Electronic Journal of Probability

Vol. 9 (2004), Paper no. 14, pages 411-435.

Journal URL
http://www.math.washington.edu/~ejpecp/

Intrinsic Coupling on Riemannian Manifolds and Polyhedra

Max-K. von Renesse[1]
mrenesse@math.tu-berlin.de
Institut für Mathematik
Technische Universität Berlin
Germany

Abstract

Starting from a central limit theorem for geometric random walks we give an elementary construction of couplings between Brownian motions on Riemannian manifolds. This approach shows that cut locus phenomena are indeed inessential for Kendall's and Cranston's stochastic proof of gradient estimates for harmonic functions on Riemannian manifolds with lower curvature bounds. Moreover, since the method is based on an asymptotic quadruple inequality and a central limit theorem only it may be extended to certain non smooth spaces which we illustrate by the example of Riemannian polyhedra. Here we also recover the classical heat kernel gradient estimate which is well known from the smooth setting.

AMS 2000 Subject Classification: Primary 60J60; secondary 60J45,58J50.

Keywords: Coupling, Gradient Estimates, Central Limit Theorem, Quadruple Inequality.

Submitted December 12 11, 2003. Final version accepted April 19, 2004.

[1]This paper was written while the author was member of the Collaborative Research Center SFB611 of the German Research Council (DFG) at Bonn University.

1. Introduction. The Riemannian *coupling by reflection* technique by Kendall [Ken86] was used by Cranston [Cra91] to give an elegant stochastic proof of L^∞-gradient estimates for harmonic functions on Riemannian manifolds with lower Ricci curvature bounds. The stronger pointwise bounds due to Yau [Yau75] are usually proved by analytic arguments based on the fundamental Bochner-Lichnerowicz formula which is difficult to transfer to non smooth (cf. [CH98]) or even non Riemannian situations. Therefore one may turn to the more flexible stochastic methods for the analysis of second order differential operators on non smooth or metric measure spaces (X, d, m) in terms of their associated Markov processes. This general agenda constitutes the background motivation for the present attempt to simplify Cranston's stochastic proof of gradient estimates whose beautiful geometric content is hidden behind a sophisticated and, as it turns out, dispensable technical superstructure.

The basic idea behind the coupling by reflection method on Riemannian manifolds (M, g) is to construct a stochastic process Ξ on the product $M \times M$ such that

i) each factor $\Xi_1 = \pi_1(\Xi)$ and $\Xi_2 = \pi_2(\Xi)$ is a Brownian motion on (M, g)

ii) the compound process $d(\Xi)$ of Ξ with the intrinsic distance function d on M is dominated by a real semi-martingale ξ whose hitting time at zero $T_N(\xi)$ can be estimated from above.

The standard construction of Ξ put forth in [Ken86] uses SDE theory on $M \times M$ and hence requires a certain degree of smoothness on the coefficients. Unfortunately these coefficients typically become singular on the diagonal and, more severely, on the cut locus $\mathrm{Cut}(M) \subset M \times M$. Even if cut locus phenomena in geometric stochastic analysis have been addressed occasionally (cf. [CKM93, Wan94, MS96]) they remain a delicate issue especially for Kendall's coupling method.

We overcome these difficulties by introducing a more intrinsic construction of the coupling process on $M \times M$ which yields the essential coupling probability estimate irrespective if the manifold (M, g) is Cartan-Hadamard or not. Moreover, a brief analysis of the proofs reveals that except an asymptotic quadruple inequality for geodesics in (M, g) and a central limit property for random walks no further regularity of (M, g) seen as a metric measure space (X, d, m) is required. Hence, the method may be applied in more general situations as we indicate by the example of certain Riemannian polyhedra. Finally we point out that the simpler *coupling by parallel transport* method works well also in the polyhedron case from which we deduce gradient estimates for the corresponding heat semigroup.

Acknowledgments: To Karl-Theodor Sturm for suggesting the problem.

2. Preliminaries. *2.1. Riemannian Central Limit Theorem [Jør75].* Let (M, g) be a smooth Riemannian manifold of dimension n and fix for every $x \in M$ an isometry $\Phi_x : \mathbb{R}^n \overset{\simeq}{\to} T_x M$ such that the resulting function

$$\Phi(.) : M \to O(M), \Phi(x) : \mathbb{R}^n \overset{\simeq}{\to} T_x M \quad \forall x \in M$$

is measurable. Let $(\xi_k)_{k \in \mathbb{N}}$ be a sequence of $\mathbb{R}^n$-valued and independent random variables defined on some probability space $(\Omega, \mathcal{O}, P)$ whose distribution equals the normalized uniform distribution on S^{n-1}. A geodesic random walk $(\Xi_k^{\epsilon,x})_{k\in\mathbb{N}}$ with step size $\epsilon > 0$ and starting point $x \in M$ is given inductively by

$$\begin{aligned}\Xi_0^{\epsilon,x} &= x\\ \Xi_{k+1}^{\epsilon,x} &= \exp_{\Xi_k^{\epsilon,x}}(\epsilon\Phi_{\Xi_k^{\epsilon,x}}\xi_{k+1}),\end{aligned}$$

where exp is the exponential map of (M,g). Then one may consider the sequence of subordinated process $\tilde{\Xi}_k^x(t) := \Xi_{\tau_k(t)}^{1/\sqrt{k},x}$ for $k \in \mathbb{N}_0$, where τ_k is a Poisson jump process on $\mathbb{N}$ with parameter k. For k fixed $\Xi_k^x(t)$ is a (time homogeneous) Markov process with transition function

$$\begin{aligned}P(\tilde{\Xi}_k^x(t) \in A \,|\, \tilde{\Xi}_k^x(s) = y) &= e^{-(t-s)k}\sum_{i\geq 0}\frac{((t-s)k)^i}{i!}\mu_{1/\sqrt{k}}^i(y,A)\\ &=: (\tilde{P}_{t-s}^k 1\!\!1_A)(y)\end{aligned}$$

with $\mu_\epsilon(z,A) = ⨍_{S_z^{n-1}\subset T_zM} 1\!\!1_A(\exp_z(\epsilon\theta))d\theta$ and $\mu^i = \mu\circ\mu\cdots\circ\mu$. The generator of $\tilde{\Xi}^k$, or equivalently of the semigroup $(\tilde{P}_t^k)_{t\geq 0}$, is therefore given by

$$A_kf(x) = k(⨍_{S_x^{n-1}} f(\exp_x(\frac{1}{\sqrt{k}}\theta)d\theta - f(x))) \stackrel{k\to\infty}{\longrightarrow} Af(x) \qquad (1)$$

where $Af(x) = 1/2n\Delta f(x)$ with the Laplace-Beltrami operator $\Delta f(x) = \mathrm{trace}(\mathrm{Hess}\, f)(x)$ on (M,g), see [Blu84, vR02]. Using (1) and Kurtz' semigroup approximation theorem it is easy to show that

$$\tilde{P}_t^k \longrightarrow P_t \text{ for } k \longrightarrow \infty$$

in the strong operator sense where $P_t = e^{tA} = P_{t/2n}^M$ is the heat semigroup $P_t^M = e^{t\Delta}$ on (M,g) after a linear time change $t \to t/2n$. Thus weak convergence for the family $\tilde{\Xi}_\cdot^k$ to a (time changed) Brownian motion $\Xi_\cdot^x$ starting in x is obtained from standard arguments showing that the sequence of distributions $\tilde{\Xi}_\cdot^k$ is tight on the Skorokhod path space $D_{\mathbb{R}_+}(M)$.

2.2. Coupled Random Walks. Instead of using the SDE approach to the construction of a coupling of two Brownian motions on (M, g) we follow the lines of the Markov chain approximation scheme for solutions to martingale problems for degenerate diffusion operators in the sense of [SV79] where the approximating processes are coupled geodesic random walks on $M \times M$.

Let $D(M) = \{(x, x) \mid x \in M\}$ be the diagonal in $M \times M$. Then for all $x, y \in M \times M \setminus D(M)$ choose some minimal geodesic $\gamma_{xy} : [0,1] \to M$ connecting x and y and fix a function

$$\Phi(.,.) : M \times M \setminus D(M) \to O(M) \times O(M)$$
$$\Phi_1(x, y) := \pi_1 \circ \Phi(x, y) : \mathbb{R}^n \overset{\simeq}{\to} T_x M$$
$$\Phi_2(x, y) := \pi_2 \circ \Phi(x, y) : \mathbb{R}^n \overset{\simeq}{\to} T_y M$$

with the additional property that

$$\Phi_1(x, y)e_1 = \frac{\dot{\gamma}_{xy}(0)}{\|\dot{\gamma}_{xy}(0)\|}, \Phi_2(x, y)e_1 = \frac{\dot{\gamma}_{yx}(0)}{\|\dot{\gamma}_{yx}(0)\|} \text{ if } x \neq y \tag{$*$}$$

where e_i is the i-th unit vector in $\mathbb{R}^n$. On the diagonal $D(M)$ we set

$$\Phi(x, x) := (\phi(x), \phi(x)) \in O_x(M) \times O_x(M) \tag{$*_D$}$$

where $\phi : M \to O(M)$ is some choice of bases as in the previous paragraph.

In the existence and regularity statement for a possible choice of Φ below the set $\mathrm{Cut}(M) \subset M \times M$ is defined as the collection of all pairs of points (x, y) which can be joined by at least two distinct minimal geodesics, hence $\mathrm{Cut}(M)$ itself is symmetric and measurable.

Lemma 1. *There is some choice of a minimal geodesic γ_{xy} (parameterized on $[0,1]$) for each $(x, y) \in M \times M$ such that the resulting map $\gamma : M \times M \to C^1([0,1], M)$, $(x, y) \mapsto \gamma_{xy}$ is measurable, symmetric, i.e. $\gamma_{xy}(t) = \gamma_{yx}(1-t)$ for all $t \in [0,1]$, and continuous on $M \times M \setminus (D(M) \cup \mathrm{Cut}(M))$. Furthermore, for any measurable frame map $\phi : M \to \Gamma(O(M))$ it is possible to find a measurable function $\Phi : M \times M \to O(M) \times O(M)$ satisfying the conditions ($*$) and ($*_D$) above and which is continuous on $M \times M \setminus (D(M) \cup \mathrm{Cut}(M))$.*

Proof. Suppose first that we found a measurable symmetric function $\gamma : M \times M \to C^1([0,1], M)$ as above and let $\psi_i \in \Gamma(O(M))$, $i = 1, 2$ be two arbitrary continuous frame maps on M. For $(x, y) \in M \times M \setminus D(M)$ we construct a new orthonormal frame on $T_x M \oplus T_y M$ by $\Phi(x, y) = \{\dot{\gamma}_{xy}/\|\dot{\gamma}_{xy}\|, \tilde{\psi}_1^2, \ldots, \tilde{\psi}_1^n, \dot{\gamma}_{yx}/\|\dot{\gamma}_{yx}\|, \tilde{\psi}_2^2, \ldots, \tilde{\psi}_2^n\}$ out of the frame $\{\psi_1(x),$
$\psi_2(y)\}$ via Schmidt's orthogonalization procedure applied to the vectors $\psi_i(e_k)$, $k = 1, \ldots, n$ in the orthogonal complements of $\dot{\gamma}_{xy}$ and $\dot{\gamma}_{yx}$ in $T_x M$ and $T_y M$ respectively. Since the maps $\partial_{s_{|s=0}}$ and $\partial_{s_{|s=1}} : C^1([0,1], M) \to TM$ are continuous and the construction of the basis Φ in $T_x M \oplus T_y M$ depends continuously on the data $\{\psi_1, \psi_2\}$, $\dot{\gamma}_{xy}$ and

$\dot\gamma_{yx}$ it is clear that the map Φ inherits the regularity properties of the function γ on $M \times M \setminus D(M)$. Since $D(M)$ is closed in $M \times M$ and hence measurable any extension of Φ by a measurable $(\phi(.), \phi(.))$ as above on $D(M)$ yields a measurable map on the whole $M \times M$. This proves the second part of the lemma.

