

SUPPLEMENTARY MATERIALS: Landmark Alternating Diffusion*

Sing-Yuan Yeh[†], Hau-Tieng Wu[‡], Ronen Talmon[§], and Mao-Pei Tsui[¶]

SM1. Proof of the Main Theorem. With the preparation in Section 4, in this section, we will prove Theorem 4.6 and 4.10.

SM1.1. Some background material for the proof. We collect some lemmas that are needed for the proof in this subsection.

Lemma SM1.1 ([SM4] Lemma A.2). *In the normal coordinate around $x \in \mathcal{M}$, when $\|v\|_g \ll 1$, $v \in T_x \mathcal{M}$, the Riemannian measure satisfies*

$$dV(\exp_x v) = \left(1 - \frac{1}{6} \text{Ric}_{kl}(x) v^k v^l + O(\|v\|^3)\right) dv^1 \wedge dv^2 \dots \wedge dv^n,$$

where Ric is the Ricci curvature.

First, we evaluate the ambient distance of two points by the same metric.

Lemma SM1.2 ([SM4] Lemma A.2). *Fix $x \in \mathcal{M}$ and $y = \exp_x^{(2)}(v)$, where $v \in T_x \mathcal{M}$ with $\|v\|_{g^{(2)}} \ll 1$. We have*

$$\iota^{(2)}(y) = \iota^{(2)}(x) + d\iota^{(2)}(v) + Q_2^{(2)}(v) + Q_3^{(2)}(v) + O\left(\|v\|_{g^{(2)}}^4\right),$$

where $\Pi^{(2)}$ is the second fundamental form of $\iota^{(2)}$, $Q_2^{(2)}(v) := \frac{1}{2} \Pi^{(2)}(v, v)$ and $Q_3^{(2)}(v) := \frac{2}{6} \nabla_v^{(2)} \Pi^{(2)}(v, v)$. Further, for $z = \exp_x^{(2)}(u)$, where $u \in T_x \mathcal{M}$ with $\|u\|_{g^{(2)}} \ll 1$, we have

$$\left\| \iota^{(2)}(z) - \iota^{(2)}(y) \right\| = \|u - v\|_{g^{(2)}} + \tilde{Q}_2^{(2)}(u, v) + \tilde{Q}_3^{(2)}(u, v) + O\left(\|u\|_{g^{(2)}}^4, \|v\|_{g^{(2)}}^4\right),$$

where

$$\begin{aligned} \tilde{Q}_2^{(2)}(u, v) &= \frac{\left\| Q_2^{(2)}(u) - Q_2^{(2)}(v) \right\|^2}{2\|u - v\|_{g^{(2)}}}, \\ \tilde{Q}_3^{(2)}(u, v) &= \frac{\left\langle d\iota^{(2)}(u - v), Q_3^{(2)}(u) - Q_3^{(2)}(v) \right\rangle}{\|u - v\|_{g^{(2)}}}. \end{aligned}$$

*Supplementary material for SIMODS MS#M165668.

<https://doi.org/10.1137/24M1656682>

[†]Corresponding author. Data Science Degree Program, National Taiwan University and Academia Sinica, Taipei, 106, Taiwan (d10948003@ntu.edu.tw).

[‡]Courant Institute of Mathematical Sciences, New York University, New York, NY 10012 USA (hauwu@cims.nyu.edu).

[§]Viterbi Faculty of Electrical and Computer Engineering, Technion - Israel Institute of Technology, Haifa, 3200003 Israel (ronen@ef.technion.ac.il).

[¶]Department of Mathematics, National Taiwan University, Taipei, 106, Taiwan, and National Center for Theoretical Sciences, Mathematics Division, Taipei, 106, Taiwan (maopei@math.ntu.edu.tw).

Recall that we could find a diffeomorphism map $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $\Phi \circ \iota^{(2)} = \iota^{(1)}$. Its behavior is summarized below. Now, we evaluate the ambient distance by the diffeometric metric.

Lemma SM1.3 ([SM4] Lemma A.3). *Fix $x \in \mathcal{M}$. To simplify the notation, we ignore the subscription of $\nabla^{(2)}\Phi|_{\iota^{(2)}(x)}$ and use $\nabla\Phi$. Similar simplification holds for $\nabla^{(2)^2}\Phi, d\iota^{(2)}, \Pi^{(2)}$, etc. Suppose $y = \exp_x^{(2)} u$, where $u \in T_x\mathcal{M}$ and $\|u\|_{g^{(2)}}$ is small enough. Then we have*

$$\iota^{(1)}(y) = \iota^{(1)}(x) + \nabla\Phi d\iota^{(2)}u + Q_2^{(1)}(u) + Q_3^{(1)}(u) + O\left(\|u\|_{g^{(2)}}^4\right),$$

where $Q_2^{(1)}$ and $Q_3^{(1)}$ are quadratic and cubic polynomials respectively defined by $Q_2^{(1)}(u) := \frac{1}{2}\Pi^{(1)}(u, u)$ and $Q_3^{(1)}(u) := \frac{1}{6}\nabla\Phi\nabla_u^{(2)}\Pi^{(2)}(u, u) + \nabla^2\Phi(u, \Pi^{(2)}(u, u))$. Further, for $x \in \mathcal{M}$ defined as above, we have

$$\|\iota^{(1)}(x) - \iota^{(1)}(y)\| = \|\nabla\Phi d\iota^{(2)}(u)\| + \tilde{Q}_3^{(1)}(u) + O\left(\|u\|^4\right),$$

where

$$\tilde{Q}_3^{(1)}(u) := \frac{2\left\langle \nabla\Phi d\iota^{(2)}(u), Q_3^{(1)}(u) \right\rangle + \|Q_2^{(1)}(u)\|^2}{\|\nabla\Phi d\iota^{(2)}(u)\|}.$$

We also need the following truncation lemma.

Lemma SM1.4 ([SM4] Lemma A.5). *Suppose Assumption 4.1 holds, $F \in L^\infty(\mathcal{M})$ and $0 < \gamma < 1/2$. Then, when ϵ is small enough, for all pairs of $x, y \in \mathcal{M}$ so that $y = \exp_x^{(2)} v$ and $\|v\|_{g^{(2)}} > 2\epsilon^\gamma$, the following holds:*

$$\left| \int_{\mathcal{M}} K_\epsilon^{(1)}(x, z) K_\epsilon^{(2)}(z, y) F(z) dV^{(2)}(z) \right| = O\left(\epsilon^{d/2+3/2}\right).$$

SM1.2. Proof of Theorem 4.6. We start with the asymptotical behavior of the kernel associated with LAD. First, for convenience, denote

$$\tilde{B}_h^{(2)}(x) := \exp_x^{(2)}\left(B_h^{(2)}\right),$$

where $B_h^{(2)} = \{u \in T_x\mathcal{M} \mid \|u\|_{g^{(2)}} \leq h\} \subset T_x\mathcal{M}$ is a d -dim disk with the center 0 and the radius $h > 0$.

Lemma SM1.5. *Take $f \in C^3(\mathcal{M})$, $0 < \gamma < 1/2$ and $x, y \in \mathcal{M}$ so that $y = \exp_x^{(2)} v$, where $v \in T_x\mathcal{M}$ and $\|v\|_{g^{(2)}} \leq 2\epsilon^\gamma$. Fix normal coordinates around x associated with $g^{(1)}$ and $g^{(2)}$ and fix orthonormal coordinates associated with $g^{(1)}$ and $g^{(2)}$. Under these orthonormal coordinates, denote*

$$(SM1.1) \quad R_x := [d\exp_x^{(2)}|_0]^{-1} [d\iota^{(2)}]^{-1} \nabla\Phi [d\iota^{(1)}] [d\exp_x^{(1)}|_0],$$

which maps $\mathbb{R}^d$ to $\mathbb{R}^d$ associated with the basis $\{E_i\}_{i=1}^d$. Then, when ϵ is sufficiently small, the following holds:

$$\begin{aligned} & \int_{\mathcal{M}} K_{\epsilon}^{(1)}(x, z) K_{\epsilon}^{(2)}(z, y) F(z) dV^{(2)}(z) \\ &= \epsilon^{d/2} [F(x) A_{0,\epsilon}(v) + \epsilon^{1/2} A_{1,\epsilon}(F, v) + \epsilon A_{2,\epsilon}(F, v)] + O(\epsilon^{d/2+3/2}), \end{aligned}$$

where

$$\begin{aligned} A_{0,\epsilon}(v) &:= \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \tilde{K}^{(2)}(\|w - v/\sqrt{\epsilon}\|) dw, \\ A_{1,\epsilon}(F, v) &:= \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \tilde{K}^{(2)}(\|w - v/\sqrt{\epsilon}\|) \nabla_w^{(2)} F(x) dw, \\ A_{2,\epsilon}(F, v) &:= F(x) B_{21,\epsilon}(v) + F(x) B_{20,\epsilon}(v) + B_{22,\epsilon}(F, v), \end{aligned}$$

where

$$\begin{aligned} B_{21,\epsilon}(v) &:= \int_{\mathbb{R}^d} [\tilde{K}^{(1)}]'(\|R_x w\|) \tilde{K}^{(2)}(\|w - v/\sqrt{\epsilon}\|) \tilde{Q}_3^{(2)}(u) dw \\ &\quad + \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) [\tilde{K}^{(2)}]'(\|w - v/\sqrt{\epsilon}\|) \tilde{Q}_3^{(2)}(w, v/\sqrt{\epsilon}) dw, \\ B_{20,\epsilon}(v) &:= \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \tilde{K}^{(2)}(\|w - v/\sqrt{\epsilon}\|) \text{Ric}_{ij}^{(1)}(x) w^i w^j dw, \\ B_{22,\epsilon}(F, v) &:= \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \tilde{K}^{(2)}(\|w - v/\sqrt{\epsilon}\|) \frac{\nabla_{w,w}^{(2)} F(x)}{2} dw, \end{aligned}$$

where $\text{Ric}^{(\ell)}$ is the Ricci curvature and $\Pi^{(\ell)}$ is the second fundamental associated with $g^{(\ell)}$. Note that $B_{21,\epsilon}(v)$ depends on the first derivative of kernel functions, $B_{20,\epsilon}(v)$ depends on Ricci curvature, and $B_{22,\epsilon}(F, v)$ depends on the second derivative of F . Moreover, $A_{0,\epsilon}(v) = A_{0,\epsilon}(-v)$, $A_{1,\epsilon}(F, v) = -A_{1,\epsilon}(F, -v)$ and $A_{2,\epsilon}(F, v) = A_{2,\epsilon}(F, -v)$.

