2] \times M \times M$.*

PROOF. Since the topology on $[T_1, T_2] \times M \times M$ is metrizable, it suffices to show $\lim_{n\to\infty} d_{g(t_n)}(x_n, y_n) = d_{g(t)}(x, y)$ when $(t_n, x_n, y_n) \to (t, x, y)$ as $n \to \infty$. By the triangle inequality,

$$\begin{aligned} \left|d_{g(t_n)}(x_n, y_n) - d_{g(t)}(x, y)\right| &\le \left|d_{g(t_n)}(x, y) - d_{g(t)}(x, y)\right| \\ &\quad + d_{g(t_n)}(x, x_n) + d_{g(t_n)}(y, y_n). \end{aligned} \tag{2.2}$$

Take $R > 0$ so that $B_R^{(t)}(x)$ includes a minimal $g(t)$-geodesic joining x and y. Take $\kappa = \kappa(\bar{B}_{4R}^{(t)}(x))$ according to Lemma 2.2. We can easily see that every minimal $g(t)$-geodesic joining y and y_n is included in $B_{2R}^{(t)}(x)$ for sufficiently large $n \in \mathbb{N}$. Thus Lemma 2.2 yields

$$\limsup_{n\to\infty} d_{g(t_n)}(y, y_n) \le \limsup_{n\to\infty} \mathrm{e}^{\kappa|t-t_n|} d_{g(t)}(y, y_n) = 0.$$

We can show $d_{g(t_n)}(x, x_n) \to 0$ similarly. Take a minimal $g(t_n)$-geodesic $\gamma_n : [a, b] \to M$ joining x and y. By our choice of R, Lemma 2.2 again yields

$$d_{g(t_n)}(x, \gamma_n(u)) \le d_{g(t_n)}(x, y) \le \mathrm{e}^{\kappa|t-t_n|} d_{g(t)}(x, y) \le \mathrm{e}^{\kappa|t-t_n|} R.$$

It implies $\limsup_{n\to\infty} d_{g(t_n)}(x, y) \le d_{g(t)}(x, y)$. In addition, γ_n is included in $B_{4R/3}^{(t_n)}(x)$ for sufficiently large n. Thus Lemmas 2.3 and 2.2 yield $d_{g(t)}(x, y) \le \mathrm{e}^{\kappa|t-t_n|} d_{g(t_n)}(x, y)$. Hence the conclusion follows by combining these estimates with (2.2). $\square$

Before closing this section, we will provide a local lower bound of the injectivity radius which is uniform in time parameter.

LEMMA 2.6. *For every $M_1 \subset M$ compact, there is $\tilde{r}_0 = \tilde{r}_0(M_1) > 0$ such that $d_{g(t)}(y, z) < \tilde{r}_0$ implies $(t, y, z) \notin \mathrm{Cut}_{\mathrm{ST}}$ for any $(t, y, z) \in [T_1, T_2] \times M_1 \times M_1$.*

PROOF. Take $R > 1$ so that $\sup_{t\in[T_1,T_2]} \sup_{x\in M_1} d_{g(t)}(o, x) < R - 1$. By Lemma 2.4, there exists a compact set $M_0 \subset M$ such that (2.1) holds. For every $t \in [T_1, T_2]$ and $x \in M_1$, $(t, x, x) \notin \mathrm{Cut}_{\mathrm{ST}}$. It implies that there is $\eta_{t,x} \in (0, 1)$

such that $(s,y,z) \notin \mathrm{Cut}_{\mathrm{ST}}$ whenever $d_{g(t)}(x,y) \vee d_{g(t)}(x,z) \vee |t-s| < \eta_{t,x}$ since $\mathrm{Cut}_{\mathrm{ST}}$ is closed. Thus there exist $N \in \mathbb{N}$ and $(t_i, x_i) \in [T_1, T_2] \times M_1$ $(i = 1, \dots, N)$ such that

$$[T_1, T_2] \times M_1 \subset \bigcup_{i=1}^{N} \left(t_i - \frac{\eta_{t_i,x_i}}{2}, t_i + \frac{\eta_{t_i,x_i}}{2} \right) \times B^{(t_i)}_{\eta_{t_i,x_i}/2}(x_i).$$

Set $\tilde{r}_0 > 0$ by

$$\tilde{r}_0 := \frac{1}{2} \exp\left(-\frac{\kappa}{2} \max_{1 \le i \le N} \eta_{t_i,x_i} \right) \min_{1 \le i \le N} \eta_{t_i,x_i},$$

where $\kappa = \kappa(M_0) > 0$ is as in Lemma 2.2. Take $(s,y,z) \in [T_1,T_2] \times M_1 \times M_1$ with $d_{g(s)}(y,z) < \tilde{r}_0$. Take $j \in \{1, \dots, N\}$ so that $|s - t_j| \vee d_{g(t_j)}(x_j, y) < \eta_{t_j,x_j}/2$. By virtue of the choice of R and M_0, Lemma 2.4 yields that every $g(s)$-geodesic joining y and z is included in M_0. Thus Lemma 2.2 yields

$$d_{g(t_j)}(y,z) \le \mathrm{e}^{\kappa|s-t_j|} d_{g(s)}(y,z) < \frac{\eta_{t_j,x_j}}{2}.$$

It implies $|s-t_j| \vee d_{g(t_j)}(x_j,y) \vee d_{g(t_j)}(x_j,z) < \eta_{t_j,x_j}$ and hence $(s,y,z) \notin \mathrm{Cut}_{\mathrm{ST}}$. $\square$

3. Approximation via geodesic random walks. Let $(Z(t))_{t \in [T_1,T_2]}$ be a family of smooth vector fields continuously depending on the parameter $t \in [T_1, T_2]$. Let $X(t)$ be the diffusion process associated with the time-dependent generator $\mathscr{L}_t = \Delta_{g(t)}/2 + Z(t)$; see [8] for a construction of $X(t)$ by solving a SDE on the frame bundle. Note that $(t, X(t))$ is a unique solution to the martingale problem associated with $\partial_t + \mathscr{L}_\cdot$ on $[T_1, T_2] \times M$; see [12] for the time-homogeneous case. Its extension to time-inhomogeneous case is straightforward; see [25] also.

In what follows, we will use several notions in Riemannian geometry such as exponential map exp, Levi–Civita connection ∇, Ricci curvature Ric etc. To clarify the dependency on the metric $g(t)$, we put (t) on superscript or $g(t)$ on subscript. For instance, we use the following symbols: $\exp^{(t)}$, $\nabla^{(t)}$ and $\mathrm{Ric}_{g(t)}$. We refer to [7] for basics in Riemannian geometry which will be used in this paper.

For each $t \in [T_1, T_2]$, we fix a measurable section $\Phi^{(t)} : M \to \mathscr{O}^{(t)}(M)$ of the $g(t)$-orthonormal frame bundle $\mathscr{O}^{(t)}(M)$ of M. Take a sequence of independent, identically distributed random variables $\{\xi_n\}_{n \in \mathbb{N}}$ which are uniformly distributed on the unit disk in $\mathbb{R}^m$. Given $x_0 \in M$, let us define a continuously-interpolated geodesic random walk $(X^\alpha(t))_{t \in [T_1,T_2]}$ on M starting from x_0 with a scale parameter $\alpha > 0$ inductively. Let $t_n^{(\alpha)} := (T_1 + \alpha^2 n) \wedge T_2$ for $n \in \mathbb{N}_0$. For $t = T_1 = t_0^{(\alpha)}$, set $X^\alpha(T_1) := x_0$. After $X^\alpha(t)$ is defined for $t \in [T_1, t_n^{(\alpha)}]$, we extend it to $t \in [t_n^{(\alpha)}, t_{n+1}^{(\alpha)}]$ by

$$\tilde{\xi}_{n+1} := \sqrt{m+2}\,\Phi^{(t_n^{(\alpha)})}\big(X^\alpha(t_n^{(\alpha)})\big)\xi_{n+1},$$

$$X^\alpha(t) := \exp^{(t_n^{(\alpha)})}_{X^\alpha(t_n^{(\alpha)})}\left(\frac{t - t_n^{(\alpha)}}{\alpha^2} \big(\alpha \tilde{\xi}_{n+1} + \alpha^2 Z(t_n^{(\alpha)})\big) \right).$$

For later use, we define $N^{(\alpha)} := \inf\{n \in \mathbb{N}_0 | t_{n+1}^{(\alpha)} - t_n^{(\alpha)} < \alpha^2\}$. This is the total number of discrete steps of our geodesic random walks with scale parameter α. Set $\mathscr{C} := C([T_1, T_2] \to M)$ and $\mathscr{D} := D([T_1, T_2] \to M)$, the space of right continuous paths on M parametrized with $[T_1, T_2]$ possessing a left limit at every point. By using a distance $d_{g(T_1)}$ on M, we metrize $\mathscr{C}$ and $\mathscr{D}$ as usual so that $\mathscr{C}$ and $\mathscr{D}$ become Polish spaces; see [10] for a distance function on $\mathscr{D}$, for example. Set $\mathscr{C}_1 := C([T_1, T_2] \to [0, \infty))$. Let us define a time-dependent $(0, 2)$-tensor field $(\nabla Z(t))^\flat$ by

$$(\nabla Z(t))^\flat(X, Y) := \frac{1}{2}\bigl(\langle \nabla_X^{(t)} Z(t), Y\rangle_{g(t)} + \langle \nabla_Y^{(t)} Z(t), X\rangle_{g(t)}\bigr).$$

ASSUMPTION 1. There exists a locally bounded nonnegative measurable function b on $[0, \infty)$ such that:

(i) For all $t \in [T_1, T_2)$,

$$2(\nabla Z(t))^\flat + \partial_t g(t) \le \mathrm{Ric}_{g(t)} + b\bigl(d_{g(t)}(o, \cdot)\bigr) g(t).$$

(ii) For each $C > 0$, a one-dimensional diffusion process y_t given by

$$dy_t = d\beta_t + \frac{1}{2}\left(C + \int_0^{y_t} b(s)\, ds\right) dt,$$

where β_t is a standard Brownian motion, does not explode. (This is the case if and only if

$$\int_1^\infty \exp\left(-\int_1^y \mathbf{b}(z)\, dz\right) \int_1^y \exp\left(\int_1^z \mathbf{b}(\xi)\, d\xi\right) dz\, dy = \infty,$$

where $\mathbf{b}(y) := C + \int_0^y b(s)\, ds$; see, e.g., [13], Theorem VI.3.2.)

Note that (1.1) is a special case of Assumption 1. Now, we are in position to state the main theorem of this paper.

THEOREM 3.1. *Under Assumption 1, X^α converges in law to X in $\mathscr{C}$ as $\alpha \to 0$.*

Most of arguments in this section will be devoted to show the tightness, that is:

PROPOSITION 3.2. *$(X^\alpha)_{\alpha \in (0,1)}$ is tight in $\mathscr{C}$.*

In fact, as we will see in the following, Proposition 3.2 easily implies Theorem 3.1.

PROOF OF THEOREM 3.1. By virtue of Proposition 3.2, for any subsequence of $(X^\alpha)_{\alpha \in (0,1)}$ there exists a further subsequence $(X^{\alpha_k})_{k \in \mathbb{N}}$ which converges in law in $\mathscr{C}$ as $k \to \infty$. Thus it suffices to show that this limit has the same law as X.

Let $(\beta^\alpha(t))_{t\in[0,\infty)}$ be a Poisson process of intensity α^{-2} which is independent of $\{\xi_n\}_{n\in\mathbb{N}}$. Set

$$\bar{\beta}^\alpha(t) := \big(T_1 + \alpha^2\beta^\alpha(t - T_1)\big) \wedge t^{(\alpha)}_{N^{(\alpha)}}.$$

Then the Poisson subordination $X^{\alpha_k}(\bar{\beta}^{\alpha_k}(\cdot))$ also converges in law in $\mathscr{D}$ to the same limit; see [5], for instance. Note that $(\bar{\beta}^\alpha(t), X^\alpha(\bar{\beta}^\alpha(t)))_{t\in[T_1,T_2]}$ is a time-inhomogeneous Markov process. The associated semigroup $P^{(\alpha)}_t$ and its generator $\tilde{\mathscr{L}}^{(\alpha)}$ are given by

$$\begin{aligned} P^{(\alpha)}_t f &:= \mathrm{e}^{-(t-T_1)\alpha^{-2}}\Bigg(\sum_{l=1}^{N^{(\alpha)}} \frac{((t-T_1)\alpha^{-2})^l}{l!}(q^{(\alpha)})^l f \\ &\qquad\qquad + \sum_{l>N^{(\alpha)}} \frac{((t-T_1)\alpha^{-2})^l}{l!}(q^{(\alpha)})^{N^{(\alpha)}} f\Bigg), \\ \tilde{\mathscr{L}}^{(\alpha)} f &:= \alpha^{-2}\big(q^{(\alpha)} f - f\big), \end{aligned}$$

where

$$q^{(\alpha)} f(t,x) := \mathbb{E}\big[f\big(t+\alpha^2, \exp^{(t)}_x\big(\alpha\sqrt{m+2}\Phi^{(t)}(x)\xi_1 + \alpha^2 Z(t)\big)\big)\big].$$

We can easily prove $\tilde{\mathscr{L}}^{(\alpha)} f \to (\partial_t + \mathscr{L}_\cdot) f$ uniformly as $\alpha \to 0$ for $f \in C^\infty_0([T_1,T_2]\times M)$. Since $(\bar{\beta}^\alpha(t), X^\alpha(\bar{\beta}^\alpha(t)))_{t\in[T_1,T_2]}$ is a solution to the martingale problem associated with $\tilde{\mathscr{L}}^{(\alpha)}$, the limit in law of $(\bar{\beta}^{\alpha_k}(\cdot), X^\alpha(\bar{\beta}^{\alpha_k}(\cdot)))$ solves the martingale problem associated with $\partial_t + \mathscr{L}_\cdot$. By the uniqueness of the martingale problem, this limit has the same law as that of $(t, X(t))_{t\in[T_1,T_2]}$. It completes the proof. □

REMARK 3.3. Proposition 3.2 also asserts that any subsequential limit in law is a probability measure on $\mathscr{C}$. Since we have not added any cemetery point to M in the definition of $\mathscr{C}$, Theorem 3.1 implies that X cannot explode. It almost recovers the result in [17]. Our assumption is slightly stronger than that in [17] on the point where we require (ii) for *all* $C > 0$, not a given constant. Note that we will use Assumption 1(ii) only for a specified constant $2C_0$ given in Lemma 3.9. However, its expression looks complicated, and it seems to be less interesting to provide an explicit bound.