Thus it remains to find a map γ as desired. In order to deal with the symmetry condition we first introduce a continuous complete ordering $\geq$ on M (which can be obtained as an induced ordering from an embedding of M into a high dimensional Euclidean space $\mathbb{R}^l$ and some complete ordering on $\mathbb{R}^l$) and restrict the discussion to the closed subset $D_-(M) = \{(x,y) \,|\, x \geq y\} \subset M \times M$ endowed with its Borel σ-algebra which is the trace of $\mathcal{B}(M \times M)$ on $D_-(M)$. We define a measurable set-valued map $\Gamma : D_-(M) \to 2^{C^1([0,1],M)}$ as follows: for each $\epsilon > 0$ choose some ϵ-net $P^\epsilon = \{p_i^\epsilon \,|\, i \in \mathbb{N}\}$ in $D_-(M)$ and choose some minimal geodesic $\gamma_{p_i^\epsilon, p_j^\epsilon}$ for each pair of points p_i^ϵ, $p_j^\epsilon \in P^\epsilon$. Arrange the set of pairs $(p_i^\epsilon, p_j^\epsilon)$ into a common sequence $\{(p_{i_k}^\epsilon, p_{j_k}^\epsilon) \,|\, k \in \mathbb{N}\}$ and let $\gamma^\epsilon : D_-(M) \to C^1([0,1],M)$ be the map defined inductively by

$$\gamma^\epsilon(x,y) = \gamma_{p_{i_0}^\epsilon p_{j_0}^\epsilon} \text{ for } (x,y) \in B_{2\epsilon}(p_{i_0}^\epsilon, p_{j_0}^\epsilon)$$

$$\gamma^\epsilon(x,y) = \gamma_{p_{i_{k+1}}^\epsilon p_{j_{k+1}}^\epsilon} \text{ for } (x,y) \in B_{2\epsilon}(p_{i_{k+1}}^\epsilon, p_{j_{k+1}}^\epsilon) \setminus \bigcup_{l=0}^{k} B_{2\epsilon}(p_{i_l}^\epsilon, p_{j_l}^\epsilon)$$

It is clear from the definition that the functions γ^ϵ are measurable and, moreover, using the geodesic equation in (M,g) together with the Arzela-Ascoli-theorem it is easy to see that for each $(x,y) \in D_-(M)$ the set of curves $\{\gamma_{xy}^\epsilon\}_{\epsilon>0}$ are relatively compact in $C^1([0,1],M)$. Trivially any limit point of $\{\gamma_{xy}^\epsilon\}_{\epsilon>0}$ for ϵ tending to zero will be a minimal geodesic from x to y. Let us choose a priori some sequence $\epsilon_k \to 0$ for $k \to \infty$ then we define the set valued function $\Gamma : D_-(M) \to 2^{C^1([0,1],M)}$ for $(x,y) \in D_-(M)$ as the collection of all possible limit points of $\Gamma^{\epsilon_k}(x,y)$, i.e. $\Gamma : M \times M \to\subset C^1([0,1],M)$ with

$$\Gamma(x,y) := \left\{ \gamma_{xy} \;\middle|\; \begin{array}{c} \exists \text{ subsequence } \epsilon_{k'} \text{ and } \gamma_{xy}^{\epsilon_{k'}} \in \Gamma^{\epsilon_{k'}}(x,y) : \\ \gamma_{xy}^{\epsilon_{k'}} \to \gamma_{xy} \text{ in } C^1([0,1],M) \text{ for } k' \to \infty \end{array} \right\}$$

The fact that we can find a measurable 'selector', i.e. a measurable map $\gamma : D_-(M) \to C^1([0,1],M)$ with $\gamma(x,y) \in \Gamma(x,y)$ follows from a measurable selection theorem as formulated in the subsequent lemma. Furthermore, the uniqueness of γ_{xy} and compactness arguments imply that any such selector obtained from the map Γ above must be continuous on $D_-(M) \cap (\mathrm{Cut}(M) \cup (D(M))^c$. It is also clear that γ_{xx} is the constant curve in x for all $x \in M$ and hence we may extend our chosen γ from $D_-(M)$ continuously onto the whole $M \times M$ by putting $\gamma_{yx}(t) := \gamma_{xy}(1-t)$ if $(x,y) \in D_-(M)$. This proves the first assertion of the lemma and the proof is completed. □

Lemma 2. *Let $(X, \mathcal{S})$ be a measurable and (Y,d) be a complete separable metric space endowed with its Borel σ-algebra $\mathcal{B}(Y)$. Let furthermore $f_k : X \to Y$ be a sequence of measurable functions which are pointwise relatively compact, i.e. for all x in X the set $\{f_k(x)\}_{k\in\mathbb{N}}$ is relatively compact in Y. Let the set valued map $F : X \to 2^Y$ be defined by pointwise collecting all possible limit points of the sequence f_k. Then there is a measurable function $f : X \to Y$ with $f(x) \in F(x)$ for all $x \in X$.*

Proof. Since the set $F(x)$ is obviously closed for any x in X it remains to check the measurability of F, i.e. we need to show that $F^{-1}(O) := \{x | F(x) \cap O \neq \emptyset\}$ is measurable in X for any $O \subset Y$ open. Since any open $O \subset Y$ can be exhausted by countably many set of the type $\overline{B_\delta(y)}$ with $\delta > 0$, $y \in Y$ we may replace O by $\overline{B_\delta(y)}$ in the condition above. But using the pointwise compactness of the sequence f_k and a diagonal sequence argument it is easy to show that

$$F^{-1}\left(\overline{B_\delta(y)}\right) = \bigcap_{\delta' > \delta} \limsup_{k \to \infty} f_k^{-1}\left(B_{\delta'}(y)\right).$$

Choosing some sequence $\delta_l' \searrow \delta$ we see that in fact $F^{-1}\left(\overline{B_\delta(y)}\right)$ is measurable. Hence we may apply the measurable selection theorem of Kuratowksi and Ryll-Nardzewski to the function F which yields the claim. $\square$

We now take two independent sequences $(\xi_k)_{k\in\mathbb{N}}$ and $(\eta_k)_{k\in\mathbb{N}}$ of $\mathbb{R}^n$-valued i.i.d. random variables with normalized uniform distribution on S^{n-1} and define a coupled geodesic random walk $\Xi_k^{\epsilon,(x,y)} = (\Xi_{1,k}^{\epsilon,(x,y)}, \Xi_{2,k}^{\epsilon,(x,y)})$ with step size ϵ and starting point (x, y) in $M \times M$ inductively by

$$\Xi_0^{\epsilon,(x,y)} = (x, y)$$

and if $\Xi_k^{\epsilon,(x,y)} \in M \times M \setminus D(M)$:

$$\begin{aligned}\Xi_{k+1}^{\epsilon,(x,y)} = \Big(&\exp_{\pi_1(\Xi_k^{\epsilon,(x,y)})}[\epsilon \Phi_1(\Xi_k^{\epsilon,(x,y)})\xi_{k+1}],\\ &\exp_{\pi_2(\Xi_k^{\epsilon,(x,y)})}[\epsilon \Phi_2(\Xi_k^{\epsilon,(x,y)})\xi_{k+1}]\Big)\end{aligned} \tag{2}$$

if $\Xi_k^{\epsilon,(x,y)} \in D(M)$:

$$\begin{aligned}\Xi_{k+1}^{\epsilon,(x,y)} = \Big(&\exp_{\pi_1(\Xi_k^{\epsilon,(x,y)})}[\epsilon \phi(\pi_1(\Xi_k^{\epsilon,(x,y)}))\xi_{k+1}],\\ &\exp_{\pi_1(\Xi_k^{\epsilon,(x,y)})}[\epsilon \phi(\pi_1(\Xi_k^{\epsilon,(x,y)}))\eta_{k+1}]\Big)\end{aligned} \tag{3}$$

where π_i, $i = 1, 2$ are the projections of $M \times M$ on the first and second factor respectively. We have two canonical possibilities to extend $\Xi_k^{\epsilon,(x,y)}$ to a process with continuous time parameter $t \in \mathbb{R}_+$, namely

i) by geodesic interpolation $\hat{\Xi}_t^{\epsilon,(x,y)}$,
ii) by Poisson subordination $\tilde{\Xi}_{\tau_\lambda(t)}^{\epsilon,(x,y)}$.

In particular, choosing $\epsilon = 1/\sqrt{k}$ and $\lambda = k$ in ii) for $k \in \mathbb{N}$ one obtains a sequence of Markov processes $\tilde{\Xi}_t^{k,(x,y)} = \Xi_{\tau_k(t)}^{1/\sqrt{k},(x,y)}$ on $M \times M$ with transition function

$$\begin{aligned}&P\big(\tilde{\Xi}^{k,(x,y)}(t) \in A \times B \big| \tilde{\Xi}^{k,(x,y)}(s) = (u, v)\big)\\ &\qquad = e^{-(t-s)k} \sum_{i \geq 0} \frac{((t-s)k)^i}{i!} \mu_{1/\sqrt{k}}^i((u, v), A \times B)\end{aligned}$$

where the kernel $\mu_\epsilon : M^2 \times \mathcal{B}(M^2) \to \mathbb{R}$ is given by

$$\mu_\epsilon((u,v), A \times B) = \begin{cases} \displaystyle\fint_{S_0^{n-1} \subset \mathbb{R}^n} 1\!\!1_A(\exp_u(\epsilon \Phi^1_{(u,v)}\theta)) 1\!\!1_B(\exp_u(\epsilon \Phi^2_{(u,v)}\theta)) d\theta & \text{if } (u,v) \in D(M)^c \\ \displaystyle\fint_{S_u^{n-1} \subset T_u M} 1\!\!1_A(\exp_u(\epsilon\theta)) d\theta \cdot \fint_{S_v^{n-1} \subset T_v M} 1\!\!1_B(\exp_v(\epsilon\theta)) d\theta & \text{else.} \end{cases}$$

The generator of the semigroup $(\tilde{P}_t^{k,(x,y)})_{t \geq 0}$ induced by $\tilde{\Xi}^{k,(x,y)}$ is

$$L_k = k(\mu_{1/\sqrt{k}} - Id).$$

3. Coupling Central Limit Theorem. The following lemma is a partial but for our aim sufficient characterization of any limit of the sequence of operators $(L_k)_{k \in \mathbb{N}}$.

Lemma 3. *Let $F : M \times M \to \mathbb{R}$ be a smooth function. Then*

$$L_k F(u,v) \longrightarrow L_c F(u,v) \quad \forall (u,v) \in M \times M, \textit{locally uniformly on } D(M)^c$$

for $k \to \infty$, where the operator $L_c = L_c^{M,\Phi}$ is defined by

$$L_c(f \otimes g) = Af \otimes g + f \otimes Ag + 1\!\!1_{D(M)^c} \langle \nabla f, \nabla g \rangle_\Phi \tag{4}$$

with $A = 1/2n\, \Delta^M$ and the bilinear form $\langle .,. \rangle_{\Phi(x,y)} : T_x M \times T_y M \to \mathbb{R}$

$$\langle U, V \rangle_{\Phi(x,y)} := \frac{1}{n} \langle \Phi_1^{-1}(x,y) U, \Phi_2^{-1}(x,y) V \rangle_{\mathbb{R}^n}$$

whenever $F : M \times M \to \mathbb{R}$ is of the form $F = f \otimes g$ for smooth $f, g : M \to \mathbb{R}$. Moreover, in the case $F = f \otimes 1$ or $F = 1 \otimes g$ one finds $L_k f \otimes 1 \to Af \otimes 1$ and $L_k 1 \otimes g \to 1 \otimes Ag$ locally uniformly on $M \times M$ for k tending to infinity.