Proof. By the definition of kernels,

$$\begin{aligned} & \int_{\mathcal{M}} K_{\epsilon}^{(1)}(x, z) K_{\epsilon}^{(2)}(z, y) F(z) dV^{(2)}(z) \\ &= \int_{\mathcal{M}} \tilde{K}^{(1)}\left(\frac{\|\iota^{(1)}(x) - \iota^{(1)}(z)\|}{\sqrt{\epsilon}}\right) \tilde{K}^{(2)}\left(\frac{\|\iota^{(2)}(z) - \iota^{(2)}(y)\|}{\sqrt{\epsilon}}\right) F(z) dV^{(2)}(z). \end{aligned}$$

Suppose $z = \exp_x^{(2)}(u)$, where $u \in T_x \mathcal{M}$ with $\|u\|_{g^{(2)}} \ll 1$. By Lemma SM1.3,

$$\begin{aligned} & \tilde{K}^{(1)}\left(\frac{\|\iota^{(1)}(x) - \iota^{(1)}(z)\|}{\sqrt{\epsilon}}\right) \\ &= \tilde{K}^{(1)}\left(\frac{\|\nabla \Phi d\iota^{(2)}(u)\|}{\sqrt{\epsilon}}\right) + [\tilde{K}^{(1)}]' \left(\frac{\|\nabla \Phi d\iota^{(2)}(u)\|}{\sqrt{\epsilon}}\right) \frac{\tilde{Q}_3^{(1)}(u)}{\sqrt{\epsilon}} + O\left(\frac{\|u\|^5}{\epsilon}\right). \end{aligned}$$

Note that $\tilde{Q}_3^{(1)}(u)$ is an even function. On the other hand, by Lemma SM1.2,

$$\begin{aligned} & \tilde{K}^{(2)} \left(\frac{\|\iota^{(2)}(y) - \iota^{(2)}(x'')\|}{\sqrt{\epsilon}} \right) \\ &= \tilde{K}^{(2)} \left(\frac{\|\mathrm{d}\iota^{(2)}(u-v)\|}{\sqrt{\epsilon}} \right) + [\tilde{K}^{(2)}]' \left(\frac{\|\mathrm{d}\iota^{(2)}(u-v)\|}{\sqrt{\epsilon}} \right) \frac{\tilde{Q}_3^{(2)}(u,v)}{\sqrt{\epsilon}} + O \left(\frac{\|u\|^4}{\sqrt{\epsilon}}, \frac{\|v\|^4}{\sqrt{\epsilon}} \right). \end{aligned}$$

Note that $\tilde{Q}_3^{(2)}(-u, -v) = -\tilde{Q}_3^{(2)}(u, v)$. By the Lemma SM1.4, exist a $\gamma \in (0, 1)$ such that we could replace the integral domain $\tilde{B}_{\epsilon^\gamma}^{(2)}(x) \subset T_x\mathcal{M}$ by $T_x\mathcal{M}$ with order $O(\epsilon^{d/2+3/2})$ where ϵ is small enough. Now, it is sufficient to apply the Taylor expansion,

$$\begin{aligned} & \int_{\mathcal{M}} K_\epsilon^{(1)}(x, z) K_\epsilon^{(2)}(z, y) F(z) \mathrm{d}V^{(2)}(z) \\ &= \int_{\mathbb{R}^d} \left[\tilde{K}^{(1)} \left(\frac{\|R_x u\|}{\sqrt{\epsilon}} \right) + [\tilde{K}^{(1)}]' \left(\frac{\|R_x u\|}{\sqrt{\epsilon}} \right) \frac{\tilde{Q}_3^{(2)}(u)}{\sqrt{\epsilon}} + O \left(\frac{\|u\|^5}{\epsilon} \right) \right] \\ & \quad \times \left[\tilde{K}^{(1)} \left(\frac{\|u-v\|}{\sqrt{\epsilon}} \right) + [\tilde{K}^{(2)}]' \left(\frac{\|u-v\|}{\sqrt{\epsilon}} \right) \frac{\tilde{Q}_3^{(2)}(u, v)}{\sqrt{\epsilon}} + O \left(\frac{\|u\|^4}{\sqrt{\epsilon}}, \frac{\|v\|^4}{\sqrt{\epsilon}} \right) \right] \\ & \quad \times \left[F(x) + \nabla_u^{(1)} F(x) + \frac{\nabla_{u,u}^{(1)2} F(x)}{2} + O(\|u\|^3) \right] \\ & \quad \times \left[1 - \mathrm{Ric}_{ij}^{(1)}(x) u^i u^j + O(\|u\|^3) \right] \mathrm{d}u + O(\epsilon^{d/2+3/2}) \\ &= \epsilon^{d/2} [F(x) A_{0,\epsilon}(v) + \epsilon^{1/2} A_{1,\epsilon}(F, v) + \epsilon A_{2,\epsilon}(F, v) + O(\epsilon^{3/2})]. \end{aligned}$$

Let $w = u/\sqrt{\epsilon}$. By the change of variables, the result follows. ■

We remark that if the chosen kernels are both Gaussian, the α -landmark alternative kernel is Gaussian. Specifically, when $\tilde{K}^{(1)}$ and $\tilde{K}^{(2)}$ are both Gaussian, that is, $\tilde{K}^{(1)}(t) = \tilde{K}^{(2)}(t) = e^{-t^2}/\sqrt{\pi}$, we have

$$(SM1.2) \quad A_0(v) = \frac{\pi^{d/2}}{\sqrt{\det(I + \Lambda_x^2)}} e^{-\|(I + \Lambda_x^2)^{-1/2} \Lambda_x V_x^T v\|^2 / \epsilon},$$

which satisfies the exponential decay property of the kernel functions.

Lemma SM1.6. Fix $x \in \mathcal{M}$ and pick $F \in C^3(\mathcal{M})$. Fix normal coordinates around x associated with $g^{(1)}$ and $g^{(2)}$ and denote $\{E_i\}_{i=1}^d \subset T_x\mathcal{M}$ to be an orthonormal basis associated with $g^{(1)}$. Set R_x like (SM1.1) and by the SVD $R_x = U_x \Lambda_x V_x^T$, where $\Lambda_x = \mathrm{diag}[\lambda_1, \dots, \lambda_d]$. Then, when ϵ is sufficiently small, we have

$$\begin{aligned} & \int_{\mathcal{M}} \int_{\mathcal{M}} K_\epsilon^{(1)}(x, z) K_\epsilon^{(2)}(z, y) F(z) \mathrm{d}V^{(2)}(z) G(y) \mathrm{d}V^{(2)}(y) \\ &= \epsilon^d \left[\frac{G(x)F(x)}{\det(\Lambda_x)} + \epsilon C^F(G) \right] + O(\epsilon^{d+3/2}), \end{aligned}$$

where

$$C^F(G) := F(x)G(x)W(x) + C_1^F(G) + C_2^F(G),$$

where

$$W(x) = \frac{2\mu_{2,0}^{(1)}}{d \det(\Lambda_x)} \left(\sum_{i=1}^d \lambda_i \operatorname{Ric}_{ii}^{(2)}(x) \right) + \frac{\mu_{2,0}^{(2)}}{d \det(\Lambda_x)} s^{(2)}(x) \\ + \int_{\mathbb{R}^d} B_{21,\epsilon}(\sqrt{\epsilon}v) dv,$$

$$C_1^F(G) = \frac{\mu_{2,0}^{(1)}}{d \cdot \det(\Lambda_x)} \left(\sum_{i=1}^d \lambda_i \left[\nabla_{E_i}^{(2)} F(x) \nabla_{E_i}^{(2)} G(x) \right] \right),$$

and

$$C_2^F(G) = \frac{\mu_{2,0}^{(1)}}{2d \cdot \det(\Lambda_x)} F(x) \left(\sum_{i=1}^d \lambda_i \nabla_{E_i, E_i}^{(2)^2} G(x) \right) + \frac{\mu_{2,0}^{(2)}}{2d \cdot \det(\Lambda_x)} F(x) \Delta^{(2)} G(x) \\ + \frac{\mu_{2,0}^{(1)}}{2d \cdot \det(\Lambda_x)} G(x) \left(\sum_{i=1}^d \lambda_i \nabla_{E_i, E_i}^{(2)^2} F(x) \right).$$

where $\operatorname{Ric}^{(\ell)}$ is the Ricci curvature and $s^{(\ell)}$ is the scalar curvature associated with $g^{(\ell)}$. Moreover, $C_1^F(G)$ depends on the first derivative of F and G , $C_2^F(G)$ depends on the second derivative of F or G and W depends on the Ricci curvatures, scalar curvatures and the second fundamental form.