Now we introduce some additional notation which will be used in the rest of this paper. For $t \in [T_1, T_2]$, we define $\lfloor t\rfloor_\alpha$ by

$$\lfloor t\rfloor_\alpha := \sup\{\alpha^2 n + T_1 \,|\, n \in \mathbb{N}_0, \alpha^2 n + T_1 < t\}.$$

Set $\mathscr{F}_n := \sigma(\xi_1, \dots, \xi_n)$. For $R > 1$, let us define $\sigma_R : \mathscr{C}_1 \to [T_1, T_2] \cup \{\infty\}$ by

$$\sigma_R(w) := \inf\{t \in [T_1, T_2] \,|\, w(t) > R - 1\},$$

where $\inf\varnothing = \infty$. We write $\hat{\sigma}_R := \sigma_R(d_{g(\cdot)}(o, X^\alpha(\cdot)))$ and $\bar{\sigma}_R := \alpha^{-2}(\lfloor\hat{\sigma}_R\rfloor_\alpha - T_1) + 1$. Note that $\bar{\sigma}_R$ is an $\mathscr{F}_n$-stopping time. For each $t \in [T_1, T_2]$ and $x, y \in M$ with $x \neq y$, we choose a minimal unit-speed $g(t)$-geodesic $\gamma_{xy}^{(t)} : [0, d_{g(t)}(x, y)] \to M$ from x to y. Note that we can choose $\gamma_{xy}^{(t)}$ so that $(x, y) \mapsto \gamma_{xy}^{(t)}$ is measurable in an appropriate sense; see, for example, [27]. We use the same symbol $\gamma_{xy}^{(t)}$ for its range $\gamma_{xy}^{(t)}([0, d_{g(t)}(x, y)])$.

3.1. *A uniform bound for the escape probability.* The goal of this subsection is to show the following:

PROPOSITION 3.4. $\lim_{R\uparrow\infty} \limsup_{\alpha\downarrow 0} \mathbb{P}[\hat{\sigma}_R \le T_2] = 0$.

For the proof, we will establish a discrete analog of a comparison argument for the radial process as discussed in [17]. From now on, we fix $R > 1$ sufficiently large so that $d_{g(T_1)}(o, x_0) < R - 1$ until the final line of the proof of Proposition 3.4. We also fix a compact set $M_0 \subset M$ satisfying (2.1). Set $r_0 := \tilde{r}_0 \wedge (1/2)$, where $\tilde{r}_0 = \tilde{r}_0(M_0)$ is as in Lemma 2.6.

The first step for proving Proposition 3.4 is to show a difference inequality for the radial process $d_{g(t)}(o, X^\alpha(t))$ (Lemma 3.7). It will play the role of the Itô formula for the radial process in our discrete setting. We introduce some notation to discuss how to avoid the singularity of $d_{g(\cdot)}(o, \cdot)$ on $\{o\} \cup \mathrm{Cut}_{g(\cdot)}(o)$. For $r > 0$, let us define a set A_r', A_r'' and A_r as follows:

$$\begin{aligned} A_r' &:= \bigl\{(t, x, y) \in [T_1, T_2] \times M_0 \times M_0 \,|\, d_{g(t)}(x, x') + d_{g(t)}(y, y') + |t - t'| \ge r \\ &\qquad\qquad \text{for any } (t', x', y') \in \mathrm{Cut}_{\mathrm{ST}}\bigr\}, \\ A_r'' &:= \bigl\{(t, x, y) \in [T_1, T_2] \times M_0 \times M_0 \,|\, d_{g(t)}(x, y) \ge r\bigr\}, \\ A_r &:= A_r' \cap A_r''. \end{aligned}$$

Note that A_r is compact and that $d_{g(\cdot)}(\cdot, \cdot)$ is smooth on A_r. For $t \in [T_1, T_2]$ and $p \in M$, let us define $o_p^{(t)} \in M_0$ by

$$o_p^{(t)} := \begin{cases} \gamma_{op}^{(t)}\left(\dfrac{r_0}{2}\right), & \text{if } (t, o, p) \notin A_{r_0}', \\ o, & \text{otherwise.} \end{cases}$$

For simplicity of notation, we denote $o_{X^\alpha(t_n^{(\alpha)})}^{(t_n^{(\alpha)})}$ by o_n. Similarly, we use the symbol γ_n for $\gamma_{o_n X^\alpha(t_n^{(\alpha)})}^{(t_n^{(\alpha)})}$ throughout this section. Note that $(t, o_p^{(t)}, p) \notin \mathrm{Cut}_{\mathrm{ST}}$ holds. Furthermore, it is uniformly separated from $\mathrm{Cut}_{\mathrm{ST}}$ in the following sense:

LEMMA 3.5. *There exist $r_1 > 0$ and $\delta_1 > 0$ such that the following holds: let $t_0, t \in [T_1, T_2]$ with $t - t_0 \in [0, \delta_1]$. Let $p_0 \in B_{R-1}^{(t_0)}(o)$ and $p \in B_{\delta_1}^{(t_0)}(p_0)$. Then we have*:

(i) $d_{g(t)}(o, p) \le \mathrm{e}^{\kappa(t-t_0)}(d_{g(t_0)}(o, p_0) + d_{g(t_0)}(p_0, p))$;

(ii) $(t, o_{p_0}^{(t_0)}, p) \in A_{r_1}$ *when* $p_0 \notin B_{r_0}^{(t_0)}(o)$.

Here $\kappa = \kappa(M_0) > 0$ *is given according to Lemma* 2.2.

By applying Lemma 3.5 to X^α, we obtain the following:

COROLLARY 3.6. *There exist* $\alpha_0 > 0$ *and* $h : [0, \alpha_0] \to [0, 1]$ *with* $\lim_{\alpha \downarrow 0} h(\alpha) = 0$ *such that the following holds*: *for* $\alpha \le \alpha_0$, $n \in \mathbb{N}_0$ *and* $s, t \in [t_n^{(\alpha)}, t_{n+1}^{(\alpha)}]$, *when* $n < \bar{\sigma}_R$:

(i) $d_{g(t)}(o, X^\alpha(s)) \le \mathrm{e}^{\kappa\alpha^2}(d_{g(t_n^{(\alpha)})}(o, X^\alpha(t_n^{(\alpha)})) + h(\alpha))$;

(ii) $(t, o_n, X^\alpha(s)) \in A_{r_1}$ *when* $X^\alpha(t_n^{(\alpha)}) \notin B_{r_0}^{(t_n^{(\alpha)})}(o)$.

Here r_1 *is the same as in Lemma* 3.5.

PROOF. Set $\bar{Z} := \sup_{t \in [T_1, T_2], x \in M_0} |Z(t)|_{g(t)}(x)$. Note that we have

$$d_{g(t_n^{(\alpha)})}\big(X^\alpha\big(t_n^{(\alpha)}\big), X^\alpha(t)\big) \le \sqrt{m+2}\alpha + \bar{Z}\alpha^2$$

by the definition of X^α. Take $\alpha_0 > 0$ so that $\sqrt{m+2}\alpha_0 + \bar{Z}\alpha_0^2 \le \delta_1$ and $\alpha^2 \le \delta_1$ hold, where δ_1 is as in Lemma 3.5. Then the conclusion follows by applying Lemma 3.5 with $t_0 = t_n^{(\alpha)}$, $p_0 = X^\alpha(t_n^{(\alpha)})$ and $p = X^\alpha(s)$. $\square$

PROOF OF LEMMA 3.5. We show that (i) holds with $\delta_1 = 1$. By the triangle inequality, the proof is reduced to showing the following two inequalities:

$$d_{g(t)}(o, p_0) \le \mathrm{e}^{\kappa(t-t_0)} d_{g(t_0)}(o, p_0); \tag{3.1}$$

$$d_{g(t)}(p_0, p) \le \mathrm{e}^{\kappa(t-t_0)} d_{g(t_0)}(p_0, p). \tag{3.2}$$

Our condition (2.1) yields that $\gamma_{op_0}^{(t_0)}$ is included in M_0. Thus Lemma 2.2 yields (3.1). When $p \in B_1^{(t_0)}(p_0)$, we have $\gamma_{p_0 p}^{(t_0)} \subset B_R^{(t_0)}(o)$. Hence (2.1) and Lemma 2.2 yield (3.2) in a similar way as (3.1).

Let us consider (ii). For simplicity of notation, we denote $o_{p_0}^{(t_0)}$ by o' in this proof. We assume that $t - t_0 \in [0, \delta]$ and $p \in B_\delta^{(t_0)}(p_0)$ hold for $\delta > 0$. First we will show $(t, o', p) \in A''_{r_0/4}$ when δ is sufficiently small. Note that $(t_0, o', p_0) \in A''_{r_0/2}$ holds since $p_0 \notin B_{r_0}^{(t_0)}(o)$ and $d_{g(t_0)}(o, o') \in \{r_0/2, 0\}$. Let $q \in \gamma_{o' p_0}^{(t)}$. By the triangle inequality,

$$d_{g(t)}(o, q) \le d_{g(t)}(o, o') + d_{g(t)}(o', p_0). \tag{3.3}$$

Since $r_0/2 < 1 < R$ holds, (2.1) yields $\gamma^{(t_0)}_{oo'} \subset M_0$ when $o' \neq o$. We can easily see that $\gamma^{(t_0)}_{o'p_0} \subset \gamma^{(t_0)}_{op_0} \subset M_0$. Thus, by applying Lemma 2.2 to (3.3),

$$
\begin{aligned}
d_{g(t)}(o,q) &\le \mathrm{e}^{\kappa(t-t_0)}\bigl(d_{g(t_0)}(o,o') + d_{g(t_0)}(o',p_0)\bigr) \\
&\le (R-1)\mathrm{e}^{\kappa\delta}.
\end{aligned}
\tag{3.4}
$$

Take $\delta_2 := 1 \wedge (\kappa^{-1}\log(R/(R-1)))$. Then, for any $\delta \in (0,\delta_2)$, (3.4) and (2.1) imply $\gamma^{(t)}_{o'p_0} \subset M_0$. Hence the triangle inequality, Lemma 2.2 and (3.2) yield

$$
\begin{aligned}
d_{g(t)}(o',p) &\ge d_{g(t)}(o',p_0) - d_{g(t)}(p_0,p) \\
&\ge \mathrm{e}^{-\kappa(t-t_0)} d_{g(t_0)}(o',p_0) - \mathrm{e}^{\kappa(t-t_0)} d_{g(t_0)}(p_0,p) \\
&\ge \frac{\mathrm{e}^{-\kappa\delta} r_0}{2} - \mathrm{e}^{\kappa\delta}\delta,
\end{aligned}
\tag{3.5}
$$

when $\delta \le \delta_2$. Thus there exists $\delta_3 = \delta_3(\kappa, r_0, R) \in (0,\delta_2]$ such that the right-hand side of (3.5) is greater than $r_0/4$ whenever $\delta \in (0,\delta_3)$. Hence $(t,o',p) \in A''_{r_0/4}$ holds in such a case.

Next we will show that there exists $r'_1 > 0$ such that $(t,o',p) \in A'_{r'_1}$ holds for sufficiently small δ. Once we have shown it, the conclusion holds with $r_1 = r'_1 \wedge (r_0/4)$. As we did in showing $(t,o',p) \in A''_{r_0/4}$, we begin with studying the corresponding statement for (t_0,o',p_0). More precisely, we claim that there exists $r''_1 \in (0,1)$ such that $(t_0,o',p_0) \in A_{r''_1}$. When $o' = o$, $(t_0,o',p_0) \in A'_{r_0}$ directly follows from the definition of $o' = o^{(t_0)}_{p_0}$. When $o' \neq o$, set

$$
\begin{aligned}
H := \bigl\{(t,x,y) \in [T_1,T_2] \times M_0 \times M_0 \,|\, &r_0 \le d_{g(t)}(o,y) \le R-1, \\
&d_{g(t)}(o,x) = r_0/2, d_{g(t)}(x,y) = d_{g(t)}(o,y) - d_{g(t)}(o,x)\bigr\}.
\end{aligned}
$$

Note that H is compact and that $H \cap \mathrm{Cut}_{\mathrm{ST}} = \varnothing$ holds since $(t,x,y) \in H$ implies that x is on a minimal $g(t)$-geodesic from y to o. Since $(t_0,o',p_0) \in H$ by the definition of o', it suffices to show that there exists $\tilde{r}_1 > 0$ such that $H \subset A'_{\tilde{r}_1}$. Indeed, the claim will be shown with $r''_1 = \tilde{r}_1 \wedge r_0$ once we have proved it. Suppose that $H \subset A'_r$ does not hold for any $r \in (0,1)$. Then there are sequences $(t_j,x_j,y_j) \in H$, $(t'_j,x'_j,y'_j) \in \mathrm{Cut}_{\mathrm{ST}}$, $j \in \mathbb{N}$, such that $|t_j - t'_j| + d_{g(t_j)}(x_j,x'_j) + d_{g(t_j)}(y_j,y'_j) \to 0$ as $j \to \infty$. We may assume that $((t_j,x_j,y_j))_j$ converges. Since $(t_j,x_j,y_j) \in H$, $x'_j, y'_j \in M_0$ holds for sufficiently large j. Thus we can take a convergent subsequence of $((t'_j,x'_j,y'_j))_j$. Since $\mathrm{Cut}_{\mathrm{ST}}$ and H are closed, and $d_{g(\cdot)}(\cdot,\cdot)$ is continuous, it contradicts with $H \cap \mathrm{Cut}_{\mathrm{ST}} = \varnothing$.