Proof. Suppose first that $(u,v) \in D(M)^c$ and let U be some neighborhood with $(u,v) \in U \subset D(M)^c$. Now for any $(u',v') \in U$ the Taylor expansion of $F = f \otimes g$ about (u',v') and the definition of the exponential map yield

$$\begin{aligned} f\big(\exp_{u'}(\tfrac{1}{\sqrt{k}} \Phi^1_{(u',v')}\theta)\big) g\big(\exp_{v'}(\tfrac{1}{\sqrt{k}} \Phi^2_{(u',v')}\theta)\big) &= f(u')g(v') \\ &+ \frac{1}{\sqrt{k}} f(u') \langle \nabla g(v'), \Phi^2_{(u',v')}\theta \rangle_{T_{v'}M} + \frac{1}{\sqrt{k}} g(v') \langle \nabla f(u'), \Phi^1_{(u',v')}\theta \rangle_{T_{u'}M} \\ &+ \frac{1}{k} \langle \nabla f(u'), \Phi^1_{(u',v')}\theta \rangle_{T_{u'}M} \cdot \langle \nabla g(v'), \Phi^2_{(u',v')}\theta \rangle_{T_{v'}M} \\ &+ \frac{1}{2k} f(u') \operatorname{Hess} g_{v'}(\Phi^2_{(u',v')}\theta, \Phi^2_{(u',v')}\theta) \\ &+ \frac{1}{2k} g(v') \operatorname{Hess} f_{u'}(\Phi^1_{(u',v')}\theta, \Phi^2_{(u',v')}\theta) + o_{u',v'}(\tfrac{1}{k}) \end{aligned}$$

where $o(.)$ is a "little o" Landau function. In fact, $o_{u',v'}(\frac{1}{k})$ can be replaced by some uniform $o_U(\frac{1}{k})$ due to the smoothness of the data (M,g) and the function F. Inserting this into the definition of L_k gives

$$
\begin{aligned}
L_k(F)(u',v') &= \Delta f(u')g(v') + f(u')\Delta g(v') \\
&+\frac{1}{n}\langle \Phi_1^{-1}(u',v')\nabla f(u'), \Phi_2^{-1}(u',v')\nabla g(v')\rangle_{\mathbb{R}^n} + \vartheta_U(\frac{1}{k})
\end{aligned}
\tag{5}
$$

with $\vartheta_U(\frac{1}{k}) \to 0$ for $k \to \infty$ because

$$
\begin{aligned}
\fint_{S^{n-1}} \langle \nabla f(u'), \Phi^1_{(u',v')}\theta\rangle_{T_{v'}M} d\theta &= \fint_{S^{n-1}} \langle \nabla g(v'), \Phi^2_{(u',v')}\theta\rangle_{T_{v'}M} d\theta = 0 \\
\frac{1}{2}\fint_{S^{n-1}} \operatorname{Hess} f_{u'}(\Phi^1_{(u',v')}\theta, \Phi^2_{(u',v')}\theta) d\theta &= \frac{1}{2n}\Delta f(u') \\
\frac{1}{2}\fint_{S^{n-1}} \operatorname{Hess} g_{v'}(\Phi^2_{(u',v')}\theta, \Phi^2_{(u',v')}\theta) &= \frac{1}{2n}\Delta g(v') \\
\fint_{S^{n-1}} \langle \nabla f(u'), \Phi^1_{(u',v')}\theta\rangle_{T_{u'}M} &\cdot \langle \nabla g(v'), \Phi^2_{(u',v')}\theta\rangle_{T_{v'}M} \\
&= \frac{1}{n}\langle \Phi_1^{-1}(u',v')\nabla f(u'), \Phi_2^{-1}(u',v')\nabla g(v')\rangle_{\mathbb{R}^n}.
\end{aligned}
$$

Now if $(u,v) \in D(M)$ by definition of L_k the coupling term $\langle \nabla f, \nabla g\rangle_\Phi$ does not appear and thus the claim is proved. □

The operator L_c has two irregular properties, one being its degeneracy, i.e. the second order part acts only in n of the $2n$ directions, and the other one being the discontinuity of the coefficients on $D(M) \cup \operatorname{Cut}(M)$. Both features together cause problems for the definition of a semigroup e^{tL_c} via the Hille-Yosida theorem. Therefore we confine ourselves to the construction of a solution Ξ to the martingale problem for L_c in a restricted sense by showing compactness of the laws of the sequence $(\tilde{\Xi}^{k,(x,y)}_{\cdot})_k$ on the space $D_{\mathbb{R}_+}(M \times M)$ of cadlag paths equipped with the Skorokhod topology.

Theorem 1 (Coupling Central Limit Theorem). *The sequence of the laws of* $\left(\tilde{\Xi}^{k,(x,y)}_{\cdot}\right)_{k\geq 0}$ *is tight on* $D_{\mathbb{R}_+}(M \times M)$ *and any weak limit of a converging subsequence* $(\tilde{\Xi}^{k',(x,y)}_{\cdot})_{k'}$ *is a solution to the martingale problem for* L_c *in the following restricted sense: let*

$$
(\Omega^\infty, P^\infty_{x,y}, (\tilde{\Xi}^{\infty,(x,y)}_s)_{s\geq 0}) = (D_{\mathbb{R}_+}(M \times M), \underset{k'\to\infty}{w\text{-}\lim}(\tilde{\Xi}^{k',(x,y)}_{\cdot})_* P, (\pi_s)_{s\geq 0})
$$

denote the canonical process on $M \times M$ *induced from a limit measure* $w\text{-}\lim_{k'\to\infty}(\tilde{\Xi}^{k',(x,y)}_{\cdot})_* P$ *on* $D_{\mathbb{R}_+}(M \times M)$ *and the natural coordinate projections* $\pi_s : D_{\mathbb{R}_+}(M \times M) \to M \times M$, *then for all* $F \in C^\infty_0(M \times M \setminus (D(M) \cup \operatorname{Cut}(M)))$, $F = f \otimes 1$ *or* $F = 1 \otimes g$ *with smooth* f, $g : M \to \mathbb{R}$ *the process*

$$
F(\Xi^{(x,y)}_t) - F(x,y) - \int_0^t L_c F(\Xi^{(x,y)}_s) ds
$$

is a $P^\infty_{x,y}$ martingale with initial value 0. *In particular, under $P^\infty_{x,y}$ both marginal processes $(\Xi^1_s := \pi^1_s)_{s\geq 0}$ and $(\Xi^2_s := \pi^2_s)_{s\geq 0}$ are Brownian motions on (M,g) starting in x and y respectively.*

Any probability measure on $D_{\mathbb{R}_+}(M \times M)$ with the properties above is called a solution to the (restricted) coupling martingale problem. We do not claim uniqueness nor a Markov property. Note that we circumvented the problem of the cut locus by the choice of admissible test functions F.

Proof of theorem 1. The proof is more or less standard. Concerning the tightness part we may engage the arguments of chapter 8 in [Dur96], for instance, with small modifications where we use the following tightness criterion for probability measures on the Sorokhod path space $D_{\mathbb{R}_+}(X)$ with $(X,d) = (M\times M, d)$ and $d((x_1,y_1),(x_2,y_2)) = \sqrt{d^2(x_1,x_2) + d^2(y_1,y_2)}$ (cf. thm. 15.5 in [Bil68] and thm. VI.1.5 in [JS87]).

Lemma 4 (Tightness criterion on $D_{\mathbb{R}_+}(X)$). *Let (X,d) be a complete and separable metric space and let $(\Omega_l, P_l, (\Xi^l_t)_{t\geq 0})_{l\in\mathbb{N}}$ be a sequence of cadlag processes on X. Then the following condition is sufficient for tightness of $(\Xi^l_\cdot)_*(P_l)$ on $D_{\mathbb{R}_+}(X)$: For all $N \in \mathbb{N}$, and η, $\epsilon > 0$ there are $x_0 \in X$, $l_0 \in \mathbb{N}$ and M, $\delta > 0$ such that*

i) $P_l(d(x_0, \Xi^l_0) > M) \leq \epsilon$ *for all* $l \geq l_0$

ii) $P_l(w(\Xi^l, \delta, N) \geq \eta) \leq \epsilon$ *for all* $l \geq l_0$

with $w(\Xi^l, \delta, N)(\omega) := \sup_{0\leq s,\, t\leq N,\, |s-t|\leq\delta} d(\Xi^l_s(\omega), \Xi^l_t(\omega))$.

We omit the details which can be found in [vR02] and turn to the martingale problem for L_c. Since L_k generates the process $(\tilde{\Xi}^{k,(x,y)}_\cdot)$ this is also true for its realization $(D_{\mathbb{R}_+}(M \times M), (\Xi^{k,(x,y)}_\cdot)_* P, (\pi_s)_{s\geq 0})$ on the path space and which is in this case equivalent to

$$F(\pi_t) - F(x,y) - \int_0^t L_k F(\pi_s) ds \text{ is a } (\Xi^{k,(x,y)}_\cdot)_* P\text{-Martingale} \tag{6}$$

for all $F \in \mathrm{Dom}(L_k) \supset C^3_0(M \times M)$. If $F \in C^\infty_0(M \times M \setminus (D(M) \cup \mathrm{Cut}(M)))$ by lemma 3 $\|L_k F - L_c F\|_\infty \to 0$ for $k \to \infty$ and thus by a general continuity argument (cf. lemma 5.5.1. in [EK86]) we may pass to the limit in the statement above provided k' is a subsequence such that $\text{w-}\lim_{k'\to\infty} (\tilde{\Xi}^{k',(x,y)}_\cdot)_* P$ exists. Finally, the assertion concerning the marginal processes follows either from [Jør75] or from putting $F = 1 \otimes f$ and $F = f \otimes 1$ in (6) respectively, in which case one may pass to the limit for k' tending to infinity without further restriction on the support of f. □

For the derivation of the coupling estimate it is easier to work with the continuous interpolated processes $(\hat{\Xi}^{k,(x,y)}_\cdot)_k$ as approximation of a suitable limit $\Xi^{(x,y)}_\cdot$. Therefore we need the following

Corollary 1. *The sequence of processes $\hat{\Xi}^{k,(x,y)}_\cdot$ is tight. For any subsequence k' the sequence of measures $(\hat{\Xi}^{k',(x,y)}_\cdot)_* P$ on $D_{\mathbb{R}_+}(M \times M)$ is weakly convergent if and only if*

$(\tilde{\Xi}^{k',(x,y)}_{\cdot})_* P$ *is, in which case the limits coincide. In particular the family* $((\hat{\Xi}^{k,(x,y)}_{\cdot})_* P)_k$ *is weakly precompact and any weak accumulation point is a solution of the (restricted) coupling martingale problem, which is supported by* $C_{\mathbb{R}_+}(M \times M)$.

Proof. By construction of $(\hat{\Xi}^{k,(x,y)}_{\cdot})$ and $(\tilde{\Xi}^{k,(x,y)}_{\cdot})$ we have $\hat{\Xi}^{k,(x,y)}_t = \tilde{\Xi}^{k,(x,y)}_{\frac{1}{k}\tau_k(s)}$ for all $t \geq 0$, $k \in \mathbb{N}$, i.e.

$$(\tilde{\Xi}^{k,(x,y)}_{\cdot}) = (\hat{\Xi}^{k,(x,y)}_{\cdot}) \circ \Theta_k$$

with the random time transformation $\Theta_k(s,\omega) = \frac{1}{k}\tau_k(s,\omega)$. Moreover, every process $(\hat{\Xi}^{k,(x,y)}_{\cdot})$ has continuous paths, so that

$$\operatorname{supp}\left(\underset{k'\to\infty}{\text{w-}\lim} (\hat{\Xi}^{k',(x,y)}_{\cdot})_* P \right) \subset C_{\mathbb{R}_+}(M \times M)$$

for every possible weak limit of a converging subsequence $(\hat{\Xi}^{k',(x,y)}_{\cdot})_* P$ and since the sequence Θ_k converges to $\mathrm{Id}_{\mathbb{R}_+}$ weakly, the continuity argument in section 17. of [Bil68] can be applied, giving

$$\underset{k'\to\infty}{\text{w-}\lim} (\tilde{\Xi}^{k',(x,y)}_{\cdot})_* P = \underset{k'\to\infty}{\text{w-}\lim} (\hat{\Xi}^{k',(x,y)}_{\cdot})_* P \tag{7}$$

for that specific subsequence, i.e. we have shown that for any subsequence $k' \to \infty$

$$\left((\hat{\Xi}^{k',(x,y)}_{\cdot}) \Rightarrow \nu \right) \Longrightarrow \left((\tilde{\Xi}^{k',(x,y)}_{\cdot}) \Rightarrow \nu \right).$$

To prove the other implication note first that the trivial estimate

$$d(\tilde{\Xi}^{k,(x,y)}_s, \tilde{\Xi}^{k,(x,y)}_{s-}) \leq \sqrt{\frac{1}{k}} \quad P\text{-a.s.}$$

implies the almost sure continuity of the coordinate process $\pi_{\cdot}$ w.r.t. any weak limit of $(\hat{\Xi}^{k,(x,y)}_{\cdot})_* P$. We may also write

$$(\tilde{\Xi}^{k,(x,y)}_{\cdot}) \circ \overline{\Theta}_k = (\hat{\Xi}^{k,(x,y)}_{\cdot}) \circ \Theta_k(\overline{\Theta}_k)$$

where $\overline{\Theta}_k(s,\omega) = \inf\{t \geq 0 \,|\, t > \Theta_k(s,\omega)\}$ is the (right continuous) generalized upper inverse of Θ_k which converges to $\mathrm{Id}_{\mathbb{R}_+}$ weakly, too. Since $\left\| \Theta_k(\overline{\Theta}_k) - \mathrm{Id}_{\mathbb{R}_+} \right\|_\infty \leq \frac{1}{k}$ and $(\hat{\Xi}^{k,(x,y)}_{\cdot})$ is a continuous process, it is easy to see that $(\hat{\Xi}^{k,(x,y)}_{\cdot})$ is tight on $C_{\mathbb{R}_+}(M \times M)$ if and only if $(\hat{\Xi}^{k,(x,y)}_{\cdot}) \circ \Theta_k(\overline{\Theta}_k)$ is tight on $D_{\mathbb{R}_+}(M \times M)$ and thus we may argue as before. Finally, the compactness itself comes from theorem 1 as well as as the characterization of any limit as a solution to the (restricted) coupling martingale problem via equation (7). $\square$

3.1. Coupling Probability and Gradient Estimates for Harmonic Functions. The curvature condition on the Riemannian manifold (M, g) enters our probabilistic proof of gradient estimates through the following lemma, in which we confine ourselves to the only nontrivial case of strictly negative lower (sectional) curvature bounds.