Proof. By Lemma SM1.4, let $y = \exp_x^{(2)}(v)$, where $v \in T_x \mathcal{M}$ and we directly compute

$$\int_{\mathcal{M}} \int_{\mathcal{M}} \tilde{K}_\epsilon^{(1)}(x, z) \tilde{K}_\epsilon^{(2)}(z, y) F(z) dV^{(2)}(z) G(y) dV^{(2)}(y) \\ = \epsilon^{d/2} \int_{\mathbb{R}^d} \left[F(x) A_{0,\epsilon}(v) + \epsilon^{1/2} A_{1,\epsilon}(F, v) + \epsilon A_{2,\epsilon}(F, v) + O(\epsilon^{3/2}) \right] \\ \times \left[G(x) + \nabla_v^{(2)} G(x) + \frac{\nabla_{v,v}^{(2)^2} G(x)}{2} + O(\|v\|^3) \right] \\ \times \left[1 - \operatorname{Ric}_{ij}^{(2)}(x) v^i v^j + O(\|v\|^3) \right] dv + O(\epsilon^{d+3/2}) \\ := \epsilon^{d/2} [I_0 + \epsilon^{1/2} I_1 + I_2 + \epsilon I_3 + I_4] + O(\epsilon^{d+3/2}),$$

where

$$I_0 = F(x)G(x) \int_{\mathbb{R}^d} A_{0,\epsilon}(v) dv, \quad I_1 = \int_{\mathbb{R}^d} A_{1,\epsilon}(F, v) \nabla_v^{(2)} G(x) dv, \\ I_2 = F(x) \int_{\mathbb{R}^d} A_{0,\epsilon}(v) \frac{\nabla_{v,v}^{(2)^2} G(x)}{2} dv, \quad I_3 = \int_{\mathbb{R}^d} A_{2,\epsilon}(F, v) G(x) dv,$$

$$I_4 = F(x)G(x) \int_{\mathbb{R}^d} A_{0,\epsilon}(v) \operatorname{Ric}_{ij}^{(2)}(x) v^i v^j dv.$$

We compute the right-hand side term by term and apply the Equation (A.60) in [SM4]. For I_0 , let $s = w - v/\sqrt{\epsilon}$ and we obtain

$$\begin{aligned} F(x)G(x) \int_{\mathbb{R}^d} A_{0,\epsilon}(v) dv &= \epsilon^{d/2} F(x)G(x) \int_{\mathbb{R}^d} K^{(1)}(\|R_x w\|) \left[\int_{\mathbb{R}^d} K^{(2)}(\|w - v/\sqrt{\epsilon}\|) dv \right] dw \\ &= \epsilon^{d/2} \frac{F(x)G(x)}{\det(\Lambda_x)}. \end{aligned}$$

For I_1 , set $s = w - v/\sqrt{\epsilon}$ and $u = R_x w$, and obtain

$$\begin{aligned} &\int_{\mathbb{R}^d} A_{1,\epsilon}(F, v) \nabla_v^{(2)} G(x) dv \\ &= \epsilon^{d/2+1/2} \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \left[\int_{\mathbb{R}^d} \tilde{K}^{(2)}(\|w - v/\sqrt{\epsilon}\|) \nabla_w^{(2)} F(x) \nabla_v^{(2)} G(x) dv \right] dw \\ &= \epsilon^{d/2+1/2} \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \left[\int_{\mathbb{R}^d} \tilde{K}^{(2)}(\|s\|) (w - s)^\top \nabla^{(2)} G(x) ds \right] w^\top \nabla^{(2)} F(x) dw \\ &= \epsilon^{d/2+1/2} \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) w^\top \nabla^{(2)} G(x) w^\top \nabla^{(2)} F(x) dw \\ &= \epsilon^{d/2+1/2} \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \nabla^{(2)} G(x)^\top (w w^\top) \nabla^{(2)} F(x) dw \\ &= \frac{\epsilon^{d/2+1/2} \mu_{2,0}^{(1)}}{d \det(\Lambda_x)} \left(\sum_{i=1}^d \lambda_i \left[\nabla_{E_i}^{(2)} F(x) \nabla_{E_i}^{(2)} G(x) \right] \right), \end{aligned}$$

where the third equality holds since the symmetric property. For I_2 , let $s = w - v/\sqrt{\epsilon}$ and $u = R_x w$, and we have

$$\begin{aligned} &\int_{\mathbb{R}^d} A_{0,\epsilon}(v) \frac{\nabla_{r,r}^{(2)2} G(x)}{2} dv \\ &= \epsilon^{d/2+1} F(x) \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \left[\int_{\mathbb{R}^d} \tilde{K}^{(2)}(\|w - v/\sqrt{\epsilon}\|) \frac{\nabla_{v,v}^{(2)2} G(x)}{2} dv \right] dw \\ &= \epsilon^{d/2+1} \frac{F(x)}{2} \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \left[\int_{\mathbb{R}^d} \tilde{K}^{(2)}(\|s\|) (w - s)^\top \nabla^{(2)2} G(x) (w - s) ds \right] dw \\ &= \epsilon^{d/2+1} \frac{F(x)}{2} \left[\int_{\mathbb{R}^d} \tilde{K}^{(2)}(\|s\|) ds \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) w^\top \nabla^{(2)2} G(x) w dw \right. \\ &\quad \left. + \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) dw \int_{\mathbb{R}^d} \tilde{K}^{(2)}(\|s\|) s^\top \nabla^{(2)2} G(x) s ds \right]. \end{aligned}$$

By the similar argument in I_1 , we have

$$I_2 = \epsilon^{d/2+1} \left[\frac{\mu_{2,0}^{(1)}}{2d \det(\Lambda_x)} F(x) \left(\sum_{i=1}^d \lambda_i \nabla_{E_i, E_i}^{(2)} G(x) \right) + \frac{\mu_{2,0}^{(2)}}{2d \det(\Lambda_x)} F(x) \Delta^{(2)} G(x) \right].$$

For I_3 , recall that $A_{2,\epsilon}(F, v) = F(x)B_{21,\epsilon}(v) + F(x)B_{20,\epsilon}(v) + B_{22,\epsilon}(F, v)$, and hence

$$\int_{\mathbb{R}^d} A_{2,\epsilon}(F, v)G(x)dv := I_5 + I_6 + I_7,$$

where

$$\begin{aligned} I_5 &= G(x)F(x) \int_{\mathbb{R}^d} B_{21,\epsilon}(v)dv, & I_6 &= G(x)F(x) \int_{\mathbb{R}^d} B_{20,\epsilon}(v)dv \\ I_7 &= G(x) \int_{\mathbb{R}^d} B_{22,\epsilon}(F, v)dv. \end{aligned}$$

The expansion of I_5 , I_6 , and I_7 is by the same direct expansion. For I_5 ,

$$\begin{aligned} G(x)F(x) \int_{\mathbb{R}^d} B_{21,\epsilon}(v)dv &= \epsilon^{d/2}G(x)F(x) \int_{\mathbb{R}^d} B_{21,\epsilon}(\sqrt{\epsilon}v)dv \\ &:= \epsilon^{d/2}G(x)F(x)W_1(x), \end{aligned}$$

where W_1 depends on second fundamental form. For I_7 , let $s = w - v/\sqrt{\epsilon}$ and $u = R_x w$, and we have

$$\begin{aligned} &G(x) \int_{\mathbb{R}^d} B_{22,\epsilon}(F, v)dv \\ &= \epsilon^{d/2+1}G(x) \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) \left[\int_{\mathbb{R}^d} \tilde{K}^{(2)}(\|w - v/\sqrt{\epsilon}\|) \frac{\nabla_{w,w}^{(2)} F(x)}{2} dv \right] dw \\ &= \epsilon^{d/2+1} \frac{G(x)}{2} \int_{\mathbb{R}^d} \tilde{K}^{(2)}(\|s\|) ds \int_{\mathbb{R}^d} \tilde{K}^{(1)}(\|R_x w\|) w^\top \nabla^{(2)} F(x) w dw \\ &= \frac{\epsilon^{d/2+1} \mu_{2,0}^{(1)}}{2d \det(\Lambda_x)} G(x) \left(\sum_{i=1}^d \lambda_i \nabla_{E_i, E_i}^{(2)} F(x) \right). \end{aligned}$$