To complete the proof, we show that there exists $\delta_1 \in (0,\delta_3]$ such that $(t,o',p) \in A'_{r''_1/2}$ when $\delta \in (0,\delta_1)$. Suppose that there exists $(t',x',y') \in \mathrm{Cut}_{\mathrm{ST}}$ such that $|t-t'| + d_{g(t)}(o',x') + d_{g(t)}(p,y') < r''_1/2$. For any $q \in \gamma^{(t)}_{py'}$, the triangle inequality and assertion (i) yield

$$
d_{g(t)}(o,q) \le d_{g(t)}(o,p) + d_{g(t)}(p,y') \le \mathrm{e}^{\kappa\delta}(R-1+\delta) + r''_1/2.
\tag{3.6}
$$

A similar observation implies $d_{g(t)}(o,q') \le (\mathrm{e}^{\kappa\delta} r_0 + r_1'')/2$ for $q' \in \gamma^{(t)}_{o'x'}$. Thus there is $\delta_4 = \delta_4(\kappa, R) \in (0, \delta_3]$ such that the right-hand side of (3.6) is less than R and $(\mathrm{e}^{\kappa\delta} r_0 + r_1'')/2 \le R$ whenever $\delta \in (0,\delta_4)$. In such a case, $\gamma^{(t)}_{py'} \subset M_0$ and $\gamma^{(t)}_{o'x'} \subset M_0$ hold. Since $(t_0, o', p_0) \in A'_{r_1''}$, Lemma 2.2 yields

$$\begin{aligned}
&|t-t'| + d_{g(t)}(o',x') + d_{g(t)}(p,y') \\
&\quad \ge |t_0 - t'| - \delta + \mathrm{e}^{-\kappa\delta} d_{g(t_0)}(o',x') + \mathrm{e}^{-\kappa\delta} d_{g(t_0)}(p,y') \\
&\quad \ge \mathrm{e}^{-\kappa\delta} r_1'' + (1-\mathrm{e}^{-\kappa\delta})|t_0 - t'| - \delta - \mathrm{e}^{-\kappa\delta}\delta.
\end{aligned} \tag{3.7}$$

Take $\delta_1 = \delta_1(\kappa, r_1'') \in (0,\delta_4]$ so that the right-hand side of (3.7) is greater than $r_1''/2$ when $\delta \in (0,\delta_1)$. Then (3.7) is absurd for any $\delta \in (0,\delta_1)$. Thus it implies the conclusion. $\square$

We prepare some notation for the second variation formula for the arc length. Let $\nabla^{(t)}$ be the $g(t)$-Levi–Civita connection and $\mathcal{R}^{(t)}$ the $g(t)$-curvature tensor associated with $\nabla^{(t)}$. For a smooth curve γ and smooth vector fields U, V along γ, the index form $I^{(t)}_\gamma(U,V)$ is given by

$$I^{(t)}_\gamma(U,V) := \int_\gamma \bigl(\langle \nabla^{(t)}_{\dot\gamma} U, \nabla^{(t)}_{\dot\gamma} V\rangle_{g(t)} - \langle \mathcal{R}^{(t)}(U,\dot\gamma)\dot\gamma, V\rangle_{g(t)}\bigr)\, ds.$$

We write $I^{(t)}_\gamma(U,U) =: I^{(t)}_\gamma(U)$ for simplicity of notation. Let $G_{t,x,y}(u)$ be the solution to the following initial value problem on $[0, d(x,y)]$:

$$\begin{cases} G''_{t,x,y}(u) = -\dfrac{\mathrm{Ric}_{g(t)}(\dot\gamma^{(t)}_{xy}(u), \dot\gamma^{(t)}_{xy}(u))}{m-1} G_{t,x,y}(u), \\ G_{t,x,y}(0) = 0, \qquad G'_{t,x,y}(0) = 1. \end{cases}$$

Note that $G_{t,x,y}(u) > 0$ for $u \in (0, d(x,y)]$ if $y \notin \mathrm{Cut}_{g(t)}(x)$; see [17], proof of Lemma 9. For simplicity, we write $G_n := G_{t_n^{(\alpha)}, o_n, X^\alpha(t_n^{(\alpha)})}$. When $X^\alpha(t_n^{(\alpha)}) \notin B_{r_0}^{(t_n^{(\alpha)})}(o)$, we define a vector field $V^\dagger$ along γ_n for each $V \in T_{X^\alpha(t_n^{(\alpha)})}M$ by

$$V^\dagger(\gamma_n(u)) := \frac{G_n(u)}{G_n(d_{g(t_n^{(\alpha)})}(o_n, X^\alpha(t_n^{(\alpha)})))} \bigl(/\!\!/^{(t_n^{(\alpha)})}_{\gamma_n} V\bigr)(\gamma_n(u)),$$

where $/\!\!/^{(t_n^{(\alpha)})}_{\gamma_n} V$ is the parallel vector field along γ_n of V associated with $\nabla^{(t_n^{(\alpha)})}$. Take $v \in \mathbb{R}^m$. By using these notations, for $n \in \mathbb{N}_0$ with $n < N^{(\alpha)}$, let us define λ_{n+1} and Λ_{n+1} by

$$\begin{aligned}
\lambda_{n+1} &:= \langle \tilde\xi_{n+1}, \dot\gamma_n\rangle_{g(t_n^{(\alpha)})}, \\
\Lambda_{n+1} &:= \partial_t d_{g(t_n^{(\alpha)})}(o, o_n) + \partial_t d_{g(t_n^{(\alpha)})}\bigl(o_n, X^\alpha\bigl(t_n^{(\alpha)}\bigr)\bigr) \\
&\qquad + \bigl\langle Z\bigl(t_n^{(\alpha)}\bigr), \dot\gamma_n\bigr\rangle_{g(t_n^{(\alpha)})}\bigl(X^\alpha\bigl(t_n^{(\alpha)}\bigr)\bigr) + \tfrac{1}{2} I^{(t_n^{(\alpha)})}_{\gamma_n}(\tilde\xi^\dagger_{n+1}),
\end{aligned}$$

when $X^\alpha(t_n^{(\alpha)}) \notin B_{r_0}^{(t_n^{(\alpha)})}(o)$, and $\lambda_{n+1} := \sqrt{m+2}\langle \xi_{n+1}, v\rangle_{\mathbb{R}^m}$ and $\Lambda_{n+1} := 0$ otherwise.

LEMMA 3.7. *If $n < \bar{\sigma}_R \wedge N^{(\alpha)}$, $\alpha < \alpha_0$ is small enough and $X^\alpha(t_n^{(\alpha)}) \notin B_{r_0}^{(t_n^{(\alpha)})}(o)$, then*

$$d_{g(t_{n+1}^{(\alpha)})}\big(o, X^\alpha\big(t_{n+1}^{(\alpha)}\big)\big) \le d_{g(t_n^{(\alpha)})}\big(o, X^\alpha\big(t_n^{(\alpha)}\big)\big) + \alpha\lambda_{n+1} + \alpha^2\Lambda_{n+1} + o(\alpha^2)$$

almost surely, where α_0 is as in Corollary 3.6. In addition, $o(\alpha^2)$ is controlled uniformly.

PROOF. By virtue of Corollary 3.6, for sufficiently small α, the Taylor expansion together with the second variation formula yields

$$\begin{aligned} &d_{g(t_{n+1}^{(\alpha)})}\big(o_n, X^\alpha\big(t_{n+1}^{(\alpha)}\big)\big) \\ &\quad = d_{g(t_n^{(\alpha)})}\big(o_n, X^\alpha\big(t_n^{(\alpha)}\big)\big) + \alpha\lambda_{n+1} + \alpha^2 \partial_t d_{g(t_n^{(\alpha)})}\big(o_n, X^\alpha\big(t_n^{(\alpha)}\big)\big) \\ &\qquad + \alpha^2 \big\langle Z\big(t_n^{(\alpha)}\big), \dot{\gamma}_n\big\rangle_{g(t_n^{(\alpha)})}\big(X^\alpha\big(t_n^{(\alpha)}\big)\big) + \frac{\alpha^2}{2} I_{\gamma_n}^{(t_n^{(\alpha)})}(J_{\tilde{\xi}_{n+1}}) \\ &\qquad + o(\alpha^2), \end{aligned} \tag{3.8}$$

where $J_{\tilde{\xi}_{n+1}}$ is a $g(t_n^{(\alpha)})$-Jacobi field along γ_n with a boundary value condition $J_{\tilde{\xi}_{n+1}}(o_n) = 0$ and $J_{\tilde{\xi}_{n+1}}(X^\alpha(t_n^{(\alpha)})) = \tilde{\xi}_{n+1}$. Note that $o(\alpha^2)$ can be chosen uniformly since this expansion can be done on the compact set A_{r_1}, and every geodesic variation is included in M_0. By the index lemma, we have $I_{\gamma_n}^{(t_n^{(\alpha)})}(J_{\tilde{\xi}_{n+1}}) \le I_{\gamma_n}^{(t_n^{(\alpha)})}(\tilde{\xi}_{n+1}^\dagger)$. Hence the desired inequality follows when $o_n = o$. In the case $o_n \ne o$, we have

$$\begin{aligned} d_{g(t_{n+1}^{(\alpha)})}\big(o, X^\alpha\big(t_{n+1}^{(\alpha)}\big)\big) &\le d_{g(t_{n+1}^{(\alpha)})}(o, o_n) + d_{g(t_{n+1}^{(\alpha)})}\big(o_n, X^\alpha\big(t_{n+1}^{(\alpha)}\big)\big), \\ d_{g(t_n^{(\alpha)})}\big(o, X^\alpha\big(t_n^{(\alpha)}\big)\big) &= d_{g(t_n^{(\alpha)})}(o, o_n) + d_{g(t_n^{(\alpha)})}\big(o_n, X^\alpha\big(t_n^{(\alpha)}\big)\big). \end{aligned}$$

Note that $(t_n^{(\alpha)}, o, o_n)$ is uniformly away from $\mathrm{Cut}_{\mathrm{ST}}$ because of our choice of r_0 and Lemma 2.6. Therefore the conclusion follows by combining them with (3.8). $\square$

Before turning into the next step, we show the following two complementary lemmas (Lemmas 3.8 and 3.9) which provide a nice control of the second-order term Λ_n in Lemma 3.7. Set $\bar{\Lambda}_n = \mathbb{E}[\Lambda_n | \mathscr{F}_{n-1}]$.

LEMMA 3.8. *Let $(a_n)_{n\in\mathbb{N}_0}$ be a uniformly bounded $\mathscr{F}_n$-predictable process. Then*

$$\lim_{\alpha\to 0}\alpha^2\sup\left\{\left|\sum_{j=n}^{N+1}a_j(\Lambda_j-\bar{\Lambda}_j)\right|\middle| n,N\in\mathbb{N},n\le N\le N^{(\alpha)}\wedge\bar{\sigma}_R\right\}=0$$

in probability.

PROOF. Note that the map $(t,x,y)\mapsto G_{t,x,y}(d(x,y))$ is continuous on A_{r_1}. Since we have $G_{t,x,y}(d(x,y))>0$ on A_{r_1}, there exists $K>0$ such that $K^{-1}<G_{t,x,y}(d(x,y))<K$. This fact together with Corollary 3.6 yields $|\Lambda_j|$ and $|\bar{\Lambda}_j|$ are uniformly bounded if $j<\bar{\sigma}_R$. Since $\sum_{j=1}^n a_j(\Lambda_j-\bar{\Lambda}_j)$ is an $\mathscr{F}_n$-local martingale and $\bar{\sigma}_R$ is $\mathscr{F}_n$-stopping time, the Doob inequality yields

$$\lim_{\alpha\to 0}\alpha^2\sup_{0\le N\le N^{(\alpha)}\wedge\bar{\sigma}_R}\left|\sum_{j=1}^{N+1}a_j(\Lambda_j-\bar{\Lambda}_j)\right|=0\qquad\text{in probability.}\tag{3.9}$$

Here we used the fact $\lim_{\alpha\to 0}\alpha^2N^{(\alpha)}=T_2-T_1$. Note that

$$\begin{aligned}&\bigcup_{N=1}^{N^{(\alpha)}\wedge\bar{\sigma}_R}\bigcup_{n=0}^{N}\left\{\alpha^2\left|\sum_{j=n+1}^{N+1}a_j(\Lambda_j-\bar{\Lambda}_j)\right|>\delta\right\}\\&\quad\subset\bigcup_{N=1}^{N^{(\alpha)}\wedge\bar{\sigma}_R}\bigcup_{n=1}^{N}\left\{\alpha^2\left|\sum_{j=1}^{n}a_j(\Lambda_j-\bar{\Lambda}_j)\right|>\frac{\delta}{2}\right\}\cup\left\{\alpha^2\left|\sum_{j=1}^{N+1}a_j(\Lambda_j-\bar{\Lambda}_j)\right|>\frac{\delta}{2}\right\}\\&\quad=\left\{\alpha^2\sup_{0\le N\le N^{(\alpha)}\wedge\bar{\sigma}_R}\left|\sum_{j=1}^{N+1}a_j(\Lambda_j-\bar{\Lambda}_j)\right|>\frac{\delta}{2}\right\}.\end{aligned}$$

Thus the conclusion follows from (3.9). □

LEMMA 3.9. *There exists a deterministic constant $C_0>0$ being independent of α and R such that the following holds*:

$$\bar{\Lambda}_{n+1}\le C_0+\frac{1}{2}\int_0^{d_{g(t_n^{(\alpha)})}(o,X^\alpha(t_n^{(\alpha)}))}b(u)\,du,$$

where b is what appeared in Assumption 1.