Lemma 5. *Let (M^n, g) be a smooth Riemannian manifold with* $\sec M \geq -\kappa < 0$ *and let* $x, y \in M$, $x \neq y$, *be joined by a unit speed geodesic* $\gamma = \gamma_{xy}$. *Then for any* $\xi, \eta \in S_x^{n-1} \subset T_x M$

$$
\begin{aligned}
d(\exp_x(t\xi), \exp_y(t /\!\!/_{\gamma}\eta)) \leq d(x,y) + t\langle \eta - \xi, \dot\gamma(0)\rangle_{T_xM} + o_\gamma(t^2) \\
+\frac{1}{2}t^2 \frac{\sqrt{\kappa}}{s_\kappa(d(x,y))}\big((|\xi^\perp|^2 + |\eta^\perp|^2)c_\kappa(d(x,y)) - 2\langle \eta^\perp, \xi^\perp\rangle_{T_xM}\big)
\end{aligned}
\tag{8}
$$

with $c_\kappa(t) = \cosh(\sqrt{\kappa}t)$, $s_\kappa(t) = \sinh(\sqrt{\kappa}t)$, *where* $/\!\!/_{\gamma}$ *denotes parallel translation on* (M, g) *along* γ *and* $\xi^\perp$, $\eta^\perp$ *denote the normal (w.r.t.* $\dot\gamma(0)$*) part of* ξ *and* η *respectively. In particular, if* $\xi^\perp = \eta^\perp$ *and* $\xi^{\parallel} = -\eta^{\parallel}$

$$
\begin{aligned}
d(\exp_x(t\xi), \exp_y(t /\!\!/_{\gamma}\eta)) \leq d(x,y) \\
- 2t\langle \xi, \dot\gamma(0)\rangle_{T_xM} + t^2\sqrt{\kappa}|\xi^\perp|^2 + o_\gamma(t^2).
\end{aligned}
\tag{9}
$$

Note that in the statement above the emphasis lies on the fact that the estimates (8) and (9) remain true also in $\mathrm{Cut}(M)$.

Proof. The proof can be found in any textbook on Riemannian geometry as long as y is not conjugate to x along γ and is based on the second variation formula for the arc length functional. In the case that y is conjugate to x along γ one may show (8) by subpartitioning $\gamma = \gamma_1 * \gamma_2 * \cdots * \gamma_k$ into geodesic segments $\{\gamma_i\}$ without conjugate points. By triangle inequality the individual estimates (8_i) along $\{\gamma_i\}$ can be reassembled to yield (8) along γ_{xy}. The proof also shows that the error term $o_\gamma(t^2)$ in (8) may be replaced by a uniform error estimate $o(t^2)$ as long as $x \neq y$ range over a compact $K \subset M \times M \setminus D(M)$. For the remaining few details the reader is referred to [vR02]. $\square$

In order to apply the previous lemma to the the sequence $(\hat{\Xi}^{l,(x,y)}_{.})_l$ we require in addition to condition $(*)$ on page 11 that $\Phi(.,.)$ satisfies

$$
\Phi_2(u,v) \circ \Phi_1^{-1}(u,v) = /\!\!/_{\gamma_{uv}} \text{ on } (T_uM)^{\perp_{\gamma_{xy}}} \tag{$**$}
$$

for all $(u, v) \in M \times M \setminus D(M)$ where γ_{uv} corresponds to some (w.r.t. u, $v \in M \times M \setminus D(M)$ symmetric) choice of connecting unit speed geodesics. A function $\Phi(.,.)$ satisfying $(*)$ and $(**)$ realizes the *coupling by reflection* method (cf. [Ken86, Cra91]) in our present context. In order to see that we actually may find at least one such map $\Phi(.,.)$ which is also measurable we may proceed similarly as in the proof of lemma 1: from a given measurable and symmetric choice $\gamma_{..} : M \times M \to C^1([0,1], M)$ and a continuous frame $\psi \in \Gamma(O(M))$ we obtain $\Phi_1(x, y)$ by an appropriate rotation of $\psi(x)$ such that $\Phi_1(x,y)e_1 = \dot\gamma_{xy}/\|\dot\gamma_{xy}\|$. $\Phi_2(x, y)$ is then obtained from $\Phi_1(x, y)$ by parallel transport and reflection w.r.t. the direction of γ_{xy}. Since these operations depend continuously (w.r.t. to the C^1-norm) on the curve γ_{xy}, $\Phi(.,.)$ inherits its measurability and continuity properties from the map $\gamma_{..} : M \times M \to C^1([0,1], M)$.

For $\delta > 0$ let us introduce the functional $T_{D,\delta} : C_{\mathbb{R}_+}(M \times M) \to \overline{\mathbb{R}}$

$$
T_{D,\delta}(\omega) = \inf\{s \geq 0 \,|\, d(\omega_s^1, \omega_s^2) \leq \delta\}
$$

with $\omega_s^i = \pi_i(\omega_s)$, $i = 1, 2$ being the projections of the path ω onto the factors. Then the *coupling time* $T_D = T_{D,0}$ is the first hitting time of the diagonal $D(M) \subset M \times M$.

Theorem 2 (Coupling Probability Estimate [Ken86]). *Let Φ be chosen as above and $Sec(M) \geq -\kappa < 0$. Then for arbitrary $x, y \in M$ and any weak limit $P_{x,y}^\infty = \underset{l' \to \infty}{w\text{-}\lim} (\hat{\Xi}_\cdot^{l',(x,y)})_* P$ on $C_{\mathbb{R}_+}(M \times M)$ the following estimate holds true:*

$$P_{x,y}^\infty(T_D = \infty) \leq \frac{n-1}{2} \sqrt{\kappa}\, d(x,y).$$

Proof. For $x = y$ there is obviously nothing to prove. So let $(x,y) \notin D(M)$ and assume first that M is compact. For $\delta > 0$ let $T_{D,\delta}^l : \Omega \to \mathbb{R} \cup \{\infty\}$

$$T_{D,\delta}^l(\omega) = \inf\{t \geq 0 \,|\, d(\hat{\Xi}_s^{l,(x,y)}(\omega)) \leq \delta\} = T_{D,\delta} \circ (\hat{\Xi}_\cdot^{l,(x,y)})(\omega)$$

be the first hitting time of the set $D_\delta = \{(x,y) \in M \times M \,|\, d(x,y) \leq \delta\}$ for the process $(\hat{\Xi}_s^{l,(x,y)})_{s \geq 0}$, where $(\Omega, \mathcal{O}, P)$ is the initial probability space on which the random i.i.d. sequences $(\xi)_{i \in \mathbb{N}}$ (and $(\eta_i)_{i \in \mathbb{N}}$) are defined. By the choice of Φ for $(u,v) \in D(M)^c$, $\theta \in S^{n-1} \subset \mathbb{R}^n$ and

$$(u^\epsilon, v^\epsilon) = (\exp_u(\epsilon \Phi_1 \theta), \exp_v(\epsilon \Phi_2 \theta))$$

we obtain from lemma 5

$$d(u^\epsilon, v^\epsilon) \leq d(u,v) - 2\epsilon\lambda + \sqrt{\kappa}\epsilon^2 \chi + o(\epsilon^2)$$

where $\lambda = \mathrm{pr}_1 \theta$ is the projection of θ onto the first coordinate axis and $\chi = \left\|\theta^\perp\right\|_{\mathbb{R}^n}^2$ is the squared length of the orthogonal part of θ. This estimate inserted into the inductive definition of $(\hat{\Xi}_t^{l,(x,y)})_{t \geq 0}$ yields in the case $\hat{\Xi}_{\lfloor lt \rfloor}^{1/\sqrt{l},(x,y)} \in D(M)^c$

$$\begin{aligned} d(\hat{\Xi}_t^{l,(x,y)}) = d(\hat{\Xi}_{lt}^{1/\sqrt{l},(x,y)}) &\leq d(\hat{\Xi}_{\lfloor lt \rfloor}^{1/\sqrt{l},(x,y)}) - 2\left(\frac{lt - \lfloor lt \rfloor}{\sqrt{l}}\right) \lambda_{\lfloor lt \rfloor + 1} \\ &+ \sqrt{\kappa} \left(\frac{lt - \lfloor lt \rfloor}{\sqrt{l}}\right)^2 \chi_{\lfloor lt \rfloor + 1} + o\left[\left(\frac{lt - \lfloor lt \rfloor}{\sqrt{l}}\right)^2\right] \end{aligned}$$

with the random variables $\lambda_i = \mathrm{pr}_1 \xi_i$ and $\chi_i = \left\|\xi_i^\perp\right\|_{\mathbb{R}^n}^2$, from which one deduces by iteration

$$\begin{aligned} &\leq d(x,y) - 2\frac{1}{\sqrt{l}} \sum_{i=0}^{\lfloor lt \rfloor} \lambda_{i+1} - 2\left(\frac{lt - \lfloor lt \rfloor}{\sqrt{l}}\right) \lambda_{\lfloor lt \rfloor + 1} + \sqrt{\kappa}\frac{1}{l} \sum_{i=0}^{\lfloor lt \rfloor} \chi_{i+1} \\ &\quad + \sqrt{\kappa} \left(\frac{lt - \lfloor lt \rfloor}{\sqrt{l}}\right)^2 \chi_{\lfloor lt \rfloor + 1} + \lfloor lt \rfloor o \left(\frac{lt - \lfloor lt \rfloor}{\sqrt{l}}\right)^2 \\ &= d(x,y) - 2S_t^l + \sqrt{\kappa}\frac{1}{l} \sum_{i=0}^{\lfloor lt \rfloor} \chi_{i+1} + \rho_t(l) =: r^l(t) \end{aligned} \tag{10}$$

at least on the set $\{T^l_{D,\delta} > t\}$, with

$$S^l_t := \frac{1}{\sqrt{l}} S_{lt},\ S_t := S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor) S_{\lfloor t \rfloor + 1},\ S_k := \sum_{i=0}^{k} \lambda_i$$

and

$$\rho_t(l) \to 0 \text{ for } l \to \infty.$$

Define furthermore the stopping times

$$T^l_\delta : \Omega \to \mathbb{R} \cup \{\infty\}, \quad T^l_\delta = \inf\{t \geq 0 \,|\, r^l_t \leq \delta\}$$

then the inequality above implies $\{T^l_{D,\delta} > m\} \subset \{T^l_\delta > m\}$ for all l, $m \in \mathbb{N}$ and hence

$$\mathbb{P}_{P,\Omega}(\{T^l_\delta > m\}) \geq \mathbb{P}_{P,\Omega}(\{T^l_{D,\delta} > m\}) = \int\limits_{\{T_{D,\delta} > m\}} (\hat{\Xi}^{l,(x,y)}_{\cdot})_*(P)(d\omega) \tag{11}$$

where the second integral is taken on a subset of the path space $\Omega' = C(\mathbb{R}_+, M \times M)$ with respect to the image measure of P under $(\hat{\Xi}^{l,(x,y)}_{\cdot})$. By assumption we have $P^\infty_{x,y} = \operatorname{w-lim}_{l' \to \infty} (\hat{\Xi}^{l',(x,y)}_{\cdot})_* P$ and the lower semi-continuity of the function $T_{D,\delta}$ w.r.t. to the topology of locally uniform convergence on the path space implies that the set $\{T_{D,\delta} > m\} \subset C(\mathbb{R}_+, M \times M)$ is open. Thus from (11) it follows that

$$\begin{aligned}
\mathbb{P}_{P^\infty_{x,y}}(\{T_{D,\delta} > m\}) &= \int\limits_{\{T_{D,\delta} > m\}} (\tilde{\Xi}^{\infty,(x,y)}_{\cdot})_*(P)(d\omega) \\
&\leq \liminf_{l' \to \infty} \int\limits_{\{T_{D,\delta} > m\}} (\hat{\Xi}^{l',(x,y)}_{\cdot})_*(P)(d\omega) \leq \liminf_{l' \to \infty} \mathbb{P}_{P,\Omega}(\{T^{l'}_\delta > m\}) \\
&= \mathbb{P}(\{T_\delta(r^\infty) > m\}).
\end{aligned}$$