Finally, let $s = w - v/\sqrt{\epsilon}$, and $u = R_x w$

$$\begin{aligned} I_4 + \epsilon I_6 &= F(x)G(x) \left[\int_{\mathbb{R}^d} A_{0,\epsilon}(v) \text{Ric}_{ij}^{(2)}(x) v^i v^j dv + \int_{\mathbb{R}^d} B_{20}(v)dv \right] \\ &= \frac{\epsilon^{d/2+1} \mu_{2,0}^{(1)}}{d \det(\Lambda_x)} G(x)F(x) \left(\sum_{i=1}^d \lambda_i \text{Ric}_{ii}^{(2)}(x) \right) + \frac{\epsilon^{d/2+1} \mu_{2,0}^{(2)}}{d \det(\Lambda_x)} G(x)F(x) s^{(2)}(x) \\ &\quad + \frac{\epsilon^{d/2+1} \mu_{2,0}^{(1)}}{d \det(\Lambda_x)} G(x)F(x) \left(\sum_{i=1}^d \lambda_i \text{Ric}_{ii}^{(2)}(x) \right) := \epsilon^{d/2+1} G(x)F(x)W_2(x), \end{aligned}$$

where the term I_6 is similar to the argument in I_2 and $s^{(\ell)}(x) = \sum_{i=1}^d \text{Ric}_{ii}^{(\ell)}(x)$ is scalar curvature. By putting everything together and setting

$$W(x) := W_1(x) + W_2(x) = \frac{2\mu_{2,0}^{(1)}}{d \det(\Lambda_x)} \left(\sum_{i=1}^d \lambda_i \text{Ric}_{ii}^{(2)}(x) \right) + \frac{\mu_{2,0}^{(2)}}{d \det(\Lambda_x)} s^{(2)}(x)$$

$$+ \int_{\mathbb{R}^d} B_{21,\epsilon}(\sqrt{\epsilon}v)dv,$$

the desired result follows. ■

With the above lemma, we can prove Theorem 4.6. We restate Theorem 4.6 here for convenience.

Theorem SM1.7. *Fix $x \in \mathcal{M}$. Fix normal coordinates around x associated with $g^{(\ell)}$ and the associated exponential maps $\exp^{(\ell)}$, where $\ell = 1, 2$. Let $\mathcal{B} := \{E_i\}_{i=1}^d \subset T_x\mathcal{M}$ be an orthonormal basis associated with $g^{(2)}$. Set*

$$R_x = \left[d \exp^{(1)} \Big|_0 \right]^{-1} [d\iota^{(1)}]^{-1} \nabla \Phi [d\iota^{(2)}] \left[d \exp^{(2)} \Big|_0 \right] \in \mathbb{R}^{d \times d},$$

with all quantities represented with the basis $\mathcal{B}$. By the singular value decomposition (SVD), we have $R_x = U_x \Lambda_x V_x^T$, where $\Lambda_x = \text{diag}[\lambda_1, \dots, \lambda_d]$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$, and $U_x, V_x \in O(d)$. Then, for a $f \in C^3(\mathcal{M})$, when $\epsilon > 0$ is sufficiently small, we have

$$\begin{aligned} T_{\text{lan},\epsilon,\alpha} f(x) &= f(x) + \frac{\epsilon \mu_{2,0}^{(2)}}{2d} \Delta^{(2)} f(x) + \frac{\epsilon \mu_{2,0}^{(1)}}{2d} \sum_{i=1}^d \lambda_i \nabla_{E_i E_i}^{(2)^2} f(x) \\ &\quad + \frac{\epsilon \mu_{2,0}^{(1)}}{d} \sum_{i=1}^d \lambda_i \left(\frac{\nabla_{E_i}^{(2)} p^{(2)}(x)}{p^{(2)}(x)} + \frac{\nabla_{E_i}^{(2)} q_\alpha(x)}{q_\alpha(x)} \right) \nabla_{E_i}^{(2)} f(x) \\ &\quad + \frac{\epsilon \mu_{2,0}^{(2)}}{d} \frac{\nabla^{(2)} p^{(2)}(x) \cdot \nabla^{(2)} f(x)}{p^{(2)}(x)} + \mathcal{O}(\epsilon^{3/2}), \end{aligned}$$

where $\nabla^{(\ell)}$ is the covariant derivative associated with $\iota^{(\ell)}$, $q_\alpha(x) := \frac{p_{\mathcal{Z}}^{(2)}(x)^{1-\alpha}}{p^{(2)}(x)^\alpha}$ and the implied constant depends on the scalar curvature at x , second fundamental form at x , C^3 norm of f and C^2 norms of $p^{(2)}$ and $p_{\mathcal{Z}}^{(2)}$.

Proof. By definition, we have,

$$\begin{aligned} &T_{\text{lan},\epsilon,\alpha} f(x) \\ &= \frac{\int_{\mathcal{M}} K_{\text{lan},\epsilon,\alpha}(x,y) f(y) p^{(2)}(y) dV(y)}{\int_{\mathcal{M}} K_{\text{lan},\epsilon,\alpha}(x,y) p^{(2)}(y) dV(y)} \\ &= \frac{\int_{\mathcal{M}} \left(\int_{\mathcal{M}} K_\epsilon^{(1)}(x,z) \frac{1}{d_\epsilon^{(2)}(z)^\alpha} K_\epsilon^{(2)}(z,y) p_{\mathcal{Z}}^{(2)}(z) dV(z) \right) f(y) p^{(2)}(y) dV(y)}{\int_{\mathcal{M}} \left(\int_{\mathcal{M}} K_\epsilon^{(1)}(x,z) \frac{1}{d_\epsilon^{(2)}(z)^\alpha} K_\epsilon^{(2)}(z,y) p_{\mathcal{Z}}^{(2)}(z) dV(z) \right) p^{(2)}(y) dV(y)}. \end{aligned}$$

The numerator of the left-hand side can be expanded and organized as

$$\begin{aligned} &\int_{\mathcal{M}} \int_{\mathcal{M}} K_\epsilon^{(1)}(x,z) K_\epsilon^{(2)}(z,y) \frac{p_{\mathcal{Z}}^{(2)}(z)}{d_\epsilon^{(2)}(z)^\alpha} dV^{(2)}(z) f(y) p^{(2)}(y) dV^{(2)}(y) \\ &= \epsilon^{-d\alpha} \int_{\mathcal{M}} \int_{\mathcal{M}} K_\epsilon^{(1)}(x,z) K_\epsilon^{(2)}(z,y) p^{(2)}(z)^{-\alpha} p_{\mathcal{Z}}^{(2)}(z)^{1-\alpha} \\ &\quad \times [1 - \alpha \epsilon \bar{E}(z) + \mathcal{O}(\epsilon^2)] dV^{(2)}(z) f(y) p^{(2)}(y) dV^{(2)}(y). \end{aligned}$$

To simplify the notation, denote $q_\alpha(x) := p^{(2)}(x)^{-\alpha} p_{\mathcal{Z}}^{(2)}(x)^{1-\alpha}$. By Lemma SM1.6, the above equation can be expanded and organized as (SM1.3)

$$\begin{aligned}
& \epsilon^{-d(1-\alpha)} \int_{\mathcal{M}} \int_{\mathcal{M}} K_\epsilon^{(1)}(x, z) K_\epsilon^{(2)}(z, y) q_\alpha(z) [1 - \epsilon \alpha \bar{E}(z)] dV^{(2)}(z) f(y) p^{(2)}(y) dV^{(2)}(y) \\
&= \frac{f(x) p^{(2)}(x) q_\alpha(x)}{\det(\Lambda_x)} + \epsilon p^{(2)}(x) q_\alpha(x) W(x) f(x) - \epsilon \alpha q_\alpha(x) p^{(2)}(x) f(x) \bar{E}(x) \\
&+ \frac{\epsilon \mu_{2,0}^{(1)}}{2d \det(\Lambda_x)} \left[q_\alpha(x) \sum_{i=1}^d \lambda_i \nabla_{E_i E_i}^{(2)^2} p^{(2)}(x) + p^{(2)}(x) \sum_{i=1}^d \lambda_i \nabla_{E_i E_i}^{(2)^2} q_\alpha(x) \right. \\
&\quad \left. + 2 \left(\sum_{i=1}^d \lambda_i \left[\nabla_{E_i}^{(2)} p^{(2)}(x) \nabla_{E_i}^{(2)} q_\alpha(x) \right] \right) \right] f(x) \\
&+ \frac{\epsilon \mu_{2,0}^{(2)}}{2d \det(\Lambda_x)} \left[q_\alpha(x) \Delta^{(2)} p^{(2)}(x) \right] f(x) \\
&+ \frac{\epsilon \mu_{2,0}^{(1)}}{2d \det(\Lambda_x)} \left[2 q_\alpha(x) \left(\sum_{i=1}^d \lambda_i \left[\nabla_{E_i}^{(2)} p^{(2)}(x) \nabla_{E_i}^{(2)} f(x) \right] \right) \right. \\
&\quad \left. + 2 p^{(2)}(x) \left(\sum_{i=1}^d \lambda_i \left[\nabla_{E_i}^{(2)} q_\alpha(x) \nabla_{E_i}^{(2)} f(x) \right] \right) \right] \\
&+ \frac{\epsilon \mu_{2,0}^{(2)}}{2d \det(\Lambda_x)} 2 q_\alpha(x) \nabla^{(2)} p^{(2)}(x) \cdot \nabla^{(2)} f(x) \\
&+ \frac{\epsilon \mu_{2,0}^{(2)}}{2d \det(\Lambda_x)} q_\alpha(x) p^{(2)}(x) \Delta^{(2)} f(x) + \frac{\epsilon \mu_{2,0}^{(1)}}{2d \det(\Lambda_x)} q_\alpha(x) p^{(2)}(x) \sum_{i=1}^d \lambda_i \nabla_{E_i E_i}^{(2)^2} f(x) + \mathcal{O}(\epsilon^{3/2}).
\end{aligned}$$