PROOF. By using $(m+2)\mathbb{E}[\langle\xi_n,e_i\rangle\langle\xi_n,e_j\rangle]=\delta_{ij}$, we obtain

$$\begin{aligned}\mathbb{E}[I_{\gamma_n}^{(t_n^{(\alpha)})}(\tilde{\xi}_{n+1}^\dagger)]&=\sum_{j=2}^m I_{\gamma_n}^{(t_n^{(\alpha)})}((\Phi^{(t_n^{(\alpha)})}(X^\alpha(t_n^{(\alpha)}))e_j)^\dagger)\\&=\frac{(m-1)G_n'(d(o_n,X^\alpha(t_n^{(\alpha)})))}{G_n(d(o_n,X^\alpha(t_n^{(\alpha)})))}.\end{aligned}$$

Note that we have

$$
\begin{aligned}
&\langle Z(t_n^{(\alpha)}), \dot\gamma_n\rangle_{g(t_n^{(\alpha)})}(X^\alpha(t_n^{(\alpha)})) - \langle Z(t_n^{(\alpha)}), \dot\gamma_n\rangle_{g(t_n^{(\alpha)})}(o_n) \\
&\quad = \int_0^{d_{g(t_n^{(\alpha)})}(o_n, X^\alpha(t_n^{(\alpha)}))} \partial_s \langle Z(t_n^{(\alpha)}), \dot\gamma_n\rangle_{g(t_n^{(\alpha)})}(\gamma_n(s))\Big|_{s=u}\, du \\
&\quad = \int_0^{d_{g(t_n^{(\alpha)})}(o_n, X^\alpha(t_n^{(\alpha)}))} \langle \nabla^{(t_n^{(\alpha)})}_{\dot\gamma_n} Z(t_n^{(\alpha)}), \dot\gamma_n\rangle_{g(t_n^{(\alpha)})}(\gamma_n(u))\, du.
\end{aligned}
$$

Recall that, for $(t, x, y) \notin \mathrm{Cut}_{\mathrm{ST}}$, we have

$$
\partial_t d_{g(t)}(x, y) = \frac{1}{2}\int_0^{d_{g(t)}(x,y)} (\partial_t g(t))(\dot\gamma_{xy}^{(t)}(u), \dot\gamma_{xy}^{(t)}(u))\, du;
$$

cf. [19], Remark 6. By combining them with Assumption 1,

$$
\begin{aligned}
\bar\Lambda_{n+1} &= \partial_t d_{g(t_n^{(\alpha)})}(o, o_n) \\
&\quad + \frac{1}{2}\int_0^{d_{g(t_n^{(\alpha)})}(o_n, X^\alpha(t_n^{(\alpha)}))} \partial_t g(t_n^{(\alpha)})(\dot\gamma_n(u), \dot\gamma_n(u))\, du \\
&\quad + \langle Z(t_n^{(\alpha)}), \dot\gamma_n\rangle_{g(t_n^{(\alpha)})}(X^\alpha(t_n^{(\alpha)})) + \frac{(m-1)G_n'(d(o_n, X^\alpha(t_n^{(\alpha)})))}{2G_n(d(o_n, X^\alpha(t_n^{(\alpha)})))} \\
&\le \frac{1}{2}\int_0^{d_{g(t_n^{(\alpha)})}(o, X^\alpha(t_n^{(\alpha)}))} b(u)\, du \\
&\quad + \partial_t d_{g(t_n^{(\alpha)})}(o, o_n) + \langle Z(t_n^{(\alpha)}), \dot\gamma_n\rangle_{g(t_n^{(\alpha)})}(o_n) \\
&\quad + \frac{1}{2}\int_0^{d_{g(t_n^{(\alpha)})}(o_n, X^\alpha(t_n^{(\alpha)}))} \mathrm{Ric}_{g(t_n^{(\alpha)})}(\dot\gamma_n(u), \dot\gamma_n(u))\, du \\
&\quad + \frac{(m-1)G_n'(d(o_n, X^\alpha(t_n^{(\alpha)})))}{2G_n(d(o_n, X^\alpha(t_n^{(\alpha)})))}.
\end{aligned}
\tag{3.10}
$$

Here we used the fact $b(u) \ge 0$ in the case $o_n \ne o$. Note that

$$
\int_0^r \mathrm{Ric}_{g(t_n^{(\alpha)})}(\dot\gamma_n(u), \dot\gamma_n(u))\, du + \frac{(m-1)G_n'(r)}{G_n(r)}
$$

is nonincreasing as a function of r. Indeed, we can easily verify it by taking a differentiation. Set

$$
\begin{aligned}
C_1 := \sup_{t\in[T_1,T_2]} \sup_{x \in B_{r_0}^{(t)}(o)} \Bigl(&|Z(t)|_{g(t)}(x) \\
&+ \sup_{\substack{V\in T_xM \\ |V|_{g(t)}\le 1}} \bigl(\partial_t g(t)(V, V) + |\mathrm{Ric}_{g(t)}(V, V)|\bigr)\Bigr).
\end{aligned}
$$

By virtue of Lemma 2.2, $C_1 < \infty$ holds. By applying a usual comparison argument to $G'_n(r_0)/G_n(r_0)$, we obtain

$$\begin{aligned}\int_0^{d_{g(t_n^{(\alpha)})}(o_n, X^\alpha(t_n^{(\alpha)}))} &\mathrm{Ric}_{g(t_n^{(\alpha)})}(\dot\gamma_n(u), \dot\gamma_n(u))\,du \\ &+ \frac{(m-1)G'_n(d(o_n, X^\alpha(t_n^{(\alpha)})))}{G_n(d(o_n, X^\alpha(t_n^{(\alpha)})))} \\ &\le C_1\bigl(r_0 + \coth(C_1 r_0)\bigr).\end{aligned}$$

Hence the conclusion with $C_0 = C_1(1 + 3r_0/4 + \coth(C_1 r_0)/2)$ follows from (3.10). □

In the next step, we will introduce a comparison process to give a control of the radial process. Let us define a function φ on $(2r_0, \infty)$ by

$$\varphi(r) := C_0 + \frac{1}{2}\int_0^r b(u)\,du,$$

where C_0 is as in Lemma 3.9. Let us define another function ψ on $(2r_0, \infty)$ so that ψ is a locally Lipschitz nonincreasing function satisfying $\psi(r) := 2(r - 2r_0)^{-1}$ for $r \in (2r_0, 2r_0 + 1]$ and $\psi(r) := 0$ for $r \ge 2r_0 + 2$. Let us define a comparison process $\rho^\alpha(t)$ taking values in $[0, \infty)$ inductively by

$$\begin{aligned}\rho^\alpha(T_1) &:= d_{g(T_1)}(o, x_0) + 3r_0, \\ \rho^\alpha(t) &:= \rho^\alpha\bigl(t_n^{(\alpha)}\bigr) + \frac{t - t_n^{(\alpha)}}{\alpha^2}\bigl(\alpha\lambda_{n+1} + \alpha^2\bigl(\varphi\bigl(\rho^\alpha\bigl(t_n^{(\alpha)}\bigr)\bigr) + \psi\bigl(\rho^\alpha\bigl(t_n^{(\alpha)}\bigr)\bigr)\bigr)\bigr), \\ &\qquad\qquad t \in \bigl[t_n^{(\alpha)}, t_{n+1}^{(\alpha)}\bigr].\end{aligned}$$

The term $\psi(\rho^\alpha(t_n^{(\alpha)}))$ is inserted to avoid a difficulty coming from the absence of the estimate in Lemma 3.7 on a neighborhood of o. By virtue of this extra term, $\rho^\alpha(t) > 2r_0$ holds for all $t \in [T_1, T_2]$ if α is sufficiently small. Let $\hat\sigma'_R$ and $\bar\sigma'_R$ be given by $\hat\sigma'_R := \sigma_R(\rho^\alpha)$ and $\bar\sigma'_R := \alpha^{-2}(\lfloor\hat\sigma'_R\rfloor_\alpha - T_1) + 1$. The following is a modification of an argument in the proof of [12], Theorem 3.5.3, into our discrete setting.

LEMMA 3.10. *For $\delta > 0$, there exist a family of events $(E_\delta^\alpha)_\alpha$ with $\lim_{\alpha\to 0}\mathbb{P}[E_\delta^\alpha] = 1$ and a constant $K(\delta) > 0$ with $\lim_{\delta\to 0} K(\delta) = 0$ such that, on E_δ^α,*

$$d_{g(t)}(o, X^\alpha(t)) \le \rho^\alpha(t) + K(\delta)$$

for $t \in [T_1, \hat\sigma_R \wedge \hat\sigma'_R \wedge T_2]$ and sufficiently small α relative to δ and R^{-1}.

PROOF. It suffices to show the assertion in the case $t = t_n^{(\alpha)}$ for some $n \in \mathbb{N}_0$. Indeed, once we have shown it, Corollary 3.6(i) yields

$$\begin{aligned} d_{g(t)}(o, X^\alpha(t)) &\le e^{\kappa\alpha^2}\big(d_{g(\lfloor t\rfloor_\alpha)}(o, X^\alpha(\lfloor t\rfloor_\alpha)) + h(\alpha)\big) \\ &\le \rho^\alpha_{\lfloor t\rfloor_\alpha} + K(\delta) + (e^{\kappa\alpha^2} - 1)R + e^{\kappa\alpha^2}h(\alpha) \\ &\le \rho^\alpha_t + K(\delta) + \alpha + (e^{\kappa\alpha^2} - 1)R + e^{\kappa\alpha^2}h(\alpha) \end{aligned}$$

for $t \in [T_1, \hat{\sigma}_R \wedge T_2]$. Here we used the facts $\varphi \ge 0$ and $\psi \ge 0$. From this estimate, we can easily deduce the conclusion.

For simplicity of notation, we denote $d_{g(t_n^{(\alpha)})}(o, X^\alpha(t_n^{(\alpha)}))$ and $\rho^\alpha(t_n^{(\alpha)})$ by d_n and ρ_n, respectively, in the rest of this proof. Let us define a sequence of $\mathscr{F}_n$-stopping times S_l by $S_0 := 0$ and

$$S_{2l+1} := \inf\big\{j \ge S_{2l} \,|\, X^\alpha\big(t_j^{(\alpha)}\big) \in B_{r_0}^{(t_j^{(\alpha)})}(o)\big\} \wedge N^{(\alpha)},$$

$$S_{2l} := \inf\big\{j \ge S_{2l-1} \,|\, X^\alpha\big(t_j^{(\alpha)}\big) \notin B_{3r_0/2}^{(t_j^{(\alpha)})}(o)\big\} \wedge N^{(\alpha)}.$$

Since $\rho_n > 2r_0$, it suffices to show the assertion in the case $S_{2l} \le n < S_{2l+1} \wedge \bar{\sigma}_R \wedge \bar{\sigma}'_R$ for some $l \in \mathbb{N}_0$. Now Lemmas 3.7 and 3.9 imply

$$d_{j+1} - \rho_{j+1} \le d_j - \rho_j + \alpha^2\big(\varphi(d_j) - \varphi(\rho_j)\big) + \alpha^2(\Lambda_{j+1} - \bar{\Lambda}_{j+1}) + o(\alpha^2)$$

for $j \in [S_{2l}, S_{2l+1} \wedge \sigma'_R \wedge \bar{\sigma}'_R)$. Here we used the fact $\psi \ge 0$. Let f_α be a C^2-function on $\mathbb{R}$ satisfying:

(i) $f_\alpha|_{(-\infty,-\alpha)} \equiv 0$; $f_\alpha|_{(\alpha,\infty)}(x) = x$;
(ii) f_α is convex;
(iii) $\alpha^2 \sup_{x\in\mathbb{R}} f''_\alpha(x) = o(1)$.