The last equality is a consequence of Donsker's invariance principle applied to the sequence of processes $(r^l_{\cdot})_{l \in \mathbb{N}}$: since each $(\lambda_i)_i$ and $(\chi_i)_i$ are independent sequences of i.i.d. random variables on $\{\Omega, P, \mathcal{A}\}$ with

$$\mathbb{E}(\lambda_i) = 0, \quad \mathbb{E}(\lambda_i^2) = \frac{1}{n}, \quad \mathbb{E}(\chi_i) = \frac{n-1}{n}$$

one finds that $(r^l_{\cdot})_l$ converges weakly to the process $r^\infty_{\cdot}$ with

$$r^\infty_t = d(x,y) + \frac{2}{\sqrt{n}} b_t + \sqrt{\kappa} \frac{n-1}{n} t \tag{12}$$

such that in particular $\mathbb{P}_{P,\Omega}(\{T_\delta(r^\infty) = m\}) = 0$ and we can pass to the limit for $l \to \infty$ in the last term on the right hand side of (11). Letting m tend to infinity leads to

$$\mathbb{P}_{P^\infty_{x,y},\Omega'}(\{T_{D,\delta} = \infty\}) \leq \mathbb{P}_{P,\Omega}(\{T_\delta(r^\infty) = \infty\}),$$

where $\delta > 0$ was chosen arbitrarily from which we finally may conclude

$$\mathbb{P}_{P^{\infty}_{x,y},\Omega'}(\{T_D = \infty\}) \leq \mathbb{P}_{P,\Omega}(\{T_0(r^{\infty}) = \infty\}) \tag{13}$$

with T_0 being the first hitting time of the origin for the semi-martingale r^{∞}. Using a Girsanov transformation of $(\Omega, \mathcal{O}, P)$ the probability on the right hand side can be computed precisely to be

$$\mathbb{P}_{P,\Omega}(\{T_0(r^{\infty}) = \infty\}) = 1 - e^{-\frac{1}{2}\sqrt{\kappa}(n-1)d(x,y)} \leq \frac{n-1}{2}\sqrt{\kappa}\, d(x,y),$$

which is the claim in the compact case. For noncompact M we choose some open precompact $A \subset M$ such that (x, y) in K. We may stop the processes $(\hat{\Xi}^{l',(x,y)}_{\cdot})$ when they leave A and repeat the previous arguments for the the stopping time $T_{A,D,\delta} = T_{D,\delta} \wedge T_{A^c}$ with $T_{A^c} = \inf\{s \geq 0 \,|\, \omega_s \in A^c\}$ which gives instead of (11)

$$\mathbb{P}_{P,\Omega}(\{T^l_\delta \wedge T_{A^c} > n\}) \geq \int\limits_{\{T_{D,\delta} \wedge T_{A^c} > n\}} (\hat{\Xi}^{l,(x,y)}_{\cdot})_*(P)(d\omega)$$

From this we obtain (13) if we successively let tend $l \to \infty$, $A \to M \times M$, $\delta \to 0$ and $n \to \infty$. □

Also in the case of lower Ricci curvature bounds the same type of arguments should yield the extension of theorem 2. However, the difficulties arise from the fact that lower Ricci bounds lead to a uniform upper estimate of the expectation of χ_i in (10) only. Since these random variables are also only asymptotically mutually independent, one has to find and apply an appropriate central limit theorem to the expression $\frac{1}{l}\sum_{i=0}^{\lfloor lt \rfloor} \chi_{i+1}$ in order to obtain the pathwise(!) upper bound for the distance process by the semimartingale (12).

For different (local or global) versions of the following result as well as for extensions to harmonic maps the reader is referred in particular to the works by W. Kendall, M. Cranston or more recently by F.Y. Wang.

Corollary 2 (Gradient estimate [Cra91]). *If u is a harmonic, nonnegative and uniformly bounded function on M, then*

$$|u(x) - u(y)| \leq \|u\|_{\infty} \frac{\sqrt{\kappa}(n-1)}{2}\, d(x,y). \tag{14}$$

Proof. From elliptic regularity theory we now that $u \in C^{\infty}(M)$. Let $x \neq y$ be given. Since $\Delta u = 0$ we find $L_c(u \otimes 1) = L_c(1 \otimes u) = 0$ and from theorem 1 it follows that both processes $((u \otimes 1)(\pi_s))_s$ and $((1 \otimes u)(\pi_s))_s$ are nonnegative continuous bounded martingales with respect to the probability measure $P^{\infty}_{x,y}$, where $\pi_{\cdot} = (\pi_1, \pi_2)_{\cdot}$ is the projection process on the path space $C_{\mathbb{R}_+}(M \times M)$. For any $s > 0$ we obtain by means of the optional stopping theorem

$$\begin{aligned} u(x) - u(y) &= (u \otimes 1)(\pi_0) - (1 \otimes u)(\pi_0) \\ &= \mathbb{E}_{P^{\infty}_{x,y}}[(u \otimes 1)(\pi_{s \wedge T_D}) - (1 \otimes u)(\pi_{s \wedge T_D})] \end{aligned}$$

which equals, since $(u\otimes 1)(\pi_{T_D}) = (1\otimes u)(\pi_{T_D})$ on $\{T_D<\infty\}$

$$= \mathbb{E}_{P^\infty_{x,y}}\left[\left((u\otimes 1)(\pi_{s\wedge T_D}) - (1\otimes u)(\pi_{s\wedge T_D})\right)1\!\!1_{\{s<T_D\}}\right]$$

and finally, with $\sup_{x,y}|u(x)-u(y)|\le \|u\|_\infty$ following from $u\ge 0$

$$\le \|u\|_\infty\, \mathbb{P}_{P^\infty_{x,y}}(\{T_D\ge s\})$$

Passing to the limit for $s\to\infty$ proves the claim by theorem 2. □

3.2. Coupling by Parallel Transport and Heat Kernel Gradient Estimate. Instead of the coupling by reflection one may also consider *coupling by parallel transport* when conditions $(*)$ and $(**)$ on Φ are replaced by

$$\Phi_2(u,v)\circ\Phi_1^{-1}(u,v) = /\!\!/_{\gamma_{uv}} \text{ on } T_uM\,\forall\, u,v\in M\times M. \qquad (**_P)$$

Due to (8) this leads to the estimate

$$d(\pi^1_t,\pi^2_t)\le e^{\kappa(n-1)t}d(x,y)\quad P^\infty_{x,y}-a.s.$$

for any limiting measure $P^\infty_{x,y}$ on $C_{\mathbb{R}_+}(M\times M)$. Since any such $P^\infty_{x,y}$ is a coupling for (M,g)-Brownian motions starting in x and y this implies a gradient estimate for the heat semigroup $P_t=e^{\Delta t}$ on (M,g) of the form

$$|\nabla P_t f|(x)\le e^{\kappa(n-1)t}P_t|\nabla f|(x)$$

for all $f\in C^1_c(M)$ and $x\in M$, cf. [Wan97] as well as for applications.

4. Extension to Riemannian Polyhedra. Let X be an n-dimensional topological manifold equipped with a complete metric d. We call (X,d) an n-dimensional Riemannian polyhedron with lower curvature bound $\kappa\in\mathbb{R}$ if (X,d) is isometric to locally finite polyhedron $\bigcup_i P_i$ of convex closed patches $P_i\subset M_i^n$ $(i\in I)$ of n-dimensional Riemannian manifolds with uniform lower sectional curvature bound κ, where

i) the boundary $\partial P_i=\bigcup_j S_{ij}\subset M_i$ of each patch $P_i\subset M_i$ is the union of totally geodesic hypersurfaces S_{ij} in M_i

ii) each $S_{ij}\subset X$ is contained in the intersection of at most two $P_k\subset X$ and $S_{ij}\subset M_i$ where $S_{kl}\subset M_k$ are isometric whenever two adjacent patches $P_i\subset X$ and $P_k\subset X$ have a common face $S_{ij}\simeq S_{kl}\subset (X,d)$

iii) the sum of the dihedral angles for each face of codimension 2 is less or equal 2π.

Examples. The boundary ∂K of a convex Euclidean polyhedron $K\subset\mathbb{R}^n$ (with nonempty interior $\mathring{K}$) is a (n-1)-dimensional Riemannian polyhedron with lower curvature bound 0 in our sense. A simple example for the case $\kappa<0$ in two dimensions is the surface of revolution obtained from a concave function $f:[a,b]\to\mathbb{R}_+$, $f\in C^1[a,b]\cap C^2([a,c)\cup(c,b])$ with $c\in(a,b)$ and

$$f'(c)=0 \text{ and } f''/f=\begin{cases}-k_1^2 & \text{on } [a,c)\\ -k_2^2 & \text{on } (c,b].\end{cases}$$

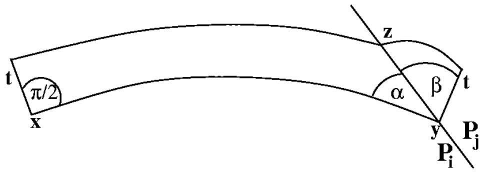
4.1. Constructions. The 2π-condition iii) above assures that (X,d) is an Alexandrov space with $\mathrm{Curv}(X) \geq \kappa$ (cf. [BBI01]), and we can use the result in [Pet98] that a geodesic segment connecting two arbitrary metrically regular points does not hit a metrically singular point, i.e. a point whose tangent cone is not the full Euclidean space. Moreover, from condition ii) it follows in particular that metrically singular points can occur only inside the $(n-2)$-skeleton X^{n-2} of X. Thus there is a natural parallel translation along any geodesic segment γ_{xy} whenever x and y are regular and which is obtained piecewisely from the parallel translation on the Riemannian patches P_i and from the natural gluing of the tangent half-spaces for points $x \in X^{n-1} \setminus X^{n-2}$ lying on the intersection of two adjacent $(n-1)$-faces $S_{ij} \simeq S_{kl} \subset X$. Similarly we can define the exponential map $\exp_x$ for every regular point $x \in X$, i.e. for given $\xi \in T_xX$ we obtain a unique 'quasi-geodesic' curve $\mathbb{R}_+ \ni t \to \exp_x(t\xi)$ (and which can be represented as a union of geodesic segments on the patches P_i). With these constructions at our disposal we can verify a non-smooth version of the asymptotic quadrangle estimate of lemma 5:

Theorem 3. *Let (X,d) be a n-dimensional Riemannian polyhedron with lower curvature bound $-\kappa < 0$ and let $x, y \in X \setminus X^{n-2}$ be connected by some segment γ_{xy}. Then for $\xi \in T_xX$, $\|\xi\| = 1$ the estimate (9) holds, where the error term $o(t^2)$ can be chosen uniform if $x \neq y$ range over a compact subset of $X \setminus X^{n-2}$.*

Proof. Let us prove (9) for fixed $x, y \in X \setminus X^{n-2}$ and $\xi \in T_xX$ first, i.e. without addressing the problem of uniformity. Suppose furthermore that for some P_i we have $\gamma_{xy} \subset P_i$, i.e. γ_{xy} is entirely contained in the (closed) patch P_i, then we distinguish three cases:

i) If $\gamma_{xy} \subset \mathring{P}_i$ then due to lemma 5 there is nothing left to prove.

ii) $x \in \mathring{P}_i$ and $y \in P_i \cap P_j$ for some j. Since y is assumed to be regular the first order part of estimate (9) is obviously true and we may focus on the second order part which corresponds to orthogonal variations of the geodesic γ, i.e. we may assume that $\xi = \xi^\perp$ in (9). If γ_{xy} is orthogonal to the hypersurface $\partial P_i \cap \partial P_j$ at y or x and y are both in $P_i \cap P_j$ then again there is nothing to prove since in this case we have to consider geodesic variations which take place completely on one of the patches $P_i \subset M_i$ or $P_j \subset M_j$ and we can apply lemma 5 on M_i or M_j respectively. Consequently we only have to treat the case that γ_{xy} is neither parallel nor orthogonal to $\partial P_i \cap \partial P_j$, i.e. $0 < \langle \frac{\dot\gamma_{xy}(d(x,y))}{\|\dot\gamma_{xy}(d(x,y))\|}, \nu \rangle_{T_yM_i} < 1$ where ν denotes the outward unit normal vector of ∂P_i.