The denominator of the left-hand side can be expanded and organized as

$$\begin{aligned}
& \epsilon^{-d(1-\alpha)} \int_{\mathcal{M}} \int_{\mathcal{M}} K_\epsilon^{(1)}(x, z) K_\epsilon^{(2)}(z, y) \frac{p_{\mathcal{Z}}^{(2)}(z)}{d_\epsilon^{(2)}(z)^\alpha} dV^{(2)}(z) p^{(2)}(y) dV^{(2)}(y) \\
&= \frac{p^{(2)}(x) q_\alpha(x)}{\det(\Lambda_x)} + \epsilon p^{(2)}(x) q_\alpha(x) W(x) - \epsilon \alpha q_\alpha(x) p^{(2)}(x) \bar{E}(x) \\
&+ \frac{\epsilon \mu_{2,0}^{(1)}}{2d \det(\Lambda_x)} \left[q_\alpha(x) \sum_{i=1}^d \lambda_i \nabla_{E_i E_i}^{(2)^2} p^{(2)}(x) + p^{(2)}(x) \sum_{i=1}^d \lambda_i \nabla_{E_i E_i}^{(2)^2} q_\alpha(x) \right. \\
&\quad \left. + 2 \left(\sum_{i=1}^d \lambda_i \left[\nabla_{E_i}^{(2)} p^{(2)}(x) \nabla_{E_i}^{(2)} q_\alpha(x) \right] \right) \right] \\
&+ \frac{\epsilon \mu_{2,0}^{(2)}}{2d \det(\Lambda_x)} \left[q_\alpha(x) \Delta^{(2)} p^{(2)}(x) \right] + \mathcal{O}(\epsilon^{3/2}).
\end{aligned}$$

Putting the above together, we obtain the claimed result

$$\begin{aligned}
T_{\text{lan},\epsilon,\alpha}f(x) &= f(x) + \frac{\epsilon\mu_{2,0}^{(2)}}{2d}\Delta^{(2)}f(x) + \frac{\epsilon\mu_{2,0}^{(1)}}{2d}\sum_{i=1}^d\lambda_i\nabla_{E_iE_i}^{(2)^2}f(x) \\
&\quad + \frac{\epsilon\mu_{2,0}^{(1)}}{d}\sum_{i=1}^d\lambda_i\left(\frac{\nabla_{E_i}^{(2)}p^{(2)}(x)}{p^{(2)}} + \frac{\nabla_{E_i}^{(2)}q_\alpha(x)}{q_\alpha(x)}\right)\nabla_{E_i}^{(2)}f(x) \\
&\quad + \frac{\mu_{2,0}^{(2)}}{d}\frac{\nabla^{(2)}p^{(2)}(x)\cdot\nabla^{(2)}f(x)}{p^{(2)}(x)} + \mathcal{O}(\epsilon^{3/2}).
\end{aligned}$$

SM1.3. Proof of Theorem 4.10. Recall the following results in [SM1], which states the variance analysis of landmark kernels.

Lemma SM1.8 (Equation (46) in [SM1]). *Take $\mathcal{X} = \{x_i\}_{i=1}^n$ and $\mathcal{Z} = \{z_k\}_{k=1}^m$, where $m = \lceil n^\beta \rceil$ for some $0 < \beta \leq 1$ and $\lceil r \rceil$ is the nearest integer of $r \in \mathbb{R}$. Then with probability higher than $1 - \mathcal{O}(1/n^2)$, we have*

$$\frac{1}{n}\sum_{i=1}^n\epsilon^{-d/2}K_\epsilon^{(2)}(z_k, x_i)K_\epsilon^{(2)}(x_i, z_l) = \epsilon^{-d/2}K_{\text{lan},\epsilon}^{(2)}(z_k, z_l) + O\left(\frac{\sqrt{\log(n)}}{n^{1/2}\epsilon^{d/4}}\right).$$

and

$$\frac{1}{m}\sum_{k=1}^m\epsilon^{-d/2}K_\epsilon^{(1)}(x_i, z_k)K_\epsilon^{(2)}(z_k, x_j) = \epsilon^{-d/2}K_{\text{lan},\epsilon,0}(z_k, z_l) + O\left(\frac{\sqrt{\log(n)}}{n^{\beta/2}\epsilon^{d/4}}\right).$$

where $K_{\text{lan},\epsilon}^{(2)}(z_i, z_j)$ is defined in Definition 4.2 and $K_{\text{lan},\epsilon,0}(z_i, z_j)$ is defined in Definition 4.4 with $\alpha = 0$.

Lemma SM1.9 (Lemma C.1 in [SM3]). *Take $\mathcal{X} = \{x_i\}_{i=1}^n$ and $\mathcal{Z} = \{z_k\}_{k=1}^m$, where $m = \lceil n^\beta \rceil$ for some $0 < \beta \leq 1$ and $\lceil r \rceil$ is the nearest integer of $r \in \mathbb{R}$. Then with probability higher than $1 - \mathcal{O}(1/n^2)$, we have*

$$\frac{1}{m}\sum_{l=1}^m\epsilon^{-d}K_{\text{lan},\epsilon}^{(2)}(z_k, z_l) = \epsilon^{-d}d_{\text{lan},\epsilon}^{(2)}(z_k) + O\left(\frac{\sqrt{\log(n)}}{n^{\beta/2}\epsilon^{d/4}}\right).$$

and

$$\frac{1}{n}\sum_{j=1}^n\epsilon^{-d}K_{\text{lan},\epsilon,0}(x_i, x_j) = \epsilon^{-d}d_{\text{lan},\epsilon,0}(x_i) + O\left(\frac{\sqrt{\log(n)}}{n^{1/2}\epsilon^{d/4}}\right).$$

where $d_{\text{lan},\epsilon}^{(2)}(z_i)$ and $d_{\text{lan},\epsilon,0}(x_i)$ are defined in Definition 4.2 and 4.4, respectively.

The following lemma shows the large deviation bound of original diffusion operator.

Lemma SM1.10 ([SM2]). Take $\mathcal{X} = \{x_i\}_{i=1}^n$ and $\mathcal{Z} = \{z_k\}_{k=1}^m$, where $m = \lceil n^\beta \rceil$ for some $0 < \beta \leq 1$ and $\lceil r \rceil$ is the nearest integer of $r \in \mathbb{R}$. Let $\epsilon = \epsilon(n)$ so that $\frac{\sqrt{\log n}}{n^{\beta/2}\epsilon^{d/2+1/2}} \rightarrow 0$ and $\epsilon \rightarrow 0$ when $n \rightarrow \infty$. Then with probability higher than $1 - \mathcal{O}(1/n^2)$, we have

$$\frac{\frac{1}{n} \sum_{j=1}^n K_{\text{lan},\epsilon,\alpha}(x_i, x_j)(f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n K_{\text{lan},\epsilon,\alpha}(x_i, x_j)} = f(x_i) - T_{\text{lan},\epsilon,\alpha}f(x_i) + \mathcal{O}\left(\frac{\sqrt{\log n}}{n^{1/2}\epsilon^{d/4-1/2}}\right).$$

Before we prove Theorem 4.10, we define some functions for convenience.

$$\hat{K}_{\text{lan},\epsilon,n}^{(2)}(z_k, z_l) := \frac{1}{n} \sum_{i=1}^n \epsilon^{-d/2} K_\epsilon^{(2)}(z_k, x_i) K_\epsilon^{(2)}(x_i, z_l),$$

$$\hat{d}_{\text{lan},\epsilon,m}^{(2)}(z_k) := \frac{1}{m} \sum_{l=1}^m \epsilon^{-d/2} \hat{K}_{\text{lan},\epsilon,m}^{(2)}(z_k, z_l),$$

$$\hat{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j) := \frac{1}{m} \sum_{k=1}^m \epsilon^{-d/2} K_\epsilon^{(1)}(x_i, z_k) K_\epsilon^{(2)}(z_k, x_j) \hat{d}_{\text{lan},\epsilon,m}^{(2)}(z_k)^{-\alpha},$$

$$\hat{d}_{\text{lan},\epsilon,\alpha,n}(x_i) := \frac{1}{n} \sum_{j=1}^n \epsilon^{-d/2} \hat{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j),$$

and

$$\tilde{d}_{\text{lan},\epsilon,m}^{(2)}(z_k) := \frac{1}{m} \sum_{l=1}^m \epsilon^{-d} K_{\text{lan},\epsilon}^{(2)}(z_k, z_l),$$

$$\tilde{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j) := \frac{1}{m} \sum_{k=1}^m \epsilon^{-d/2+d\alpha} K_\epsilon^{(1)}(x_i, z_k) K_\epsilon^{(2)}(z_k, x_j) \tilde{d}_{\text{lan},\epsilon}^{(2)}(z_k)^{-\alpha}.$$

Now, it is sufficient to prove the Theorem 4.10. We restate Theorem 4.10 here for convenience.