For example, a function f_α satisfying these conditions is constructed by setting

$$\tilde{f}(x) = \int_{-\infty}^x \int_{-\infty}^t b \exp\Big(-\frac{a}{1-s^2}\Big) 1_{(-1,1)}(s)\, ds\, dt,$$

where a, b is chosen to satisfy

$$\int_{-\infty}^1 \exp\Big(-\frac{a}{1-s^2}\Big) 1_{(-1,1)}(s)\, ds = 1,$$

$$b \int_{-\infty}^1 \int_{-\infty}^t \exp\Big(-\frac{a}{1-s^2}\Big) 1_{(-1,1)}(s)\, ds\, dt = 1$$

and $f_\alpha(x) := \alpha \tilde{f}(\alpha^{-1}x)$. By the Taylor expansion with condition (iii) of f_α, we have

$$\begin{aligned} f_\alpha(d_{j+1} - \rho_{j+1}) &\le f_\alpha(d_j - \rho_j) \\ &\quad + \alpha^2 f'_\alpha(d_j - \rho_j)\big(\varphi(d_j) - \varphi(\rho_j) + (\Lambda_j - \bar{\Lambda}_j)\big) \qquad (3.11) \\ &\quad + o(\alpha^2). \end{aligned}$$

Let $C > 0$ be the Lipschitz constant of φ on $[0, R]$. Note that we have

$$f'_\alpha(d_j - \rho_j)\big(\varphi(d_j) - \varphi(\rho_j)\big) \le C(d_j - \rho_j)_+ \tag{3.12}$$

since φ is nondecreasing. Now by using (3.11) and (3.12) combined with the fact $d_{S_{2l}} - \rho_{S_{2l}} < -\alpha$ for sufficiently small α, we obtain

$$\begin{aligned}(d_n - \rho_n)_+ &\le f_\alpha(d_n - \rho_n) \\ &\le C\alpha^2 \sum_{j=S_{2k}}^{n-1} (d_j - \rho_j)_+ + \alpha^2 \sum_{j=S_{2k}}^{n-1} f'_\alpha(d_j - \rho_j)(\Lambda_{j+1} - \bar{\Lambda}_{j+1}) \\ &\quad + o(1).\end{aligned} \tag{3.13}$$

Here the first inequality follows from condition (ii) of f_α, and $n \le \alpha^{-2}(T_2 - T_1)$ is used to derive the error term $o(1)$. Let E_δ^α be an event defined by

$$E_\delta^\alpha := \left\{ \alpha^2 \sup_{k \le k' \le N^{(\alpha)} \wedge \bar{\sigma}_R} \left| \sum_{j=k}^{k'} f'_\alpha(d_{j-1} - \rho_{j-1})(\Lambda_j - \bar{\Lambda}_j) \right| < \delta \right\}.$$

Note that $a_j = f'_\alpha(d_{j-1} - \rho_{j-1})$ is $\mathscr{F}_n$-predictable and uniformly bounded by 1. Thus, by combining Lemma 3.8 with (3.13), we obtain

$$(d_n - \rho_n)_+ \le C\alpha^2 \sum_{j=S_{2l}}^{n-1} (d_j - \rho_j)_+ + 2\delta$$

on E_δ^α for sufficiently small α. Thus, by virtue of a discrete Gronwall inequality (see [31], e.g.),

$$(d_n - \rho_n)_+ \le 2\delta\big(1 + (1 + C\alpha^2)^n\big) \le 2\delta\big(1 + \mathrm{e}^{C(T_2 - T_1)}\big).$$

This estimate implies the conclusion. $\square$

COROLLARY 3.11. *For every $R' < R$,*

$$\limsup_{\alpha \to 0} \mathbb{P}[\hat{\sigma}_R \le T_2] \le \limsup_{\alpha \to 0} \mathbb{P}[\hat{\sigma}'_{R'} \le T_2].$$

Now we turn to the proof of our destination in this section.

PROOF OF PROPOSITION 3.4. By Corollary 3.11, the proof of Proposition 3.4 is reduced to estimate $\mathbb{P}[\hat{\sigma}'_R \le T_2]$. To obtain a useful bound of it, we would like to apply the invariance principle for ρ^α. However, there is a technical difficulty coming from the unboundedness of the drift term of ρ^α. To avoid it, we introduce an auxiliary process $\tilde{\rho}^\alpha$ in the sequel.

Let $\tilde{\varphi}$ be a bounded, globally Lipschitz function on $\mathbb{R}$ such that $\tilde{\varphi}(r) = \varphi(r) + \psi(r)$ for $r \in [2r_0 + R^{-1}, R]$. Let us define an $\mathbb{R}$-valued process $\tilde{\rho}^\alpha(t)$ inductively by

$$\tilde{\rho}^\alpha(T_1) := d_{g(T_1)}(o, x_0) + 3r_0,$$
$$\tilde{\rho}^\alpha(t) := \tilde{\rho}^\alpha\bigl(t_n^{(\alpha)}\bigr) + \frac{t - t_n^{(\alpha)}}{\alpha^2}\bigl(\alpha\lambda_{n+1} + \alpha^2\tilde{\varphi}\bigl(\tilde{\rho}^\alpha\bigl(t_n^{(\alpha)}\bigr)\bigr)\bigr), \qquad t \in \bigl[t_n^{(\alpha)}, t_{n+1}^{(\alpha)}\bigr].$$

We also define two diffusion processes $\rho^0(t)$ and $\tilde{\rho}^0(r)$ as solutions to the following SDEs:

$$\begin{cases} d\rho^0(t) = dB(t) + \bigl(\varphi(\rho^0(t)) + \psi(\rho^0(t))\bigr)\,dt, \\ \rho^0(T_1) = d_{g(T_1)}(o, x_0) + 3r_0, \end{cases}$$
$$\begin{cases} d\tilde{\rho}^0(t) = dB(t) + \tilde{\varphi}(\tilde{\rho}^0(t))\,dt, \\ \tilde{\rho}^0(T_1) = d_{g(T_1)}(o, x_0) + 3r_0, \end{cases}$$

where $(B(t))_{t\in[T_1,T_2]}$ is a standard one-dimensional Brownian motion with $B(T_1) = 0$. We claim that $\tilde{\rho}^\alpha$ converges in law to $\tilde{\rho}^0$ as $\alpha \to 0$. Indeed, we can easily show the tightness of $(\tilde{\rho}^\alpha)_{\alpha>0}$ by modifying an argument for the invariance principle for i.i.d. sequences since $\tilde{\varphi}$ is bounded. Then the claim follows from the same argument as we used in the proof of Theorem 3.1 under Proposition 3.2, which is based on the Poisson subordination and the uniqueness of the martingale problem.

Let us define $\eta_R : \mathscr{C}_1 \to [T_1, T_2] \cup \{\infty\}$ by

$$\eta_R(w) := \inf\{t \in [T_1, T_2] | w(t) \le 2r_0 + R^{-1}\}.$$

Then we have

$$\mathbb{P}[\hat{\sigma}'_R \le T_2] \le \mathbb{P}[\sigma_R(\rho^\alpha) \wedge \eta_R(\rho^\alpha) \le T_2] = \mathbb{P}[\sigma_R(\tilde{\rho}^\alpha) \wedge \eta_R(\tilde{\rho}^\alpha) \le T_2].$$

Since $\{w | \sigma_R(w) \wedge \eta_R(w) \le T_2\}$ is closed in $\mathscr{C}_1$, the Portmanteau theorem implies

$$\begin{aligned}\limsup_{\alpha\to 0} \mathbb{P}[\sigma_R(\tilde{\rho}^\alpha) \wedge \eta_R(\tilde{\rho}^\alpha) \le T_2] &\le \mathbb{P}[\sigma_R(\tilde{\rho}^0) \wedge \eta_R(\tilde{\rho}^0) \le T_2] \\ &= \mathbb{P}[\sigma_R(\rho^0) \wedge \eta_R(\rho^0) \le T_2].\end{aligned}$$

Since ρ^0 is a diffusion process on $(2r_0, \infty)$ which cannot reach the boundary by Assumption 1, the conclusion follows. □

3.2. *Tightness of geodesic random walks.* Recall that we have metrized the path space $\mathscr{C}$ by using $d_{g(T_1)}$. To deal with the tightness of $(X^\alpha)_{\alpha\in(0,1)}$ in $\mathscr{C}$, we show the following lemma, which provides a tightness criterion compatible with the time-dependent metric $d_{g(t)}$.

LEMMA 3.12. *$(X^\alpha)_{\alpha\in(0,1)}$ is tight if*

$$\lim_{\delta\to 0}\frac{1}{\delta}\limsup_{\alpha\to 0}\sup_{n\in\mathbb{N}_0}\mathbb{P}\Big[\sup_{t_n^{(\alpha)}\le s\le(t_n^{(\alpha)}+\delta)\wedge T_2} d_{g(s)}\big(X^\alpha\big(t_n^{(\alpha)}\big), X^\alpha(s)\big)>\varepsilon,\ \hat{\sigma}_R=\infty\Big]=0 \tag{3.14}$$

holds for every $\varepsilon>0$ and $R>1$.

PROOF. By following a standard argument (e.g., [5], Theorems 7.3 and 7.4), we can easily show that $(X^\alpha)_{\alpha\in(0,1)}$ is tight if, for every $\varepsilon>0$,

$$\lim_{\delta\to 0}\frac{1}{\delta}\limsup_{\alpha\to 0}\sup_{t\in[T_1,T_2]}\mathbb{P}\Big[\sup_{t\le s\le(t+\delta)\wedge T_2} d_{g(T_1)}(X^\alpha(t), X^\alpha(s))>\varepsilon\Big]=0.$$

Thus, by virtue of Proposition 3.4, $(X^\alpha)_{\alpha\in(0,1)}$ is tight if

$$\lim_{\delta\to 0}\frac{1}{\delta}\limsup_{\alpha\to 0}\sup_{t\in[T_1,T_2]}\mathbb{P}\Big[\sup_{t\le s\le(t+\delta)\wedge T_2} d_{g(T_1)}(X^\alpha(t), X^\alpha(s))>\varepsilon, \hat{\sigma}_R=\infty\Big]=0$$

for every $\varepsilon>0$ and $R>1$. Given $R>1$, take M_0 and κ as in Lemmas 2.4 and 2.2, respectively. Then, for $\varepsilon<1$ and $s,t\in[T_1,T_2]$,

$$\begin{aligned}&\{d_{g(s)}(X^\alpha(s), X^\alpha(\lfloor t\rfloor_\alpha))\le\varepsilon, \hat{\sigma}_R=\infty\}\\&\qquad\subset\{d_{g(T_1)}(X^\alpha(s), X^\alpha(t))\le 2\mathrm{e}^{\kappa(T_2-T_1)}\varepsilon, \hat{\sigma}_R=\infty\},\end{aligned}$$

if α is sufficiently small. Thus we have

$$\begin{aligned}&\Big\{\sup_{t\le s\le(t+\delta)\wedge T_2} d_{g(T_1)}(X^\alpha(t), X^\alpha(s))>\varepsilon, \hat{\sigma}_R=\infty\Big\}\\&\qquad\subset\Big\{\sup_{\lfloor t\rfloor_\alpha\le s\le(\lfloor t\rfloor_\alpha+2\delta)\wedge T_2} d_{g(s)}(X^\alpha(\lfloor t\rfloor_\alpha), X^\alpha(s))>\frac{\mathrm{e}^{-\kappa(T_2-T_1)}\varepsilon}{2}, \hat{\sigma}_R=\infty\Big\}\end{aligned}$$

for $\alpha^2\le\delta$, and hence the conclusion follows. □

PROOF OF PROPOSITION 3.2. Take $R>1$. By virtue of Lemma 3.12, it suffices to show (3.14). Take $M_0\subset M$ compact and κ as in Lemmas 2.4 and 2.2, respectively. By taking smaller $\varepsilon>0$, we may assume that $\varepsilon<\tilde{r}_0/2$, where $\tilde{r}_0=\tilde{r}_0(M_0)$ is as in Lemma 2.6. Take $n\in\mathbb{N}_0$ with $n<N^{(\alpha)}$. Let us define a $\mathscr{F}_k$-stopping time ζ_ε by

$$\zeta_\varepsilon:=\inf\big\{k\in\mathbb{N}_0 \mid n\le k\le N^{(\alpha)}, d_{g(t_k^{(\alpha)})}\big(X^\alpha\big(t_n^{(\alpha)}\big), X^\alpha\big(t_k^{(\alpha)}\big)\big)>\varepsilon\big\}.$$

Then, for sufficiently small α,

$$\begin{aligned}&\Big\{\sup_{t_n^{(\alpha)}\le s\le(t_n^{(\alpha)}+\delta)\wedge T_2} d_{g(s)}\big(X^\alpha\big(t_n^{(\alpha)}\big), X^\alpha(s)\big)\ge 2\varepsilon, \hat{\sigma}_R=\infty\Big\}\\&\qquad\subset\{\alpha^2(\zeta_\varepsilon-n)<\delta, \hat{\sigma}_R=\infty\}.\end{aligned} \tag{3.15}$$

Set $p_k := X^\alpha(t_k^{(\alpha)})$ for $k \in \mathbb{N}_0$ and $f(t,x) := d_{g(t)}(p_n, x)$. Note that f^2 is smooth on $\{f < \varepsilon\}$. Let us define λ_k' by

$$\lambda_{k+1}' := \langle \tilde{\xi}_{k+1}, \dot{\gamma}_{p_n p_k}^{(t_k^{(\alpha)})} \rangle_{g(t_k^{(\alpha)})}.$$

We claim that there exists a constant $C > 0$ such that

$$f(t_{k+1}^{(\alpha)}, p_{k+1})^2 \le f(t_k^{(\alpha)}, p_k)^2 + 2\alpha f(t_k^{(\alpha)}, p_k)\lambda_{k+1}' + C\alpha^2 \tag{3.16}$$

for $k \le \zeta_\varepsilon \wedge N^{(\alpha)}$ on $\{\hat{\sigma}_R = \infty\}$. Indeed, in the same way as we did to obtain (3.8),

$$\begin{aligned} f(t_{k+1}^{(\alpha)}, p_{k+1})^2 &\le f(t_k^{(\alpha)}, p_k)^2 + 2\alpha f(t_k^{(\alpha)}, p_k)\lambda_{k+1}' + \alpha^2(\lambda_{k+1}')^2 \\ &\quad + 2\alpha^2 f(t_k^{(\alpha)}, p_k)\big(\partial_t f(t_k^{(\alpha)}, p_k) + \langle Z(t_k^{(\alpha)}), \dot{\gamma}_{p_n p_k}^{(t_n^{(\alpha)})} \rangle_{g(t_k^{(\alpha)})}(p_k)\big) \\ &\quad + \alpha^2 f(t_k^{(\alpha)}, p_k) I_{\gamma_{p_n p_k}^{(t_k^{(\alpha)})}}^{(t_k^{(\alpha)})}(J_{\tilde{\xi}_{k+1}}) + o(\alpha^2). \end{aligned} \tag{3.17}$$