Let $\eta = /\!\!/_{\gamma_{xy}} \xi$ be the parallel translate of a unit vector $\xi \in T_xM_i$ normal to $\dot\gamma_{xy}$. Then $\zeta = \eta - \frac{\langle \nu, \eta \rangle}{\langle \nu, \dot\gamma_{xy}(d(x,y)) \rangle} \dot\gamma_{xy}(d(x,y)) \in T_yM_i \cap T_yM_j$ is the unique vector in the intersection of the $\{\dot\gamma_{xy}(d(x,y)), \eta\}$-plane and the tangent hyperplane to ∂P_i in

y which is determined by its w.r.t. $\dot\gamma_{xy}$ orthogonal projection η. For its length we obtain $\|\zeta\| = \sin^{-1}\alpha$ where α is the angle enclosed by $\dot\gamma_{xy}$ and η at y. Since $\partial P_i \subset M_i$ and $\partial P_j \subset M_j$ are totally geodesic the point $z = \exp_y(t\zeta)$ also lies on $\partial P_i \cap \partial P_j$ and the triangle inequality yields

$$\begin{aligned} &d_X(\exp_x(t\xi), \exp_y(t /\!\!/_{\gamma_{xy}} \xi)) \\ &\qquad \leq d_{M_i}(\exp_x(t\xi), z) + d_{M_j}(z, \exp_y(t\eta)) \end{aligned} \tag{15}$$

where d_X, d_{M_i} and d_{M_j} denote the distance functions on X, M_i and M_j respectively. Now the estimate (8) of lemma 5 applied to ξ and ζ in M_i yields

$$d_{M_i}(\exp_x(t\xi), \exp_y(t\zeta)) \leq d_{M_i}(x,y) - t\frac{\cos\alpha}{\sin\alpha} + t^2\sqrt{k}|\xi^\perp|^2 + o_\gamma(t^2)$$

since trivially $\langle\dot\gamma, \zeta\rangle = \frac{\cos\alpha}{\sin\alpha}$ and by construction $\zeta^\perp = \eta = /\!\!/_{\gamma_{xy}}\xi$. As for the distance $d_{M_j}(z, \exp_y(t\eta))$ remember that by the smoothness assumption the curvature of M_j is locally uniformly bounded and from the Toponogov triangle comparison and the cosine formula on the model spaces $\mathbb{M}_{d,\kappa}$ we may infer with $\beta = \sphericalangle_{T_yM_j}(\zeta,\eta)$

$$d_{M_j}(z, \exp_y(t\eta)) = t\sqrt{|\zeta|^2 + 1 - |\zeta|\cos\beta} + o(t^2) = t\frac{\cos\alpha}{\sin\alpha} + o(t^2)$$

because all vectors $\dot\gamma_{xy}$, η and ζ lie on a common hyperplane and as $\eta \perp \dot\gamma$ we have $\alpha = \pi/2 - \beta$. Inserting the the last two inequalities into (15) yields (9).

iii) $x \in P_i \cap P_k$ and $y \in P_i \cap P_j$. We may argue similarly as in ii) by subdividing the quadruple into two geodesic triangles on M_j and M_k and a remaining quadruple on M_i. - Alternatively, if $z \in \gamma_{xy} \cap \mathring{P}_i \neq \emptyset$ then one may subdivide $\gamma_{xy} = \gamma_{xz} * \gamma_{zy}$ and argue as in ii). If $\gamma_{xy} \cap \mathring{P}_i = \emptyset$, then again we have to deal with variations of γ_{xy} on a single Riemannian patch P_k only, where P_k depends on the direction $\xi \perp \dot\gamma_{xy}$.

The discussion above proves (9) when $\gamma_{xy} \subset P_i$ for some P_i. In the general case when γ_{xy} is not contained in a single patch we subdivide $\gamma_{xy} = \gamma_{i_1} * \gamma_{i_2} * \cdots * \gamma_{i_m}$ into pieces $\gamma_{i_k} \subset P_{i_k}$ lying entirely on one of the patches which we consider separately: let $\{x_1, \ldots, x_m\} = \gamma_{xy} \cap X^{n-1}$ be the set of (transversal) intersections of γ_{xy} and X^{n-1} and for each $k = 1, \ldots, m$, $t > 0$ let $z_t^k = \exp_{x_k}(tz_k)$ where the direction z^k (depending on the initial direction $\xi \in T_xX$) is chosen as in ii). As before the triangle inequality yields the simple upper bound $d(x_t, y_t) \leq d(x_t, z_t^1) + d(z_t^1, z_t^2) + \cdots + d(z_t^{m-1}, z_t^m) + d(z_t^m, y_t)$ for the distance between $x_t = \exp_x(\xi_t)$ and $y_t = \exp_y(t /\!\!/_{\gamma_{xy}} \eta)$. On each patch P_{i_k} we may apply the previous discussion i) - iii) in order to derive asymptotic estimates for $d(x_t, z_t^1)$, $d(z_t^1, z_t^2)$, $\ldots$, $d(z_t^{m-1}, z_t^m)$ and $d(z_t^m, y_t)$, where it is important to note that for sufficiently small t the variations η_t^k of the pieces of γ_{i_k} which we construct on each patch P_{i_k} also lie entirely on P_{i_k}. (This follows from the fact that the segment γ_{xy} lies at a strictly positive distance away from X^{n-2} which comprises the set of points where more than just two patches intersect.) Hence we obtain an upper bound for the distance $d(x_t, y_t)$ in the global quadruple by a sum of distances $d(x_t, z_t^1)$, $d(z_t^1, z_t^2)$, $d(z_t^{m-1}, z_t^m)$ and $d(z_t^m, y_t)$ in geodesic quadrangles and triangles which are each entirely contained in a single patch. Analogously to the final step in ii) summing up the corresponding asymptotic upper estimates we recover (9) for the global quadruple due to the special choice of the directions $\{z^k | k \in 1, \ldots, m\}$. Finally, the uniformity assertion is obtained in a similar way by combining the arguments of lemma 5 on each patch P_i with the observation that for a given compact set $K \subset X \setminus X^{n-2}$ the collection of all segments $\{\gamma_{xy} | (x,y) \in K \times K\}$ also lies at a strictly positive distance away from X^{n-2}, which may be inferred from a simple compactness consideration. This implies that there is some $t_0 > 0$ such that for all $t \leq t_0$ and $x, y \in K$ all variations $\gamma_{xy,t}$ constructed in the previous paragraph determine a well-defined sequence of geodesic triangles and quadrangles located on the individual patches as above. Hence, by the smoothness of the patches (and the fact that only finitely many patches are involved for $x, y \in K \times K$) we may conclude as in lemma 5 that the estimate (9) is in fact locally uniform in the sense stated above. $\square$

As a second preparation for the probabilistic approach to a gradient estimate on (X,d) we need to state precisely what we understand by a Brownian motion in the present situation.

Definition 1. *The ('Dirichlet-')Laplacian Δ^X on (X,d) is defined as the generator of the Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ which is obtained as the $L^2(X, dm = \sum_i dm_{i\lfloor_{P_i}})$-closure of the classical energy form $\mathcal{E}(f,f) = \sum_i \int_{P_i} |\nabla f|^2 \, dm_i$ on the set of Lipschitz functions on (X,d) with compact support. A continuous Hunt process whose transition semigroup coincides with the semigroup associated to $(\mathcal{E}, D(\mathcal{E}))$ on $L^2(X, dm)$ is called a Brownian motion on (X,d).*

Equivalently we could define $(\mathcal{E}, D(\mathcal{E}))$ by the sum of the Dirichlet integrals on the patches P_i as above with the domain $D(\mathcal{E})$ equal to the set of piecewise $H^{1,2}(P_i)$-functions $f \in L^2(X, dm)$ with $\mathcal{E}(f,f) < \infty$ and whose traces $f_{|S_\pm}$ along the joint $(n-1)$-dimensional faces $S = P_i \cap P_j$ of each pair of adjacent patches coincide.

For the construction of the coupling process on $X \times X$ for two Brownian Motions on (X,d) we would like to proceed as in the smooth case by using a coupling map $\Phi(.,.)$ with $\Phi(x,y) : \mathbb{R}^n \times \mathbb{R}^n \to K_x X \times K_y X$ where K_x denotes the tangent cone of X over x (c.f. [BBI01]). Here further singularities of $\Phi(.,.)$ may be caused by the existence of non-Euclidean tangent cones K_x when $x \in X^{n-2}$. However, choosing beforehand a map $\Psi(,)$ on $X \times X^{n-2} \cup X^{n-2} \times X \cup \{(x,x) \,|\, x \in X\}$ with $\Psi(x,y) : \mathbb{R}^n \times \mathbb{R}^n \to K_x X \times K_y X$ (not necessarily isometric) and depending measurably on (x,y) we can find a globally defined measurable coupling map $\Phi(.,.)$ extending $\Psi(.)$ and satisfying $(*)$ and $(**)$ on $X \times X \setminus (X \times X^{n-2} \cup X^{n-2} \times X)$ which can be proved by slightly modifying the arguments of lemma 1. Hence we have everything we need to define a sequence of coupled (quasi-)geodesic random walks on $X \times X$ from which we obtain as before the sequences $(\tilde{\Xi}^{k,(x,y)}_{.})_k$ and $(\hat{\Xi}^{k,(x,y)}_{.})_k$ by scaling.

4.2. Coupling CLT on Polyhedra.

Proposition 1. *For any $(x,y) \in X \times X$ the sequences $(\tilde{\Xi}^{k,(x,y)}_{.})_k$ and $(\hat{\Xi}^{k,(x,y)}_{.})_k$ are tight on $D_{\mathbb{R}_+}(X \times X)$ and $C_0(\mathbb{R}_+, X \times X)$ respectively. For any subsequence k' the sequence of measures $(\hat{\Xi}^{k',(x,y)}_{.})_* P$ on $D_{\mathbb{R}_+}(X \times X)$ is weakly convergent if and only if*

$(\tilde{\Xi}^{k',(x,y)}_{.})_* P$ *is, in which case the limits coincide. For* $x, y \in X \setminus X^{n-2}$ *under any weak limit* $P^{\infty}_{x,y} = w\text{-}\lim_{k'\to\infty}(\hat{\Xi}^{k',(x,y)}_{.})_* P$ *the time changed marginal processes* $(\omega, t) \to \pi_1(\omega_{2n\cdot t})$ *and* $(\omega, t) \to \pi_2(\omega_{2n\cdot t})$ *are Brownian motions on* (X, d) *starting in* x *and* y *respectively.*