Theorem SM1.11. Take $\mathcal{X} = \{x_i\}_{i=1}^n$ and $\mathcal{Z} = \{z_k\}_{k=1}^m$, where $m = \lceil n^\beta \rceil$ for some $0 < \beta \leq 1$ and $\lceil n^\beta \rceil$ is the smallest integer greater than or equal to $x \in \mathbb{R}$. Take $f \in C^3(\mathcal{M})$ and denote $\mathbf{f} \in \mathbb{R}^n$ such that $\mathbf{f}_i = f(x_i)$. Let $\epsilon = \epsilon(n)$ so that $\frac{\sqrt{\log n}}{n^{\beta/2}\epsilon^{d/4+1}} \rightarrow 0$ and $\epsilon \rightarrow 0$ when $n \rightarrow \infty$. Then with probability higher than $1 - \mathcal{O}(1/n^2)$, we have

$$\frac{1}{\epsilon} \left[\left(I_n - \left(\mathbf{D}_\alpha^{(1)} \right)^{-1} \mathbf{W}_\alpha^{(1)} \mathbf{M}_\alpha^{(2)} \right) \mathbf{f} \right] (i) = \frac{f(x_i) - T_{\text{lan},\epsilon,\alpha}f(x_i)}{\epsilon} + \mathcal{O}\left(\frac{\sqrt{\log n}}{n^{\beta/2}\epsilon^{d/4+1}}\right),$$

for all $i = 1, 2, \dots, n$.

Proof. Fix $x_i \in \mathcal{M}$ and $\alpha \in [0, 1]$. For convenience, define

$$[L\mathbf{f}](i) := \frac{1}{\epsilon} \left[\left(I_n - \left(\mathbf{D}_\alpha^{(1)} \right)^{-1} \mathbf{W}_\alpha^{(1)} \mathbf{M}_\alpha^{(2)} \right) \mathbf{f} \right] (i),$$

for all $i = 1, 2, \dots, n$. By definition, we have

$$\begin{aligned} \epsilon[L\mathbf{f}](i) &= \frac{\frac{1}{n} \sum_{j=1}^n \hat{K}_{\text{lan}, \epsilon, \alpha, m}(x_i, x_j)(f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \hat{K}_{\text{lan}, \epsilon, \alpha, m}(x_i, x_j)} \\ &= \frac{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \hat{d}_{\text{lan}, \epsilon, m}^{(2)}(z_k)^{-\alpha} \right) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \hat{d}_{\text{lan}, \epsilon, m}^{(2)}(z_k)^{-\alpha} \right)} \\ &= \frac{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \left(\frac{1}{m} \sum_{l=1}^m \hat{K}_{\text{lan}, \epsilon, n}^{(2)}(z_k, z_l) \right)^{-\alpha} \right) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \left(\frac{1}{m} \sum_{l=1}^m \hat{K}_{\text{lan}, \epsilon, n}^{(2)}(z_k, z_l) \right)^{-\alpha} \right)}. \end{aligned}$$

First, we control the term $\hat{K}_{\text{lan}, \epsilon, m}^{(2)}$. For convenience, we rewrite

$$\begin{aligned} \hat{K}_{\text{lan}, \epsilon, n}^{(2)}(z_k, z_l) &= \epsilon^{-d/2} K_{\text{lan}, \epsilon}^{(2)}(z_k, z_l) - \left(\epsilon^{-d/2} K_{\text{lan}, \epsilon}^{(2)}(z_k, z_l) - \hat{K}_{\text{lan}, \epsilon, n}^{(2)}(z_k, z_l) \right) \\ &:= \epsilon^{-d/2} K_{\text{lan}, \epsilon}^{(2)}(z_k, z_l) - D_1(z_k, z_l). \end{aligned}$$

Calculate directly,

$$\epsilon[L\mathbf{f}](i) = J_1 + J_2 + J_3,$$

where

$$\begin{aligned} J_1 &:= \frac{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \left(\frac{1}{m} \sum_{l=1}^m K_{\text{lan}, \epsilon, n}^{(2)}(z_k, z_l) \right)^{-\alpha} \right) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \left(\frac{1}{m} \sum_{l=1}^m K_{\text{lan}, \epsilon, n}^{(2)}(z_k, z_l) \right)^{-\alpha} \right)}, \\ J_2 &:= \frac{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \left(\frac{1}{m} \sum_{l=1}^m D_1(z_k, z_l) \right)^{-\alpha} \right) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \left(\frac{1}{m} \sum_{l=1}^m \epsilon^{-d/2} K_{\text{lan}, \epsilon, \alpha}^{(2)}(z_k, z_l) \right)^{-\alpha} \right)}, \end{aligned}$$

and

$$\begin{aligned} J_3 &:= \frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \left(\frac{1}{m} \sum_{l=1}^m \hat{K}_{\text{lan}, \epsilon, n}^{(2)}(z_k, z_l) \right)^{-\alpha} \right) (f(x_i) - f(x_j)) \\ &\quad \times \left[\left(\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \left(\frac{1}{m} \sum_{l=1}^m \epsilon^{-d/2} K_{\text{lan}, \epsilon, \alpha}^{(2)}(z_k, z_l) \right)^{-\alpha} \right) \right)^{-1} \right. \\ &\quad \left. - \left(\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \left(\frac{1}{m} \sum_{l=1}^m \hat{K}_{\text{lan}, \epsilon, \alpha}^{(2)}(z_k, z_l) \right)^{-\alpha} \right) \right)^{-1} \right]. \end{aligned}$$

By Lemma SM1.8, with probability $1 - \mathcal{O}(n^{-2})$, we have

$$(SM1.5) \quad \hat{K}_{\text{lan},\epsilon,n}^{(2)}(z_k, z_l) = \epsilon^{-d/2} K_{\text{lan},\epsilon}^{(2)}(z_k, z_l) + \mathcal{O}\left(\frac{\sqrt{\log(n)}}{n^{1/2}\epsilon^{d/4}}\right).$$

Denote by Ω_1 the event space that Equation (SM1.5) holds. Under Ω_1 , we can easily control J_2 and J_3 , and hence we have

$$\begin{aligned} \epsilon[L\mathbf{f}](i) &= \frac{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \tilde{d}_{\text{lan},\epsilon,m}^{(2)}(z_k)^{-\alpha} \right) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \tilde{d}_{\text{lan},\epsilon,m}^{(2)}(z_k)^{-\alpha} \right)} \\ &\quad + \mathcal{O}\left(\alpha \frac{\sqrt{\log(n)}}{n^{1/2}\epsilon^{d/4}}\right), \end{aligned}$$

where the implied constant depends on $\|f\|_{\infty}$.

Second, again, calculate directly,

$$\begin{aligned} &\frac{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \tilde{d}_{\text{lan},\epsilon,m}^{(2)}(z_k)^{-\alpha} \right) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \tilde{d}_{\text{lan},\epsilon,m}^{(2)}(z_k)^{-\alpha} \right)} \\ &= J_4 + J_5 + J_6, \end{aligned}$$

where

$$\begin{aligned} J_4 &:= \frac{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) d_{\text{lan},\epsilon}^{(2)}(z_k)^{-\alpha} \right) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) d_{\text{lan},\epsilon}^{(2)}(z_k)^{-\alpha} \right)}, \\ J_5 &:= \frac{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) D_2(z_k) \right) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \epsilon^{d\alpha} d_{\text{lan},\epsilon}^{(2)}(z_k)^{-\alpha} \right)}, \end{aligned}$$

and

$$\begin{aligned} J_6 &:= \frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \tilde{d}_{\text{lan},\epsilon,m}^{(2)}(z_k)^{-\alpha} \right) (f(x_i) - f(x_j)) \\ &\quad \times \left[\left(\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \epsilon^{d\alpha} d_{\text{lan},\epsilon}^{(2)}(z_k)^{-\alpha} \right) \right)^{-1} \right. \\ &\quad \left. - \left(\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) \tilde{d}_{\text{lan},\epsilon,m}^{(2)}(z_k)^{-\alpha} \right) \right)^{-1} \right], \end{aligned}$$

where $D_2(z_k) := \tilde{d}_{\text{lan},\epsilon,m}^{(2)}(z_k)^{-\alpha} - \epsilon^{d\alpha} d_{\text{lan},\epsilon}^{(2)}(z_k)^{-\alpha}$. By Lemma SM1.9, with probability $1 - \mathcal{O}(n^{-2})$, we have

$$(SM1.6) \quad \tilde{d}_{\text{lan},\epsilon,m}^{(2)} = \epsilon^{-d} d_{\text{lan},\epsilon}^{(2)}(z_k) + \mathcal{O}\left(\frac{\sqrt{\log(n)}}{n^{\beta/2}\epsilon^{d/4}}\right).$$

Denote by Ω_2 the event space that Equation (SM1.6) holds. Thus, putting the above together, under $\Omega_1 \cap \Omega_2$, we can control J_5 and J_6 and hence we have

$$\begin{aligned} \epsilon[L\mathbf{f}](i) &= \frac{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) d_{\text{lan},\epsilon}^{(2)}(z_k)^{-\alpha} \right) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \left(\frac{1}{m} \sum_{k=1}^m K_{\epsilon}^{(1)}(x_i, z_k) K_{\epsilon}^{(2)}(z_k, x_j) d_{\text{lan},\epsilon}^{(2)}(z_k)^{-\alpha} \right)} \\ &\quad + \mathcal{O}\left(\alpha \frac{\sqrt{\log(n)}}{n^{\beta/2}\epsilon^{d/4}}\right), \end{aligned}$$

where $\frac{\sqrt{\log(n)}}{n^{\beta/2}\epsilon^{d/4}} \geq \frac{\sqrt{\log(n)}}{n^{1/2}\epsilon^{d/4}}$ and the implied constant depends on $\|f\|_{\infty}$.