Here $o(\alpha^2)$ is controlled uniformly. Let $K_1 > 0$ be a constant satisfying that the $g(t)$-sectional curvature on M_0 is bounded below by $-K_1$ for every $t \in [T_1, T_2]$. Such a constant exists since M_0 is compact. Then a comparison argument implies

$$f(t_k^{(\alpha)}, p_k) I_{\gamma_{p_n p_k}^{(t_k^{(\alpha)})}}^{(t_k^{(\alpha)})}(J_{\tilde{\xi}_{k+1}}) \le K_1 f(t_k^{(\alpha)}, p_k) \coth\big(K_1 f(t_k^{(\alpha)}, p_k)\big).$$

Here the right-hand side is bounded uniformly if $k < \zeta_\varepsilon \wedge N^{(\alpha)}$. The remaining estimate of the second-order term in (3.17) to show (3.16) is easy since we are on the event $\{\hat{\sigma}_R = \infty\}$. Applying (3.16) repeatedly from $k = n$ to $k = \zeta_\varepsilon$, we obtain

$$\varepsilon^2 < 2\alpha \sum_{k=n}^{\zeta_\varepsilon} f(t_k^{(\alpha)}, p_k)\lambda_{k+1}' + C\delta$$

on $\{\alpha^2(\zeta_\varepsilon - n) < \delta, \hat{\sigma}_R = \infty\}$. Set $N_\delta^{(\alpha)} := \sup\{k \in \mathbb{N}_0 | k \le \alpha^{-2}\delta + n\}$. By taking $\delta < (2C)^{-1}\varepsilon^2$, we obtain

$$\begin{aligned} &\{\alpha^2(\zeta_\varepsilon - n) < \delta, \hat{\sigma}_R = \infty\} \\ &\quad \subset \left\{ \sum_{k=n}^{\zeta_\varepsilon} f(t_k^{(\alpha)}, p_k)\lambda_{k+1}' > \frac{\varepsilon^2}{4\alpha}, \alpha^2(\zeta_\varepsilon - n) < \delta, \hat{\sigma}_R = \infty \right\} \\ &\quad \subset \left\{ \sup_{n \le N \le N_\delta^{(\alpha)}} \sum_{k=n}^{N} f(t_k^{(\alpha)}, p_k) 1_{\{f(t_k^{(\alpha)}, p_k) \le \varepsilon\}} \lambda_{k+1}' > \frac{\varepsilon^2}{4\alpha} \right\}. \end{aligned} \tag{3.18}$$

Set

$$Y_{k+1} := \frac{1}{\sqrt{m+2}} f\big(t_k^{(\alpha)}, p_k\big) 1_{\{f(t_k^{(\alpha)}, p_k) \le \varepsilon\}} \lambda'_{k+1}.$$

We can easily see that $|Y_k| \le 1$ and $\sum_{k=n+1}^{N} Y_k$ is $\mathscr{F}_N$-martingale. By [11], Theorem 1.6, with (3.18), we obtain

$$\begin{aligned}
&\mathbb{P}[\alpha^2(\zeta_\varepsilon - n) < \delta, \hat{\sigma}_R = \infty] \\
&\quad \le \mathbb{P}\Bigg[\sup_{n \le N \le N_\delta^{(\alpha)}} \sum_{k=n+1}^{N+1} Y_k > \frac{\varepsilon^2}{4\alpha\sqrt{m+2}}\Bigg] \\
&\quad \le \exp\Bigg(-\frac{\varepsilon^4}{8\sqrt{m+2}(\alpha\varepsilon^2 + 4\alpha^2\sqrt{m+2}(N_\delta^{(\alpha)} - n))}\Bigg) \\
&\quad \le \exp\Bigg(-\frac{\varepsilon^4}{8\sqrt{m+2}(\alpha\varepsilon^2 + 4\sqrt{m+2}\delta)}\Bigg).
\end{aligned}$$

Hence (3.14) follows by combining this estimate with (3.15). $\square$

4. Coupling by reflection. For $k \in \mathbb{R}$, let $U_{a,k}$ be a one-dimensional Ornstein–Uhlenbeck process defined as a solution to the following SDE:

$$\begin{aligned}
dU_{a,k}(t) &= -\frac{k}{2} U_{a,k}(t)\,dt + 2\,dB(t), \\
U_{a,k}(T_1) &= a.
\end{aligned}$$

More explicitly, $U_{a,k}(t) = \mathrm{e}^{-k(t-T_1)/2} a + 2\int_{T_1}^{t} \mathrm{e}^{k(s-t)/2}\,dB(s)$. Here $B(t)$ is the standard one-dimensional Brownian motion as in the proof of Proposition 3.4.

THEOREM 4.1. *Suppose*

$$2(\nabla Z(t))^\flat + \partial_t g(t) \le \mathrm{Ric}_{g(t)} + k g(t) \tag{4.1}$$

holds for some $k \in \mathbb{R}$. Then, for each $x_1, x_2 \in M$, there exists a coupling $\mathbf{X}(t) := (X_1(t), X_2(t))$ of two $\mathscr{L}_t$-diffusion processes starting at (x_1, x_2) satisfying

$$\begin{aligned}
\mathbb{P}\Big[\inf_{T_1 \le t \le T} d_{g(t)}(\mathbf{X}(t)) > 0\Big] &\le \mathbb{P}\Big[\inf_{T_1 \le t \le T} U_{d_{g(T_1)}(x_1,x_2),k}(t) > 0\Big] \\
&= \chi\left(\frac{d_{g(T_1)}(x_1, x_2)}{2\sqrt{\beta(T - T_1)}}\right)
\end{aligned}$$

for each $T \in [T_1, T_2]$, where

$$\chi(a) := \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} \mathrm{e}^{-u^2/2}\,du, \qquad \beta(t) := \begin{cases} \dfrac{e^{kt} - 1}{k}, & k \ne 0, \\ t, & k = 0. \end{cases}$$

In addition, for $i = 1, 2$, $X_i(t)$ is a solution to the martingale problem associated with the time-inhomogeneous generator $\mathscr{L}_t$ and the filtration generated by $\mathbf{X}$.

REMARK 4.2. (i) Our assumption (4.1) extends existing curvature assumptions in two respects. On the one hand, (4.1) is nothing but (1.1) when $Z(t) \equiv 0$ and $k = 0$. On the other hand, (4.1) can be regarded as a natural extension of a lower Ricci curvature bound by k. Indeed, Bakry–Émery's curvature-dimension condition $\mathsf{CD}(k, \infty)$ (see [2], e.g.), which is a natural extension of a lower Ricci curvature bound by k, appears in (4.1) when both $Z(t)$ and $g(t)$ are independent of t.

(ii) Given $k > 0$, a simple example satisfying (4.1) can be constructed by a scaling. Indeed, for a complete metric g whose Ricci curvature is nonnegative, $g(t) = \mathrm{e}^{-k(t-T_1)} g$ satisfies (4.1) when $Z(t) \equiv 0$.

(iii) From the first item in this remark, when $Z(t) \equiv 0$, one may expect that (4.1) works as an analog of Bakry–Émery's $\mathsf{CD}(k, N)$ condition, which is equivalent to $\mathrm{Ric}_g \geq k$ and $\dim M < N$ when $g(t)$ is independent of t, instead of $\mathsf{CD}(k, \infty)$ since $\dim M = m < \infty$ in our case. However, the following observation suggests us that we should be careful: let us consider (4.1) in the case $k > 0$ and $Z(t) \equiv 0$. When $\partial_t g(t) \equiv 0$, the Bonnet–Myers theorem tells us that the diameter of M is bounded and hence M is compact. Moreover, the Bonnet–Myers theorem still holds under $\mathsf{CD}(k, N)$ in the time-homogeneous case; see [3, 4, 23]. However, when $g(t)$ depends on t, it is no longer true that (4.1) implies the compactness of M. In fact, we can easily obtain a noncompact M enjoying (4.1) with $k > 0$ for some $g(t)$ by following the observation in the second item of this remark.

By a standard argument, Theorem 4.1 implies the following estimate for a gradient of the diffusion semigroup:

COROLLARY 4.3. *Let $((X(t))_{t \in [T_1, T_2]}, (\mathbb{P}_x)_{x \in M})$ be a $\mathscr{L}_t$-diffusion process with $\mathbb{P}_x[X(T_1) = x] = 1$. For any bounded measurable function f on M, let us define $P_t f$ by $P_t f(x) := \mathbb{E}_x[f(X(t))]$. Then, under the same assumption as in Theorem 4.1, we have*

$$\limsup_{y \to x} \left| \frac{P_t f(x) - P_t f(y)}{d_{g(T_1)}(x, y)} \right| \leq \frac{1}{\sqrt{2\pi\beta(t - T_1)}} \sup_{z, z' \in M} |f(z) - f(z')|.$$

In particular, $P_t f$ is $d_{g(T_1)}$-globally Lipschitz continuous when f is bounded.

PROOF. Let $\mathbf{X} = (X_1, X_2)$ be a coupling of $\mathscr{L}_t$-diffusions $(X(t), \mathbb{P}_x)$ and $(X(t), \mathbb{P}_y)$ given in Theorem 4.1. Let τ^* be the coupling time of $\mathbf{X}$, that is, $\tau^* := \inf\{t \in [T_1, T_2] | \mathbf{X}(t) \in D(M)\}$. Let us define $\mathbf{X}^* = (X_1^*, X_2^*)$ of $(X(t), \mathbb{P}_x)$ and $(X(t), \mathbb{P}_y)$ by

$$\mathbf{X}^*(t) := \begin{cases} \mathbf{X}(t), & \text{if } \tau^* > t, \\ (X_1(t), X_1(t)), & \text{otherwise.} \end{cases}$$

Since τ^* is a stopping time with respect to the filtration generated by $\mathbf{X}$, and X_i $(i = 1, 2)$ is a solution to the martingale problem associated with the same filtration, $\mathbf{X}^*$ is again a coupling of $\mathscr{L}_t$-diffusion processes. Since $\{\tau^* > T\} = \{\inf_{T_1 \le t \le T} d_{g(t)}(\mathbf{X}(t)) > 0\}$, Theorem 4.1 yields

$$\begin{aligned} P_t f(x) - P_t f(y) &= \mathbb{E}[f(X_1^*(t)) - f(X_2^*(t))] \\ &= \mathbb{E}\big[\big(f(X_1^*(t)) - f(X_2^*(t))\big) 1_{\{\tau^* > t\}}\big] \\ &\le \mathbb{P}[\tau^* > t] \sup_{z, z' \in M} |f(z) - f(z')| \\ &\le \chi\left(\frac{d_{g(T_1)}(x, y)}{2\sqrt{\beta(t - T_1)}}\right) \sup_{z, z' \in M} |f(z) - f(z')|. \end{aligned}$$

Hence the assertion holds by dividing the both sides of the above inequality by $d_{g(T_1)}(x, y)$ and by letting $y \to x$ after that. $\square$

As we did in the last section, let $(\gamma_{xy}^{(t)})_{x, y \in M}$ be a measurable family of unit-speed minimal $g(t)$-geodesics such that $\gamma_{xy}^{(t)}$ joins x and y. Without loss of generality, we may assume that $\gamma_{xy}^{(t)}$ is symmetric, that is, $\gamma_{xy}^{(t)}(d_{g(t)}(x, y) - s) = \gamma_{yx}^{(t)}(s)$ holds. Let us define $\tilde{m}_{xy}^{(t)} : T_y M \to T_y M$ by

$$\tilde{m}_{xy}^{(t)} v := v - 2\langle v, \dot{\gamma}_{xy}^{(t)} \rangle_{g(t)} \dot{\gamma}_{xy}^{(t)}\big(d_{g(t)}(x, y)\big).$$

This is a reflection with respect to a hyperplane which is $g(t)$-perpendicular to $\dot{\gamma}_{xy}^{(t)}$. Let us define $m_{xy}^{(t)} : T_x M \to T_y M$ by

$$m_{xy}^{(t)}(v) := \tilde{m}_{xy}^{(t)}\big(\big(/\!\!/_{\gamma_{xy}^{(t)}}^{(t)} v\big)\big(d_{g(t)}(x, y)\big)\big).$$

Clearly $m_{xy}^{(t)}$ is a $g(t)$-isometry. As in the last section, let $\Phi^{(t)} : M \to \mathscr{O}^{(t)}(M)$ be a measurable section of the $g(t)$-orthonormal frame bundle $\mathscr{O}^{(t)}(M)$ of M. Let us define two measurable maps $\Phi_i^{(t)} : M \times M \to \mathscr{O}^{(t)}(M)$ for $i = 1, 2$ by

$$\begin{aligned} \Phi_1^{(t)}(x, y) &:= \Phi^{(t)}(x), \\ \Phi_2^{(t)}(x, y) &:= \begin{cases} m_{xy}^{(t)} \Phi_1^{(t)}(x, y), & (x, y) \in M \times M \setminus D(M), \\ \Phi^{(t)}(x), & (x, y) \in D(M). \end{cases} \end{aligned}$$

Take $x_1, x_2 \in M$. By using $\Phi_i^{(t)}$, we define a coupled geodesic random walk $\mathbf{X}^\alpha(t) = (X_1^\alpha(t), X_2^\alpha(t))$ by $X_i^\alpha(T_1) = x_i$ and, for $t \in [t_n^{(\alpha)}, t_{n+1}^{(\alpha)}]$,

$$\begin{aligned} \tilde{\xi}_{n+1}^i &:= \sqrt{m+2}\, \Phi_i^{(t_n^{(\alpha)})}\big(\mathbf{X}^\alpha\big(t_n^{(\alpha)}\big)\big) \xi_{n+1}, \\ X_i^\alpha(t) &:= \exp_{X_i^\alpha(t_n^{(\alpha)})}^{(t_n^{(\alpha)})}\left(\frac{t - t_n^{(\alpha)}}{\alpha^2}\big(\alpha \tilde{\xi}_{n+1}^i + \alpha^2 Z\big(t_n^{(\alpha)}\big)\big)\right) \end{aligned}$$

for $i = 1, 2$. We can easily verify that X_i^α has the same law as X^α with $x_0 = x_i$.