Proof. The tightness assertion and the coincidence of the limits of any jointly converging subsequences $(\tilde{\Xi}^{k',(x,y)}_{.})_{k'}$ and $(\hat{\Xi}^{k',(x,y)}_{.})_{k'}$ is proved precisely in the same manner as in the smooth case. Let us denote for short $\mu := P^{\infty}_{x,y} = \text{w-}\lim_{k'\to\infty}(\hat{\Xi}^{k',(x,y)}_{.})_* P$ the weak limit of some converging subsequence. Then it remains to identify the marginals $\mu^i = \Pi^i_* \mu$ of μ under the projection map $\Pi^i : C_{\mathbb{R}_+}(X \times X) \to C_{\mathbb{R}_+}(X)$, $\omega_. \to \omega^i_.$ as the measures induced from the Dirichlet form $(\mathcal{E}^{\tau}, D(\mathcal{E}^{\tau})) := (\frac{1}{2n}\mathcal{E}, D(\mathcal{E}))$ and starting points x and y respectively. For $\rho > 0$ let C_ρ be some Lipschitz $\rho-$neighbourhood of the set X^{n-2} and let $(\mathcal{E}^{\tau}_{\rho}, D(\mathcal{E}^{\tau}_{\rho}))$ be Dirichlet form which is obtained by taking the $L^2(X, dm)$-closure of the energy form $\mathcal{E}^{\tau}$ restricted to the set of Lipschitz functions with compact support in $X \setminus \overline{C_\rho}$. Furthermore let $T^i_\rho = \inf\{t \geq 0 | \omega^i_t \in \overline{C_\rho}\}$ be the hitting time of the marginals for the set C_ρ. We claim that the T^i_ρ-stopped marginal processes under any weak limit $P^{\infty}_{x,y}$ are associated with $(\mathcal{E}^{\tau}_{\rho}, D(\mathcal{E}^{\tau}_{\rho}))$ starting in x and y respectively. For this denote by $\mathcal{A}^{\rho}$ the collection of all $f \in \bigcap_i C^{\infty}(\mathring{P}_i) \cap \text{Lip}(X) \cap C_c(X \setminus \overline{C_\rho})$ satisfying the gluing condition for the normal derivatives on adjacent $(n-1)$-dimensional faces

$$\sum_{\partial P_i \cap \partial P_j \neq \emptyset} \frac{\partial}{\partial \nu_j} f = 0 \text{ on } \partial P_i \cap \partial P_j \cap \overline{C_\rho}. \tag{+}$$

Since for the generator L_k of $\tilde{\Xi}^{k,(x,y)}_{.}$ we have that

$$L_k(f \otimes 1)(u, v) = A_k f(u) \text{ and } L_k(1 \otimes g)(u, v) = A_k g(v)$$

with A_k acting as a mean value type operator according to formula (1) for $f \in \mathcal{A}^{\rho}$ it is easy to show (see also proof of lemma 6) that

$$f(\omega^i_t) - f(\omega^i_0) - \frac{1}{2n}\int_0^t \Delta^X_\rho f(\omega_s) ds \text{ is a } \mu-\text{martingale} \tag{16}$$

where Δ^X_ρ is the restriction of the formal Laplace-Beltrami-operator $\Delta^X = \sum_i \Delta^{P_i}_{|P_i}$ onto the set $\mathcal{A}^{\rho}$. Trivially in the formula above we may replace t by $t \wedge T^i_\rho$.

Let us assume first that X is compact and, by abuse of notation, let us denote by Δ^X_ρ also the generator of $(\mathcal{E}_\rho, D(\mathcal{E}_\rho))$, whose domain equals

$$D(\Delta^X_\rho) = \left\{ f \in L^2(X \setminus \overline{C_\rho}) \;\middle|\; \begin{array}{l} \Delta^X f \in L^2(X \setminus \overline{C}_\rho),\ f = 0 \text{ on } \partial C_\rho \\ f \text{ satisfies } (+) \end{array} \right\}.$$

Then by using the existence and smoothness of the heat kernel $(p^{X,\rho}_t(.,.))_{t\geq 0}$ associated to Δ_ρ^X we see that for $f \in D(\Delta_\rho^X)$ and $\epsilon > 0$

$$f_\epsilon(x) := p_\epsilon^{X,\rho} * f(x) = \int\limits_{X\setminus\overline{C_\rho}} p_\epsilon^{X,\rho}(x,y)f(y)m(dy) \in \mathcal{A}^\rho$$

and for $\epsilon \to 0$

$$\begin{aligned} f_\epsilon &\to f \in L^2(X \setminus \overline{C_\rho}) \\ \Delta_\rho^X f_\epsilon &= p_\epsilon * \Delta_\rho^X f \to \Delta_\rho^X f, \end{aligned}$$

where in the last line we used the fact that on $D(\Delta_\rho^X)$ the heat semigroup $(P_t^{X,\rho})$ commutes with Δ_ρ^X. This implies that $(\frac{1}{2n}\Delta_\rho^X, \mathcal{A}^\rho)$ is a core for the generator Δ_ρ^X of the Dirichlet form $(\mathcal{E}_\rho^\tau, D(\mathcal{E}_\rho^\tau))$ and hence by (16), abstractly speaking, we may say that the measures $\mu_\rho^i = (\Pi^i \circ \Sigma_{T_\rho^i})_* \mu$ induced from the projection Π^i and the stopping maps $\Sigma_{T_\rho^i} : C_{\mathbb{R}_+}(X \times X) \to C_{\mathbb{R}_+}(X \times X)$, $\omega_. \to \omega_{.\wedge T_\rho^i}$ $(i = 1, 2)$ are solutions to the martingale problem for $(\frac{1}{2n}\Delta_\rho^X, \delta_x)$ and $(\frac{1}{2n}\Delta_\rho^X, \delta_y)$ respectively. Since these martingale problems are well-posed the measures μ_ρ^i must coincide with those generated by the Dirichlet form $(\mathcal{E}_\rho^\tau, D(\mathcal{E}_\rho^\tau))$ and given starting points x and y.
In order to show that $\Pi_*^i\mu$ is induced from $(\mathcal{E}^\tau, D(\mathcal{E}^\tau))$ we want to pass to the limit for $\rho \to 0$ in the last statement. It is easy to see that weak convergence $(\Sigma_\rho)_*\mu \longrightarrow \mu$ will follow from $\lim_{\rho\to 0} \mu\{T_\rho^i \le t\} = 0$ for all $t > 0$. Moreover, obviously

$$\mu\{T_\rho^i \le t\} = \mu_\epsilon^i\{T_\rho \le t\} \tag{17}$$

for $\epsilon < \rho$, where μ_ϵ^i is defined as above and T_ρ is the first hitting time functional of C_ρ on the path space $C_{\mathbb{R}_+}(X)$. Remember that μ_ϵ^i is induced from the Dirichlet form $(\mathcal{E}_\epsilon^\tau, D(\mathcal{E}_\epsilon^\tau))$. Using the fact that the set X^{n-2} is polar for $(\mathcal{E}, D(\mathcal{E}))$ it follows that for any subsequence $\epsilon' \to 0$ the Dirichlet forms $(\mathcal{E}_{\epsilon'}^\tau, D(\mathcal{E}_{\epsilon'}^\tau))$ converges to $(\mathcal{E}^\tau, D(\mathcal{E}^\tau))$ in the sense of Mosco (cf. [Mos94]), which implies the $L^2(X, dm)-$strong convergence of the semigroups $P_t^{\epsilon'}$ to the semigroup P_t generated by $(\mathcal{E}^\tau, D(\mathcal{E}^\tau))$. Hence, by standard compactness arguments and the fact that $\{T_\rho \le t\} \subset C_{\mathbb{R}_+}(X)$ is closed we may pass to the limit for $\epsilon' \to 0$ on the right hand side of (17) which yields

$$\mu\{T_\rho^i \le t\} \le \nu\{T_\rho \le t\}$$

where ν is the measure associated to the form $(\mathcal{E}^\tau, D(\mathcal{E}^\tau))$ and starting point x or y respectively. Using once more that X^{n-2} is polar for $(\mathcal{E}, D(\mathcal{E}))$ we find that indeed

$$\limsup_{\rho\to 0} \mu\{T_\rho^i \le t\} \le \lim_{\rho\to 0} \nu\{T_\rho \le t\} = 0 \quad \forall\, t > 0. \tag{18}$$

By the continuity of the maps $\Pi^i : C_{\mathbb{R}_+}(X \times X) \to C_{\mathbb{R}_+}(X)$ the measures $\mu_\rho^i = \Pi_*^i(\Sigma_{T_\rho^i *}\mu)$ also converge to $\mu^i = \Pi_*^i\mu$ for $\rho \to 0$, and since the corresponding sequence of generating Dirichlet forms $(\mathcal{E}_\rho^\tau, D(\mathcal{E}_\rho^\tau))$ converges in the sense of Mosco to $(\mathcal{E}\tau, D(\mathcal{E}\tau))$ the limiting measures μ^i $(i = 1, 2)$ must be the unique measures on $C_{\mathbb{R}_+}(X)$ generated by $(\mathcal{E}\tau, D(\mathcal{E}^\tau))$

and the starting point x and y respectively.

In the case that X is non compact we have to localize the previous arguments: Fix some point $0 \in X$ and let $B_R(0) \subset X$ be the open metric ball around 0. Then, by the same reasoning as before we establish (16) for all functions $f \in \mathcal{A}_R^\rho := \{f \in \bigcap_i C^\infty((\mathring{P}_i \cap B_R(0)) \setminus \overline{C_\rho}) \cap \mathrm{Lip}((X \cap B_R(0)) \setminus \overline{C_\rho}) \mid f \text{ satisfies } (+)\}$. By the same arguments as above this implies

$$(\Sigma_{\rho,R})_*\mu = \nu_{\rho,R} = (\Sigma_{\rho,R})_*\nu,$$

where ν is the probability measure on the path space $C_{\mathbb{R}_+}(X)$ induced from the Dirchlet form $(\mathcal{E}^\tau, D(\mathcal{E}^\tau))$ and $\Sigma_{\rho,R}$ is the endomorphism on $C_{\mathbb{R}_+}(X)$ obtained from stopping a path when it leaves $(X \cap B_R(0)) \setminus \overline{C_\rho}$. analogous Dirichlet form on $X \cap B_R(0)) \setminus \overline{C_\rho})$. Using the polarity of X^{n-2} we may pass to the limit for $\rho \to 0$ first, which gives

$$(\Sigma_R)_*\mu = \nu_R = (\Sigma_R)_*\nu. \tag{19}$$

In a final step we would like to pass to the limit for $R \to \infty$. From the lower curvature bound on X it follows that

$$\lim_{R\to\infty} \nu\{T_{B_R(0)} \le t\} = 0 \quad \forall\, t > 0,$$

and which can be transferred to the analogous statement for ν by the same argument as in (18). Hence we may send $R \to \infty$ on both sides of (19) which concludes the proof of the proposition. ☐

Lemma 6. *Let $u \in L^\infty(X, dm) \cap \mathcal{D}(\mathcal{E})$ weakly harmonic on (X, d), i.e. $\mathcal{E}(u, \xi) = 0$ for all $\xi \in \mathcal{D}(\mathcal{E})$, and let $P^\infty_{x,y} = w\text{-}\lim_{k'\to\infty}(\hat{\Xi}^{k',(x,y)}_{\cdot})_*P$ be a weak limit of a subsequence $(\hat{\Xi}^{k',(x,y)}_{\cdot})_*P$ where $x, y \in X \setminus X^{n-2}$. Then under $P^\infty_{x,y}$ the processes $t \to u(\omega^1_t) = (u \otimes 1)(\omega_t)$ and $t \to u(\omega^2_t) = (1 \otimes u)(\omega_t)$ are martingales with respect to to the canonical filtration $(\overline{\mathcal{F}}_t = \sigma\{\pi^i_s \mid s \le t,\, i = 1, 2\})_{t\ge 0}$ on $C_{\mathbb{R}_+}(X \times X)$.*

Proof. Due to elliptic regularity theory one finds $u \in C^0(X) \cap C^\infty(\mathring{X}^{n-1})$ and that u satisfies the gluing condition (+). Hence we find that $A_k(u) \to 0$ locally uniformly on $X \setminus X^{n-2}$, where A_k is approximate Laplacian operator (1). We would like to use this property when we pass to the limit for $k' \to \infty$. It remains to justify this limit. Let us call for short $\nu = P^\infty_{x,y} = \mathrm{w}\text{-}\lim_{k'\to\infty}(\hat{\Xi}^{k',(x,y)}_{\cdot})_*P$ for a suitable subsequence k' and $\hat{\nu}^k = (\hat{\Xi}^{k,(x,y)}_{\cdot})_*P$. For $\rho > 0$ we may find some open neighbourhood $C_\rho \subset X$ of X^{n-2} satisfying

i) $\overline{B_{\rho/2}(X^{n-2})} \subset C_\rho \subset \overline{C_\rho} \subset B_\rho(X^{n-2})$
ii) ∂C_ρ intersects X^{n-1} transversally and
iii) $\partial C_\rho \cap X \setminus X^{n-1}$ is smooth.