Third, again, calculate directly,

$$\begin{aligned} (SM1.7) \quad & \frac{\frac{1}{n} \sum_{j=1}^n \tilde{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \tilde{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j)} \\ &= \frac{\frac{1}{n} \sum_{j=1}^n K_{\text{lan},\epsilon,\alpha}(x_i, x_j) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n K_{\text{lan},\epsilon,\alpha}(x_i, x_j)} \\ &\quad + \frac{\frac{1}{n} \sum_{j=1}^n (\tilde{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j) - \epsilon^{-d/2+d\alpha} K_{\text{lan},\epsilon,\alpha}(x_i, x_j)) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n \epsilon^{-d/2+d\alpha} K_{\text{lan},\epsilon,\alpha}(x_i, x_j)} \\ &\quad + \frac{1}{n} \sum_{j=1}^n \tilde{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j) (f(x_i) - f(x_j)) \\ &\quad \times \left[\left(\frac{1}{n} \sum_{j=1}^n \epsilon^{-d/2+d\alpha} K_{\text{lan},\epsilon,\alpha}(x_i, x_j) \right)^{-1} - \left(\frac{1}{n} \sum_{j=1}^n \tilde{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j) \right)^{-1} \right]. \end{aligned}$$

We need to control the term $|\tilde{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j) - \epsilon^{-d/2+d\alpha} K_{\text{lan},\epsilon,\alpha}(x_i, x_j)|$, which is generalized the proof of Lemma SM1.8. Let a random variable

$$F := \epsilon^{-d/2+d\alpha} K_{\epsilon}^{(1)}(x_i, Z) K_{\epsilon}^{(2)}(Z, x_j) d_{\text{lan},\epsilon}^{(2)}(Z)^{-\alpha},$$

where $Z \sim p_Z^{(2)}$. Denote by F_k one realization of F when the realization of random variable Z is z_k . Now, we need to find the convergence rate of

$$\frac{1}{m} \sum_{k=1}^m F_k \rightarrow \mathbb{E}[F].$$

By Lemma SM1.5, let $y = \exp_{x_i}^{(2)} v$ where $v \in T_{x_i} \mathcal{M}$ and we have

$$\begin{aligned} \mathbb{E}[F] &= \epsilon^{-d/2+d\alpha} \tilde{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j) \\ &= \epsilon^{-d/2+d\alpha} \int_{\mathcal{M}} K_{\epsilon}^{(1)}(x_i, z) K_{\epsilon}^{(2)}(z, x_j) d_{\text{lan},\epsilon}^{(2)}(z)^{-\alpha} p_{\mathcal{Z}}^{(2)}(z) dV(z) \\ &= p^{(2)}(z)^{-\alpha} p_{\mathcal{Z}}^{(2)}(z)^{1-\alpha} A_0(v) + \mathcal{O}(\epsilon). \end{aligned}$$

Furthermore, by Equation (SM1.2), we know $A_0(v)$ decay fast as $\|v\|_{g^{(2)}}$ small. On the other hand, we have

$$\mathbb{E}[F^2] = \epsilon^{-d+2d\alpha} \int_{\mathcal{M}} K_{\epsilon}^{(1)}(x_i, z)^2 K_{\epsilon}^{(2)}(z, x_j)^2 d_{\text{lan},\epsilon}^{(2)}(z)^{-2\alpha} p_{\mathcal{Z}}^{(2)}(z) dV(z) \asymp \epsilon^{-d/2}.$$

Hence, we have $\text{Var}(F) \asymp \epsilon^{-d/2}$ and $F_k \asymp \epsilon^{-d/2}$. By Bernstein inequality, we know

$$\mathbb{P}\left(\frac{1}{m} \sum_{k=1}^m F_k - \mathbb{E}(F) > t\right) \leq \exp\left(-\frac{mt^2}{2\epsilon^{-d/2} + \frac{2}{3}\epsilon^{-d/2}t}\right).$$

Since we ask $\frac{t}{\mathbb{E}[F]} \rightarrow 0$ when $m \rightarrow \infty$, the exponent becomes

$$\frac{mt^2}{2\epsilon^{-d/2} + \frac{2}{3}\epsilon^{-d/2}t} \geq \frac{mt^2}{3\epsilon^{-d/2}}.$$

Then if we choose m such that $\frac{mt^2}{3\epsilon^{-d/2}} = 4\log(n)$, we have

$$t \asymp \frac{\sqrt{\log n}}{n^{\beta/2}\epsilon^{d/4}},$$

which goes to 0 by our assumption $\frac{\sqrt{\log n}}{n^{\beta/2}\epsilon^{d/4}} \rightarrow 0$. By union bound and get for all pairs (x_i, x_j) , $i, j = 1, \dots, n$, with probability $1 - \mathcal{O}(n^{-2})$

$$(SM1.8) \quad \tilde{K}_{\text{lan},\epsilon,\alpha,m}(x_i, x_j) = \epsilon^{-d/2+d\alpha} K_{\text{lan},\epsilon,\alpha}(x_i, x_j) + \mathcal{O}\left(\frac{\sqrt{\log n}}{n^{\beta/2}\epsilon^{d/4}}\right).$$

Denote by Ω_3 the event space that Equation (SM1.8) holds. Thus, putting the above together, under $\Omega_1 \cap \Omega_2 \cap \Omega_3$, we have

$$\epsilon[Lf](i) = \frac{\frac{1}{n} \sum_{j=1}^n K_{\text{lan},\epsilon,\alpha}(x_i, x_j)(f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n K_{\text{lan},\epsilon,\alpha}(x_i, x_j)} + \mathcal{O}\left(\frac{\sqrt{\log(n)}}{n^{\beta/2}\epsilon^{d/4}}\right),$$

where the implied constant depends on $\|f\|_{\infty}$.

Finally, we need to apply Lemma SM1.10 to get the convergence rate. Let random variables

$$\begin{aligned} G &:= \epsilon^{-d+d\alpha} K_{\text{lan},\epsilon,\alpha}(x_i, Y)(f(x_i) - f(Y)), \\ H &:= \epsilon^{-d+d\alpha} K_{\text{lan},\epsilon,\alpha}(x_i, Y), \end{aligned}$$

where $Y \sim p^{(2)}$.

By Equation (SM1.3), we have

$$\begin{aligned}
\mathbb{E}[G] &= \frac{\epsilon \mu_{2,0}^{(1)}}{2d \det(\Lambda_x)} 2q_\alpha(x_i) \left(\sum_{i=1}^d \lambda_i \left[\nabla_{E_i}^{(2)} p^{(2)}(x) \nabla_{E_i}^{(2)} (f(x_i) - f(x))|_{x=x_i} \right] \right) \\
&\quad + \frac{\epsilon \mu_{2,0}^{(1)}}{2d \det(\Lambda_x)} 2p^{(2)}(x_i) \left(\sum_{i=1}^d \lambda_i \left[\nabla_{E_i}^{(2)} q_\alpha(x) \nabla_{E_i}^{(2)} (f(x_i) - f(x))|_{x=x_i} \right] \right) \\
&\quad + \frac{\epsilon \mu_{2,0}^{(2)}}{2d \det(\Lambda_x)} 2q_\alpha(x_i) \nabla^{(2)} p^{(2)}(x) \cdot \nabla^{(2)} (f(x_i) - f(x))|_{x=x_i} \\
&\quad + \frac{\epsilon \mu_{2,0}^{(2)}}{2d \det(\Lambda_x)} q_\alpha(x_i) p^{(2)}(x_i) \Delta^{(2)} (f(x_i) - f(x))|_{x=x_i} \\
&\quad + \frac{\epsilon \mu_{2,0}^{(1)}}{2d \det(\Lambda_x)} q_\alpha(x_i) p^{(2)}(x_i) \sum_{i=1}^d \lambda_i \nabla_{E_i E_i}^{(2)^2} (f(x_i) - f(x))|_{x=x_i} + \mathcal{O}(\epsilon^{3/2}),
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[G^2] &= \frac{\epsilon^{1-d/2} \mu_{2,0,2}^{(2)}}{2d \det(\Lambda_x)} q_\alpha(x_i) p^{(2)}(x_i) \Delta^{(2)} (f(x_i) - f(x))^2|_{x=x_i} \\
&\quad + \frac{\epsilon^{1-d/2} \mu_{2,0,2}^{(1)}}{2d \det(\Lambda_x)} q_\alpha(x_i) p^{(2)}(x_i) \sum_{i=1}^d \lambda_i \nabla_{E_i E_i}^{(2)^2} (f(x_i) - f(x))^2|_{x=x_i} + \mathcal{O}(\epsilon^{3/2}) \\
&= \frac{\epsilon^{1-d/2} \mu_{2,0,2}^{(2)}}{2d \det(\Lambda_x)} q_\alpha(x_i) p^{(2)}(x_i) \|\nabla^{(2)} (f(x_i) - f(x))\|^2|_{x=x_i} \\
&\quad + \frac{\epsilon^{1-d/2} \mu_{2,0,2}^{(1)}}{2d \det(\Lambda_x)} q_\alpha(x_i) p^{(2)}(x_i) \sum_{i=1}^d \lambda_i |\nabla_{E_i}^{(2)} (f(x_i) - f(x))|^2|_{x=x_i} + \mathcal{O}(\epsilon^{3/2}),
\end{aligned}$$