In what follows, we assume (4.1). We can easily verify that it implies Assumption 1. Thus, by Theorem 3.1, $(\mathbf{X}^\alpha)_{\alpha>0}$ is tight under Assumption 1. In addition, a subsequential limit $\mathbf{X}^{\alpha_k} \to \mathbf{X} = (X_1, X_2)$ in law exists, and it is a coupling of two $\mathscr{L}_t$-diffusion processes starting at x_1 and x_2, respectively. We fix such a subsequence $(\alpha_k)_{k\in\mathbb{N}}$. In the rest of this paper, we use the same symbol $\mathbf{X}^\alpha$ for the subsequence $\mathbf{X}^{\alpha_k}$ and the term "$\alpha \to 0$" always means the subsequential limit "$\alpha_k \to 0$."

We will prove that the coupling $\mathbf{X}$ obtained as above is a desired one in Theorem 4.1. We first remark that we can easily verify that X_i $(i = 1, 2)$ is a solution to the martingale problem associated with the filtration generated by $\mathbf{X}$ in the same way as in the proof of Theorem 3.1. Set $\hat{\sigma}_R^i := \sigma_R(d_{g(\cdot)}(o, X_i^\alpha(\cdot)))$ for $i = 1, 2$. We fix $R > 1$ sufficiently large until the beginning of the proof of Theorem 4.1. Let $M_0 \subset M$ be a relatively compact open set satisfying (2.1) for $2R$ instead of R. We next show a difference inequality of $d_{g(t)}(\mathbf{X}^\alpha(t))$. To describe it, we will introduce several notation as in the last section. For simplicity, let us denote $\gamma^{(t_n^{(\alpha)})}_{X_1^\alpha(t_n^{(\alpha)})X_2^\alpha(t_n^{(\alpha)})}$ by $\bar{\gamma}_n$. Let us define a vector field V_{n+1} along $\bar{\gamma}_n$ by

$$V_{n+1} := /\!/^{(t_n^{(\alpha)})}_{\bar{\gamma}_n}\big(\tilde{\xi}^1_{n+1} - \langle \tilde{\xi}^1_{n+1}, \dot{\bar{\gamma}}_n\rangle_{g(t_n^{(\alpha)})}\dot{\bar{\gamma}}_n(0)\big).$$

Take $v \in \mathbb{R}^m$. Let us define λ^*_{n+1} and Λ^*_{n+1} by

$$\lambda^*_{n+1} := \begin{cases} 2\langle \tilde{\xi}^1_{n+1}, \dot{\bar{\gamma}}_n\rangle_{g(t_n^{(\alpha)})}, & \text{if } (y_1, y_2) \notin D(M), \\ 2\sqrt{m+2}\langle \xi_{n+1}, v\rangle, & \text{otherwise,} \end{cases}$$

$$\begin{aligned}\Lambda^*_{n+1} := \frac{1}{2}\bigg(\int_0^{d_{g(t_n^{(\alpha)})}(\mathbf{X}^\alpha(t_n^{(\alpha)}))} &\big(\partial_t g(t_n^{(\alpha)}) + 2(\nabla Z(t_n^{(\alpha)}))^\flat\big) \\ &\times (\dot{\bar{\gamma}}_n(s), \dot{\bar{\gamma}}_n(s))\,ds \\ &+ I^{(t_n^{(\alpha)})}_{\bar{\gamma}_n}(V_{n+1})\bigg)1_{\{\mathbf{X}^\alpha(t_n^{(\alpha)})\notin D(M)\}}.\end{aligned}$$

For $\delta \geq 0$, let us define $\tau_\delta : \mathscr{C}_1 \to [T_1, T_2] \cup \{\infty\}$ by

$$\tau_\delta(w) := \inf\{t \geq T_1 | w(t) \leq \delta\}.$$

We also define $\hat{\tau}_\delta$ by $\hat{\tau}_\delta := \tau_\delta(d_{g(\cdot)}(\mathbf{X}^\alpha(\cdot)))$.

LEMMA 4.4. *For $n \in \mathbb{N}_0$ with $n < N^{(\alpha)}$, we have*

$$\begin{aligned} e^{kt_{n+1}^{(\alpha)}/2} d_{g(t_{n+1}^{(\alpha)})}(\mathbf{X}^\alpha(t_{n+1}^{(\alpha)})) \leq \Big(1 + \frac{k}{2}\Big) &e^{kt_n^{(\alpha)}/2} d_{g(t_n^{(\alpha)})}(\mathbf{X}^\alpha(t_n^{(\alpha)})) \\ &+ e^{kt_n^{(\alpha)}/2}(\alpha\lambda^*_{n+1} + \alpha^2\Lambda^*_{n+1}) \\ &+ o(\alpha^2), \end{aligned} \tag{4.2}$$

when $n < \hat{\tau}_\delta \wedge \hat{\sigma}_R^1 \wedge \hat{\sigma}_R^2$ *and* α *is sufficiently small. Moreover, we can control the error term* $o(\alpha^2)$ *uniformly in the position of* $\mathbf{X}^\alpha$.

PROOF. When $(t_n^{(\alpha)}, \mathbf{X}^\alpha(t_n^{(\alpha)})) \notin \mathrm{Cut}_{\mathrm{ST}}$, (4.2) is just a consequence of the second variational formula for the distance function combined with the index lemma for $I_{\bar{\gamma}_n}^{(t_n^{(\alpha)})}$. To include the case $(t_n^{(\alpha)}, \mathbf{X}^\alpha(t_n^{(\alpha)})) \in \mathrm{Cut}_{\mathrm{ST}}$ and to obtain a uniform control of $o(\alpha^2)$, we extend this argument. Let us define H and $p_1, p_2 : H \to [T_1, T_2] \times \bar{M}_0 \times \bar{M}_0$ by

$$\begin{aligned} H := \bigl\{(t, x, y, z) | t \in [T_1, T_2], x, y, z \in \bar{M}_0, d_{g(t)}(x, y) \geq \delta, \\ d_{g(t)}(x, y) = 2d_{g(t)}(x, z) = 2d_{g(t)}(y, z)\bigr\}, \\ p_1(t, x, y, z) := (t, x, z), \\ p_2(t, x, y, z) := (t, y, z). \end{aligned}$$

If $\mathbf{q} = (t, x, y, z) \in H$, then $p_1(\mathbf{q}), p_2(\mathbf{q}) \notin \mathrm{Cut}_{\mathrm{ST}}$ since z is on a midpoint of a minimal $g(t)$-geodesic joining x and y. Since H is compact, $p_1(H)$ and $p_2(H)$ are also compact. Hence there is a constant $\eta > 0$ such that

$$\begin{aligned} \inf\bigl\{|t - t'| + d_{g(t)}(x, x') + d_{g(t)}(y, y') | (t, x, y) \in p_1(H) \cup p_2(H), \\ (t', x', y') \in \mathrm{Cut}_{\mathrm{ST}}\bigr\} > \eta. \end{aligned}$$

Take $\alpha > 0$ sufficiently small relative to η and δ. Set

$$\begin{aligned} p_n &:= \bar{\gamma}_n\biggl(\frac{d_{g(t_n^{(\alpha)})}(\mathbf{X}^\alpha(t_n^{(\alpha)}))}{2}\biggr), \\ p_n' &:= \exp_{p_n}^{(t_n^{(\alpha)})}\biggl(V_{n+1}\biggl(\frac{d_{g(t_n^{(\alpha)})}(\mathbf{X}^\alpha(t_n^{(\alpha)}))}{2}\biggr)\biggr). \end{aligned}$$

By the triangle inequality, we have

$$\begin{aligned} d_{g(t_n^{(\alpha)})}(\mathbf{X}^\alpha(t_n^{(\alpha)})) &= d_{g(t_n^{(\alpha)})}(X_1^\alpha(t_n^{(\alpha)}), p_n) + d_{g(t_n^{(\alpha)})}(p_n, X_2^\alpha(t_n^{(\alpha)})), \\ d_{g(t_{n+1}^{(\alpha)})}(\mathbf{X}^\alpha(t_{n+1}^{(\alpha)})) &\leq d_{g(t_{n+1}^{(\alpha)})}(X_1^\alpha(t_{n+1}^{(\alpha)}), p_n') + d_{g(t_{n+1}^{(\alpha)})}(p_n', X_2^\alpha(t_{n+1}^{(\alpha)})). \end{aligned}$$

Since $(t_n^{(\alpha)}, \mathbf{X}^\alpha(t_n^{(\alpha)}), p_n) \in H$, we can apply the second variation formula to each term on the right-hand side of the above inequality. Hence we obtain (4.2). For a uniform control of the error term, we remark that $\bar{\gamma}_n$ is included in M_0, and the $g(t_n^{(\alpha)})$-length of $\bar{\gamma}_n$ is bigger than δ. These facts follows from $n < \hat{\tau}_\delta \wedge \hat{\sigma}_R^1 \wedge \hat{\sigma}_R^2$ and the choice of M_0. Thus the every calculation of the second variation formula above is done on a compact subset of $[T_1, T_2] \times M_0 \times M_0$ which is uniformly away from $\mathrm{Cut}_{\mathrm{ST}}$. It yields the desired result. $\square$

Let us define a continuous stochastic process U_a^α on $\mathbb{R}$ starting at a by

$$U_a^\alpha(t) := \mathrm{e}^{-kt/2}a + \alpha \mathrm{e}^{-kt/2}\left(\sum_{j=1}^n \mathrm{e}^{kt_j^{(\alpha)}/2}\lambda_j^* + \frac{t - t_n^{(\alpha)}}{\alpha^2}\mathrm{e}^{kt_n^{(\alpha)}/2}\lambda_{n+1}^*\right).$$

As a final preparation of the proof of Theorem 4.1, we show the following comparison theorem for the distance process of coupled geodesic random walks.

LEMMA 4.5. *For each $\varepsilon > 0$, there exists a family of events $(E_\varepsilon^\alpha)_\alpha$ such that $\mathbb{P}[E_\varepsilon^\alpha]$ converges to 1 as $\alpha \to 0$ and*

$$d_{g(t)}(\mathbf{X}^\alpha(t)) \le U^\alpha_{d_{g(T_1)}(\mathbf{X}^\alpha(T_1))}(t) + \varepsilon \tag{4.3}$$

for all $t \in [T_1, T_2 \wedge \hat{\tau}_\delta \wedge \hat{\sigma}_R^1 \wedge \hat{\sigma}_R^2]$ on E_ε^α for sufficiently small α.

PROOF. In a similar way as in the proof of Lemma 3.10, we can complete the proof once we have found E_ε^α on which (4.3) holds when $t = t_n^{(\alpha)} \in [T_1, T_2 \wedge \hat{\tau}_\delta \wedge \hat{\sigma}_R^1 \wedge \hat{\sigma}_R^2]$. Set $\bar{\Lambda}_{n+1}^* := \mathbb{E}[\Lambda_{n+1}^* | \mathscr{F}_n]$. Then $\sum_{j=1}^n \mathrm{e}^{kt_{j-1}^{(\alpha)}/2}(\Lambda_j^* - \bar{\Lambda}_j^*)$ is an $\mathscr{F}_n$-local martingale. Indeed, Λ_{n+1}^* is bounded if $t_n^{(\alpha)} < \hat{\sigma}_R^1 \wedge \hat{\sigma}_R^2$ and so is $\bar{\Lambda}_{n+1}^*$. Let us define E_ε^α by

$$E_\varepsilon^\alpha := \left\{ \sup_{\substack{N \le N^{(\alpha)} \\ t_N^{(\alpha)} \le T_2 \wedge \hat{\sigma}_R^1 \wedge \hat{\sigma}_R^2}} \sum_{j=1}^{N+1} \mathrm{e}^{kt_j^{(\alpha)}/2}(\Lambda_j^* - \bar{\Lambda}_j^*) \le \frac{\varepsilon}{2\alpha^2(T_2 - T_1)} \right\}.$$

In a similar way as in Lemma 3.8 or [16], Lemma 6, $\lim_{\alpha \to 0} \mathbb{P}[E_\varepsilon^\alpha] = 1$ holds. On E_ε^α, we can replace $\alpha^2 \mathrm{e}^{kt_n^{(\alpha)}/2}\Lambda_{n+1}^*$ in (4.2) with $\alpha^2 \mathrm{e}^{kt_n^{(\alpha)}/2}\bar{\Lambda}_{n+1}^* + \varepsilon/(2(T_2 - T_1))$. Since we have $(m+2)\mathbb{E}[\langle \xi_i, e_k\rangle \langle \xi_i, e_l\rangle] = \delta_{kl}$, we obtain

$$\bar{\Lambda}_{n+1}^* \le -\frac{k}{2} d_{g(t_n^{(\alpha)})}(\mathbf{X}^\alpha(t_n^{(\alpha)})).$$

Thus an iteration of Lemma 4.4 implies (4.3) on E_ε^α when $t = t_n^{(\alpha)}$. $\square$

PROOF OF THEOREM 4.1. Take $\varepsilon \in (0, 1)$ arbitrarily. Let $R > 1$ be sufficiently large so that

$$\limsup_{\alpha \to 0} \mathbb{P}[\hat{\sigma}_R^1 \wedge \hat{\sigma}_R^2 \le T_2] < \varepsilon.$$