Let $T^i_\rho = \inf\{t \ge 0 | \omega^1_t \in \overline{C_\rho}\}$ for $i = 1, 2$ the hitting time for the marginals of $\overline{C_\rho}$ and let

$$D_i = \{\omega \in C_{\mathbb{R}_+}(X \times X) \mid T^i_\rho \text{ is not continuous in } \omega\},$$

then $\nu(D_i)=0$ for $i=1,2$. This is seen as follows: since the hitting time of a closed set $C\subset X$ is lower semi-continuous on $C_{\mathbb{R}_+}(X)$ for each $\omega\in D_i$ we necessarily have $T^i_\rho(\omega)<\infty$ and it exists a sequence $\omega^\epsilon\to\omega^i\in C_{\mathbb{R}_+}(X)$ such that $T^i_\rho(\omega^i)+\delta<\liminf_\epsilon T^i_\rho(\omega^\epsilon)$ for some $\delta>0$. Note that by condition ii) on C_ρ the set $X^{n-1}\cap\partial C_\rho$ has (Hausdorff-)dimension $\leq n-2$ and hence is polar for Brownian motion on (X,d) and that by proposition 1 under ν the (time changed) marginal processes are Brownian motions on X. Thus $T^i_\rho(\omega)<\infty$ implies (ν-almost surely) $\omega^i_{T^i_\rho(\omega)}\in\partial C_\rho\cap X\setminus X^{n-1}$. But then $T^i_\rho(\omega^i)+\epsilon<\liminf_\epsilon T^i_\rho(\omega^\epsilon)$ implies the existence of some $\epsilon_0>0$ such that $\omega^i_{T^i_\rho(\omega)+\epsilon'}\notin C_\rho$ for all $\epsilon'\leq\epsilon_0$. Using the strong Markov property of the marginal processes under ν and the regularity of $\partial C_\rho\cap X\setminus X^{n-1}$ we finally deduce that the set of such paths has indeed vanishing ν-measure. On account of $\hat\nu^{k'}\Rightarrow\nu$ and the ν-almost sure continuity of functional $\Sigma_\rho: D_{\mathbb{R}_+}(X\times X)\to D_{\mathbb{R}_+}(X\times X)$, $(\Sigma_\rho\omega)(t)=\omega_{t\wedge T^1_\rho(\omega)\wedge T^2_\rho(\omega)}$ we find (by thm. 5.1. of [Bil68]) that $(\Sigma_\rho)_*\,\hat\nu^{k'}\Rightarrow\nu_\rho:=(\Sigma_\rho)_*\,\nu$ for $k'\to\infty$. Set $\overline{T}_\rho=T^1_\rho\wedge T^2_\rho$, then the Markov property of $\hat{\Xi}^{k,(x,y)}_{\cdot}$ and the optional sampling theorem yield that for all $t\geq s_l\geq\ldots s_1\geq 0$ and $v,g_1,\ldots g_l\in C_b(X\times X)$

$$
\begin{aligned}
&\left\langle v(\omega_t)-v(\omega_0)-\int_0^t\overline{A}_k v(\omega_s)ds,\, g_1(\omega_{s_1})\ldots g_l(\omega_{s_l})\right\rangle_{\hat\nu^k_\rho}\\
&=\left\langle v(\omega_{t\wedge\overline{T}_\rho})-v(\omega_0)-\int_0^{t\wedge\overline{T}_\rho}\overline{A}_k v(\omega_s)ds,\, g_1(\omega_{s_1\wedge\overline{T}_\rho})\ldots g_l(\omega_{s_l\wedge\overline{T}_\rho})\right\rangle_{\hat\nu^k}\\
&=\left\langle v(\omega_{s_l\wedge\overline{T}_\rho})-v(\omega_0)-\int_0^{s_l\wedge\overline{T}_\rho}\overline{A}_k v(\omega_s)ds,\, g_1(\omega_{s_1\wedge\overline{T}_\rho})\ldots g_l(\omega_{s_l\wedge\overline{T}_\rho})\right\rangle_{\hat\nu^k}\\
&=\left\langle v(\omega_{s_l})-v(\omega_0)-\int_0^{s_l}\overline{A}_k v(\omega_s)ds,\, g_1(\omega_{s_1})\ldots g_l(\omega_{s_l})\right\rangle_{\hat\nu^k_\rho}, \qquad (20)
\end{aligned}
$$

where $\overline{A}_k$ is the generator of $\hat{\Xi}^{k,(x,y)}_{\cdot}$. If we put $v=u\otimes 1$ for u as above and use the fact that

$$\overline{A}_k(u\otimes 1)=(A_k u)\otimes 1\to 0 \text{ uniformly on } X\setminus B_\rho(X^{n-2})$$

we see that for each $t\geq s_l$ the sequence of functionals $T^k_t: D_{\mathbb{R}_+}((X\setminus B_\rho(X^{n-2}))\times(X\setminus B_\rho(X^{n-2}))\to\mathbb{R}$ defined by

$$
\begin{aligned}
T^k_t(\omega)&=(u\otimes 1)(\omega_t)-(u\otimes 1)(\omega_0)\\
&\quad-\int_0^t\overline{A}_k(u\otimes 1)(\omega_s)ds\, g_1(\omega_{s_1})\ldots g_l(\omega_{s_l})
\end{aligned}
$$

converges uniformly on compacts $K\subset D_{\mathbb{R}_+}((X\setminus B_{\rho/2}(X^{n-2}))\times(X\setminus B_{\rho/2}(X^{n-2}))$ to the functional

$$T_t(\omega)=(u(\omega^1_t)-u(\omega^1_0))g_1(\omega_{s_1})\ldots g_l(\omega_{s_l}).$$

defined on $D_{\mathbb{R}_+}(X \times X) \supset D_{\mathbb{R}_+}((X \setminus B_{\rho/2}(X^{n-2})) \times (X \setminus B_{\rho/2}(X^{n-2}))$. Hence, due to $\hat{\nu}_\rho^{k'} \Rightarrow \nu_\rho$ we may pass to the limit in (20) for $k' \to \infty$ giving

$$\begin{aligned} &\left\langle (u(\omega_t^1) - u(\omega_0^1)) g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l}) \right\rangle_{\nu_\rho} \\ &\qquad = \left\langle (u(\omega_{s_l}^1) - u(\omega_0^1)) g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l}) \right\rangle_{\nu_\rho}. \end{aligned} \tag{21}$$

Moreover, by definition of ν_ρ (21) is equivalent to

$$\begin{aligned} &\left\langle (u(\omega_{t\wedge \overline{T}_\rho}^1) - u(\omega_0^1)) g_1(\omega_{s_1\wedge \overline{T}_\rho}) \ldots g_l(\omega_{s_l\wedge \overline{T}_\rho}) \right\rangle_{\nu} \\ &\qquad = \left\langle (u(\omega_{s_l\wedge \overline{T}_\rho}^1) - u(\omega_0^1)) g_1(\omega_{s_1\wedge \overline{T}_\rho}) \ldots g_l(\omega_{s_l\wedge \overline{T}_\rho}) \right\rangle_{\nu}. \end{aligned} \tag{22}$$

Finally, using again that under ν the marginal processes are time changed Brownian motions on X and X^{n-2} is polar we deduce $\overline{T}_\rho \geq T^1_{B_\rho(X^{n-2})} \wedge T^2_{B_\rho(X^{n-2})} \to \infty$ for $\rho \to 0$ ν-almost surely, such that taking the limit for $\rho \to 0$ in (22) yields

$$\begin{aligned} &\left\langle (u(\omega_t^1) - u(\omega_0^1)) g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l}) \right\rangle_{\nu} \\ &\qquad = \left\langle (u(\omega_{s_l}^1) - u(\omega_0^1)) g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l}) \right\rangle_{\nu} \end{aligned}$$

which amounts to the statement that the process $t \to u(\omega_t^1)$ is a $((\overline{\mathcal{F}}_t), \nu)$-martingale. □

Proposition 2. *Let (X, d) be an n-dimensional Riemannian polyhedron with with lower curvature bound $-\kappa < 0$, and for arbitrary $x, y \in X \setminus X^{n-2}$ let the measure $P^\infty_{x,y} = w\text{-}\lim_{l' \to \infty} (\hat{\Xi}^{l',(x,y)}_{\cdot})_* P$ on $C_{\mathbb{R}_+}(X \times X)$ be a weak limit of some suitably chosen subsequence k'. Then the coupling probability estimate holds true as in the smooth case, i.e.*

$$P^\infty_{x,y}(T_D = \infty) \leq \frac{n-1}{2} \sqrt{k}\, d(x, y).$$

Proof. Using proposition 3 we may proceed as in the proof of theorem 2 if we restrict of the discussion onto the set of paths stopped at time $\overline{T}_\rho$ for $\rho > 0$. In analogy to the proof of lemma 6 the final step is to send $\rho \to 0$ which yields the claim. □

Collecting the results we may conclude that Cranston's gradient estimate holds also in the case of Riemannian polyhedra.

Theorem 4. *Let (X, d) be an n-dimensional Riemannian polyhedron with lower curvature bound $-\kappa < 0$ then any nonnegative bounded function $u \in \mathcal{D}(\mathcal{E})$ which is weakly harmonic on (X, d) satisfies the gradient estimate (14).*

Proof. This follows from lemma 6 with optional sampling, using proposition 2 and the continuity of weakly harmonic functions. □

Just as in the smooth situation we can easily transfer the arguments to the coupling by parallel transport method ($**_P$) which yields a corresponding gradient estimate for the heat semigroup on (X, d).

Theorem 5. *Let (X, d) be an n-dimensional Riemannian polyhedron with lower curvature bound $-\kappa < 0$ and let $(P_t)_{t \geq 0}$ be the heat semigroup associated to the Dirchlet form $(\mathcal{E}, D(\mathcal{E}))$ then*

$$\mathrm{Lip}(P_t f) \leq e^{\kappa(n-1)t}\mathrm{Lip}(f) \quad \forall f \in \mathrm{Lip}(X, d).$$

References

[BBI01] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry.* American Mathematical Society, Providence, RI, 2001.

[Bil68] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., New York, 1968.

[Blu84] Gilles Blum. A note on the central limit theorem for geodesic random walks. *Bull. Austral. Math. Soc.*, 30(2):169–173, 1984.

[CH98] Jingyi Chen and Elton P. Hsu. Gradient estimates for harmonic functions on manifolds with Lipschitz metrics. *Canad. J. Math.*, 50(6):1163–1175, 1998.

[CKM93] Michael Cranston, Wilfrid S. Kendall, and Peter March. The radial part of Brownian motion. II. Its life and times on the cut locus. *Probab. Theory Related Fields*, 96(3):353–368, 1993.

[Cra91] Michael Cranston. Gradient estimates on manifolds using coupling. *J. Funct. Anal.*, 99(1):110–124, 1991.

[Dur96] Richard Durrett. *Stochastic calculus.* CRC Press, Boca Raton, FL, 1996.

[EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes.* John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[Jør75] Erik Jørgensen. The central limit problem for geodesic random walks. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 32:1–64, 1975.

[JS87] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes.* Springer-Verlag, Berlin, 1987.

[Ken86] Wilfrid S. Kendall. Nonnegative Ricci curvature and the Brownian coupling property. *Stochastics*, 19(1-2):111–129, 1986.

[Mos94] Umberto Mosco. Composite media and asymptotic Dirichlet forms. *J. Funct. Anal.*, 123(2):368–421, 1994.

[MS96] Paul Malliavin and Daniel W. Stroock. Short time behavior of the heat kernel and its logarithmic derivatives. *J. Differential Geom.*, 44(3):550–570, 1996.

[Pet98] Anton Petrunin. Parallel transportation for Alexandrov space with curvature bounded below. *Geom. Funct. Anal.*, 8(1):123–148, 1998.

[SV79] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes.* Springer-Verlag, Berlin, 1979. Die Grundlehren der mathematischen Wissenschaften, Band 233.

[vR02] Max-K. von Renesse. Comparison Properties of Diffusion Semigroups on Spaces with lower Curvature Bounds. 2002. PhD Thesis, Univ. Bonn.

[Wan94] Feng Yu Wang. Successful couplings of nondegenerate diffusion processes on compact manifolds. *Acta Math. Sinica*, 37(1):116–121, 1994.

[Wan97] Feng-Yu Wang. On estimation of the logarithmic Sobolev constant and gradient estimates of heat semigroups. *Probab. Theory Related Fields*, 108(1):87–101, 1997.

[Yau75] Shing Tung Yau. Harmonic functions on complete Riemannian manifolds. *Comm. Pure Appl. Math.*, 28:201–228, 1975.