It is possible by Proposition 3.4. Set $a := d_{g(T_1)}(x_1, x_2)$. Take $T \in [T_1, T_2]$, and let $\delta > 0$ be $\delta > 2\varepsilon$. Then Lemma 4.5 yields

$$\begin{aligned}\mathbb{P}[\hat{\tau}_\delta > T] &\le \mathbb{P}[\{\hat{\tau}_\delta > T\} \cap E_\varepsilon^\alpha \cap \{\hat{\sigma}_R^1 \wedge \hat{\sigma}_R^2 > T\}] + 2\varepsilon \\ &\le \mathbb{P}[\tau_{\delta/2}(U_a^\alpha) > T] + 2\varepsilon.\end{aligned}$$

Thus we obtain

$$\limsup_{\alpha\to 0}\mathbb{P}[\hat{\tau}_\delta > T] \le \limsup_{\alpha\to 0}\mathbb{P}\Big[\inf_{t\in[T_1,T]} U_a^\alpha(t) \ge \delta/2\Big]$$

by letting $\varepsilon \downarrow 0$. Note that U_a^α converges in law to U_a as $\alpha \to 0$. Since

$$\{\mathbf{w} \in C([T_1, T_2] \to M \times M) | \tau_\delta(d_{g(\cdot)}(\mathbf{w}(\cdot))) > T\}$$

is open, and $\{w | \inf_{t\in[T_1,T_2]} w(t) \ge \delta/2\}$ is closed in $C([0, T] \to \mathbb{R})$, the Portmanteau theorem yields

$$\begin{aligned}\mathbb{P}\Big[\inf_{T_1\le t\le T} d_{g(t)}(\mathbf{X}(t)) > \delta\Big] &\le \liminf_{\alpha\to 0}\mathbb{P}[\hat{\tau}_\delta > T]\\ &\le \limsup_{\alpha\to 0}\mathbb{P}\Big[\inf_{t\in[T_1,T]} U_a^\alpha(t) \ge \delta/2\Big]\\ &\le \mathbb{P}\Big[\inf_{t\in[T_1,T]} U_a(t) \ge \delta/2\Big].\end{aligned}$$

Therefore the conclusion follows by letting $\delta \downarrow 0$. □

5. Coupling by parallel transport. As a final part of the paper, we will see that we can also construct a coupling by parallel transport by following our manner. In the construction of the coupling by reflection, we used a map $m_{xy}^{(t)}$. By following the same argument after omitting $\tilde{m}_{xy}^{(t)}$ in the definition of $m_{xy}^{(t)}$, we obtain a coupling by parallel transport. The difference of it from the coupling by reflection is the absence of the term corresponding to λ_n^*, which comes from the first variation of arc length. As a result, we can show the following; cf. [16]:

THEOREM 5.1. *Assume* (4.1). *For $x_1, x_2 \in M$, there is a coupling $\mathbf{X}(t) = (X_1(t), X_2(t))$ of two $\mathscr{L}_t$-diffusion processes starting at x_1 and x_2 at time T_1, respectively, such that*

$$d_{g(t)}(\mathbf{X}(t)) \le \mathrm{e}^{-k(t-s)/2} d_{g(s)}(\mathbf{X}(s))$$

for $T_1 \le s \le t \le T_2$ almost surely.

It recovers a part of results studied in [1]. In particular, a contraction type estimate for Wasserstein distances under the heat flow follows.

PROOF OF THEOREM 5.1. Let us construct a coupling by parallel transport of geodesic random walks $\mathbf{X}^\alpha = (X_1^\alpha, X_2^\alpha)$ starting at $(x_1, x_2) \in M \times M$ by following the procedure stated just before Theorem 5.1. By taking a subsequence, we may assume that $\mathbf{X}^\alpha$ converges in law as $\alpha \to 0$. We denote the limit by $\mathbf{X} = (X_1, X_2)$. In what follows, we prove

$$\mathbb{P}\Big[\sup_{T_1\le s\le t\le T_2}\big(\mathrm{e}^{kt/2} d_{g(t)}(\mathbf{X}(t)) - \mathrm{e}^{ks/2} d_{g(s)}(\mathbf{X}(s))\big) > \varepsilon\Big] = 0$$

for any $\varepsilon > 0$. By virtue of the Portmanteau theorem together with Proposition 3.4, it suffices to show

$$\lim_{\alpha\to 0}\mathbb{P}\Big[\sup_{T_1\le s\le t\le T_2}\big(\mathrm{e}^{kt/2}d_{g(t)}(\mathbf{X}^\alpha(t)) - \mathrm{e}^{ks/2}d_{g(s)}(\mathbf{X}^\alpha(s))\big) > \varepsilon, \\ \hat\sigma_R^1\wedge\hat\sigma_R^2 = \infty\Big] = 0 \tag{5.1}$$

for any $R > 1$. We write $d_n := \mathrm{e}^{kt_n^{(\alpha)}/2}d_{g(t_n^{(\alpha)})}(\mathbf{X}^\alpha(t_n^{(\alpha)}))$ in this proof for simplicity of notation. For $\delta > 0$, let us define a sequence of $\mathscr{F}_n$-stopping times S_l by $S_0 := 0$ and

$$S_{2l+1} := \inf\{j \ge S_{2l} | d_j \le \delta\} \wedge N^{(\alpha)},$$
$$S_{2l} := \inf\{j \ge S_{2l-1} | d_j \ge 2\delta\} \wedge N^{(\alpha)}.$$

Note that $d_{S_{2l-1}} \le 3\delta$ holds on $\{\hat\sigma_R^1\wedge\hat\sigma_R^2 = \infty\}$ for sufficiently small α. As mentioned just before Theorem 5.1, the difference inequality (4.2) holds with $\lambda^* = 0$ when $S_{2l-1} \le n < S_{2l}\wedge\bar\sigma_R^1\wedge\bar\sigma_R^2$ for some $l \in \mathbb{N}_0$. In this case, the error term $o(\alpha^2)$ is controlled uniformly also in l. Let us define an event E_δ^α by

$$E_\delta^\alpha := \left\{\sup_{\substack{n\le N\le N^{(\alpha)}\\ t_N^{(\alpha)}\le T_2\wedge\hat\sigma_R^1\wedge\hat\sigma_R^2}}\sum_{j=n+1}^{N+1}\mathrm{e}^{kt_j^{(\alpha)}/2}(\Lambda_j^* - \bar\Lambda_j^*) \le \frac{\delta}{2\alpha^2}\right\}.$$

Then, as in Lemmas 3.8 and 4.5, we can show $\lim_{\alpha\to 0}\mathbb{P}[E_\delta^\alpha] = 1$. On $E_\delta^\alpha \cap \{\hat\sigma_R^1\wedge\hat\sigma_R^2 = \infty\}$, we have $d_N \le d_n + \delta$ for $S_{2l-1} \le n \le N \le S_{2l}$ if α is sufficiently small. Moreover, for $n < S_{2l-1} \le N < S_{2l}$,

$$d_N - d_n \le (d_N - d_{S_{2l-1}}) + d_{S_{2l-1}} \le 5\delta.$$

In the case $S_{2l} \le N < S_{2l+1}$, we obtain $d_N - d_n \le 2\delta$. Thus $d_N - d_n \le 5\delta$ holds for all $n < N$ on $E_\delta^\alpha \cap \{\hat\sigma_R^1\wedge\hat\sigma_R^2 = \infty\}$. Take $\delta > 0$ less than $\varepsilon/10$. Then our observations yield (5.1) since $d_{g(t)}(\mathbf{X}^\alpha(t)) - d_{g(\lfloor t\rfloor_\alpha)}(\mathbf{X}(\lfloor t\rfloor_\alpha))$ becomes uniformly small on $\{\hat\sigma_R^1\wedge\hat\sigma_R^2 = \infty\}$ as $\alpha \to 0$. $\square$

REFERENCES

[1] ARNAUDON, M., COULIBALY, K. A. and THALMAIER, A. (2010). Horizontal diffusion in C^1-path space. In *Séminaire de Probabilités XLIII*. *Lecture Notes in Math*. **2006** 73–94. Springer, Berlin.

[2] BAKRY, D. (1997). On Sobolev and logarithmic Sobolev inequalities for Markov semigroups. In *New Trends in Stochastic Analysis* (*Charingworth*, 1994) 43–75. World Scientific, River Edge, NJ. MR1654503

[3] BAKRY, D. and LEDOUX, M. (1996). Sobolev inequalities and Myers's diameter theorem for an abstract Markov generator. *Duke Math. J.* **85** 253–270. MR1412446

[4] BAKRY, D. and QIAN, Z. (2005). Volume comparison theorems without Jacobi fields. In *Current Trends in Potential Theory*. *Theta Ser. Adv. Math.* **4** 115–122. Theta, Bucharest. MR2243959

[5] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley, New York. MR1700749

[6] BLUM, G. (1984). A note on the central limit theorem for geodesic random walks. *Bull. Aust. Math. Soc.* **30** 169–173. MR0759783

[7] CHAVEL, I. (1993). *Riemannian Geometry—a Modern Introduction*. *Cambridge Tracts in Mathematics* **108**. Cambridge Univ. Press, Cambridge. MR1271141

[8] COULIBALY-PASQUIER, K. A. (2011). Brownian motion with respect to time-changing Riemannian metrics, applications to Ricci flow. *Ann. Inst. Henri Poincaré Probab. Stat.* To appear. Available at arXiv:0901.1999.

[9] CRANSTON, M. (1991). Gradient estimates on manifolds using coupling. *J. Funct. Anal.* **99** 110–124. MR1120916

[10] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes*: *Characterization and Convergence*. Wiley, New York. MR0838085

[11] FREEDMAN, D. A. (1975). On tail probabilities for martingales. *Ann. Probab.* **3** 100–118. MR0380971

[12] HSU, E. P. (2002). *Stochastic Analysis on Manifolds*. *Graduate Studies in Mathematics* **38**. Amer. Math. Soc., Providence, RI. MR1882015

[13] IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. North-Holland, Amsterdam. MR1011252

[14] JØRGENSEN, E. (1975). The central limit problem for geodesic random walks. *Z. Wahrsch. Verw. Gebiete* **32** 1–64. MR0400422

[15] KENDALL, W. S. (1998). From stochastic parallel transport to harmonic maps. In *New Directions in Dirichlet Forms*. *AMS/IP Studies in Advanced Mathematics* **8** 49–115. Amer. Math. Soc., Providence, RI. MR1652279

[16] KUWADA, K. (2010). Couplings of the Brownian motion via discrete approximation under lower Ricci curvature bounds. In *Probabilistic Approach to Geometry*. *Advanced Studies in Pure Mathematics* **57** 273–292. Math. Soc. Japan, Tokyo. MR2648265

[17] KUWADA, K. and PHILIPOWSKI, R. (2011). Non-explosion of diffusion processes on manifolds with time-dependent metric. *Math. Z.* To appear. Available at arXiv:0910.1730.

[18] KUWADA, K. and PHILIPOWSKI, R. (2011). Coupling of Brownian motion and Perelman's $\mathcal{L}$-functional. *J. Funct. Anal.* **260** 2742–2766.

[19] MCCANN, R. J. and TOPPING, P. M. (2010). Ricci flow, entropy and optimal transportation. *Amer. J. Math.* **132** 711–730. MR2666905

[20] OSHIMA, Y. (2004). Time-dependent Dirichlet forms and related stochastic calculus. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **7** 281–316. MR2066136

[21] PHILIPOWSKI, R. (2009). Coupling of diffusions on manifolds with time-dependent metric. Seminar talk at Universtität Bonn.

[22] PINSKY, M. A. (1976). Isotropic transport process on a Riemannian manifold. *Trans. Amer. Math. Soc.* **218** 353–360. MR0402957

[23] QIAN, Z. (1997). Estimates for weighted volumes and applications. *Quart. J. Math. Oxford Ser.* (2) **48** 235–242. MR1458581

[24] STANNAT, W. (1999). The theory of generalized Dirichlet forms and its applications in analysis and stochastics. *Mem. Amer. Math. Soc.* **142** viii+101. MR1632609

[25] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **233**. Springer, Berlin. MR0532498

[26] TOPPING, P. (2009). $\mathcal{L}$-optimal transportation for Ricci flow. *J. Reine Angew. Math.* **636** 93–122. MR2572247

[27] VON RENESSE, M.-K. (2004). Intrinsic coupling on Riemannian manifolds and polyhedra. *Electron. J. Probab.* **9** 411–435 (electronic). MR2080605

[28] WANG, F. Y. (1994). Successful couplings of nondegenerate diffusion processes on compact manifolds. *Acta Math. Sinica* **37** 116–121. MR1272513

[29] WANG, F.-Y. (1997). On estimation of the logarithmic Sobolev constant and gradient estimates of heat semigroups. *Probab. Theory Related Fields* **108** 87–101. MR1452551

[30] WANG, F. Y. (2005). *Functional Inequalities, Markov Semigroups, and Spectral Theory*. *Mathematics Monograph Series* **4**. Science Press, Beijing, China.

[31] WILLETT, D. and WONG, J. S. W. (1965). On the discrete analogues of some generalizations of Gronwall's inequality. *Monatsh. Math.* **69** 362–367. MR0185175

[32] ZHANG, Q. S. (2011). *Sobolev Inequalities, Heat Kernels Under Ricci Flow, and the Poincaré Conjecture*. CRC Press, Boca Raton, FL. MR2676347

GRADUATE SCHOOL OF HUMANITIES AND SCIENCES
OCHANOMIZU UNIVERSITY
OHTSUKA 2-1-1, BUNKYO-KU
TOKYO 112-8610
JAPAN
E-MAIL: kuwada.kazumasa@ocha.ac.jp