where $\mu_{2,0,2}^{(\ell)} = \int_{\mathbb{R}^d} \|x\|^2 \tilde{K}^{(\ell)}(\|x\|)^2 dx$. By Equation (SM1.4), we have

$$\mathbb{E}[H] = \frac{p^{(2)}(x) q_\alpha(x)}{\det(\Lambda_x)} + \mathcal{O}(\epsilon),$$

and

$$\mathbb{E}[H^2] = \epsilon^{-d/2} \mu_{0,0,2}^{(1)} \mu_{0,0,2}^{(2)} \frac{p^{(2)}(x) q_\alpha(x)}{\det(\Lambda_x)} + \mathcal{O}(\epsilon^{-d/2+1}),$$

where $\mu_{0,0,2}^{(\ell)} = \int_{\mathbb{R}^d} \tilde{K}^{(\ell)}(\|x\|)^2 dx$. Now, it is sufficient to apply Lemma SM1.10 to conclude the proof and further look closer to what the implied constant depends. Thus, with probability $1 - \mathcal{O}(n^{-2})$, we have

$$(SM1.9) \quad \frac{\frac{1}{n} \sum_{j=1}^n K_{\text{lan}, \epsilon, \alpha}(x_i, x_j) (f(x_i) - f(x_j))}{\frac{1}{n} \sum_{j=1}^n K_{\text{lan}, \epsilon, \alpha}(x_i, x_j)}$$

$$= f(x_i) - T_{\text{lan}, \epsilon, \alpha} f(x_i) + \mathcal{O}\left(\frac{\sqrt{\log n}}{n^{1/2} \epsilon^{d/4-1/2}}\right),$$

where the implied constant depends on

$$\|\nabla^{(2)}(f(x_i) - f(x))\|^2|_{x=x_i} + \sum_{i=1}^d \lambda_i |\nabla_{E_i}^{(2)}(f(x_i) - f(x))|^2|_{x=x_i}.$$

Denote by Ω_4 the event space that Equation (SM1.9) holds. Thus, putting the above together, under $\Omega_1 \cap \Omega_2 \cap \Omega_3 \cap \Omega_4$,

$$(SM1.10) \quad [L\mathbf{f}](i) = \frac{f(x_i) - T_{\text{lan}, \epsilon, \alpha} f(x_i)}{\epsilon} + \mathcal{O}\left(\frac{\sqrt{\log n}}{n^{\beta/2} \epsilon^{d/4+1}}\right),$$

where $\frac{\sqrt{\log n}}{n^{1/2} \epsilon^{d/4-1/2}} \leq \frac{\sqrt{\log n}}{n^{\beta/2} \epsilon^{d/4}}$. That is, with probability $1 - \mathcal{O}(n^{-2})$, Equation (SM1.10) holds. ■

Before closing this section, we provide a numerical exploration of the obtained convergence rate in the variance analysis by studying the relation between n and ϵ . We conduct two experiments following [SM2] to examine the variance term in Theorem 4.10. First, consider the canonical $\mathbb{S}^1$ and the scaled circle S , defined as

$$\begin{aligned} \mathbb{S}^1 &= \{(\cos \theta, \sin \theta)\} \subset \mathbb{R}^2, \\ S_1 &= \{(1.5 \cos \theta, 1.5 \sin \theta)\} \subset \mathbb{R}^2, \end{aligned}$$

where $\theta \in [-\pi, \pi)$. We randomly select 3000 points $\{(s_i, r_i) : s_i \in S_1, r_i \in \mathbb{S}^1\}$ uniformly and we select 1500 landmark points followed the non-uniform p.d.f. $p_{\mathcal{Z}}^{(2)}(\theta) = \frac{58}{50}[0.48 \cos \theta + 0.52]$. Consider the function $f(x, y) = x$ and evaluate at $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. We have $\Delta^{(2)} f|_{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})} = -\frac{1}{\sqrt{2}}$,

$\lambda_1 \nabla_{E_1, E_1}^{(2)^2} f|_{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})} = \frac{-2}{3\sqrt{2}}$ and $\lambda_1 \frac{\nabla_{E_1}^{(2)} p_{\mathcal{Z}}^{(2)^{1-\alpha}}}{p_{\mathcal{Z}}^{(2)^{1-\alpha}}} \nabla_{E_1}^{(2)} f|_{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})} \neq 0$. Then, for fix x_i , we conduct $T = 50$ independent trials and compute the deviation

$$\sqrt{\frac{1}{T} \sum_{t=1}^T \left| \frac{1}{\epsilon} \left[\left(I_n - \left(\mathbf{D}_{\alpha, t}^{(1)} \right)^{-1} \mathbf{W}_{\alpha, t}^{(1)} \mathbf{M}_{\alpha, t}^{(2)} \right) \mathbf{f} \right] (i) - \frac{f(x_i) - T_{\text{lan}, \epsilon, \alpha} f(x_i)}{\epsilon} \right|^2},$$

where $\mathbf{D}_{\alpha, t}^{(1)}$, $\mathbf{D}_{\alpha, t}^{(1)}$ and $\mathbf{M}_{\alpha, t}^{(2)}$ are constructed in the t -th trial. The relationship between the deviation and ϵ is shown in the left subfigure of Figure SM1. According to Theorem 4.10, we expect the convergence rate to be $\epsilon^{5/4}$. However, a linear fit yields slopes 0.75, 0.71, 0.73, 0.71 and 0.73 when $\alpha = 0, 0.25, 0.5, 0.75, 1.0$ respectively. This numerical result is faster than our analysis.

Second, consider the canonical $\mathbb{S}^2$ and the scaled sphere S_2 , defined as

$$\begin{aligned} \mathbb{S}^2 &= \{(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)\} \subset \mathbb{R}^3, \\ S_2 &= \{(0.8 \sin \theta \cos \phi, 0.8 \sin \theta \sin \phi, 0.8 \cos \theta)\} \subset \mathbb{R}^3, \end{aligned}$$

where $\phi \in [-2\pi, 2\pi)$ and $\theta \in [-\pi, \pi)$. We randomly select 3000 points $\{(s_i, r_i) : s_i \in S_2, r_i \in \mathbb{S}^2\}$ uniformly and we select 2000 landmark points uniformly. Consider the function $f(x, y, z) = x + z$ and evaluate at $(0, 0, 1)$. The relationship between the deviation and

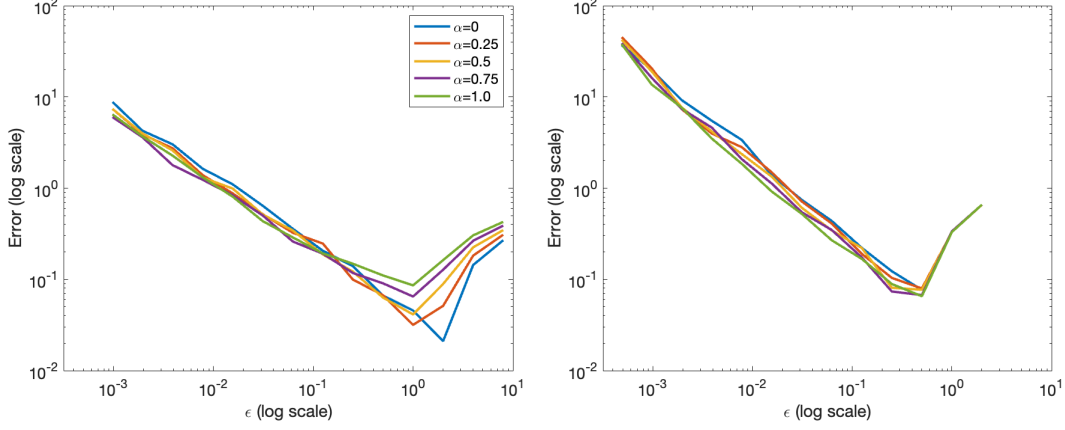


Figure SM1. *Left: The deviation between the discrete and continuous α -LAD operator on $\mathbb{S}^1$ and S_1 . Right: The deviation between the discrete and continuous α -LAD operator on $\mathbb{S}^2$ and S_2 .*

ϵ is shown in the right subfigure of Figure SM1. According to Theorem 4.10, we expect the convergence rate to be $\epsilon^{3/2}$. However, a linear fit yields slopes 0.92, 0.94, 0.94, 0.96 and 0.96 for $\alpha = 0, 0.25, 0.5, 0.75, 1.0$, respectively, which is also faster than our analysis.

According to Figure SM1, we anticipate that the convergence rate should be $\epsilon^{d/4+1/2}$ but it does not match the theoretical convergence rate we derive. The main reason is that we control the stochastic fluctuation by controlling the datasets and landmarks separately. For example, in (SM1.7), the control is divided into three terms, where the last two terms contribute to this slower theoretical convergence rate. We hypothesize that by properly handling the dependence induced by the diffusion between datasets and landmarks, we could improve the theoretical convergence rate.

SM2. Additional Examples for Section 5.3. To focus on the fact that the embedding manifold to which α -LAD converges at $\alpha = 1$ might not be the embedding manifold of AD, we simplify example in Figure 4 by setting both manifolds $\iota^{(1)}(\mathcal{M}) = \iota^{(2)}(\mathcal{M})$ be $\mathbb{S}^1$. We set define $\Phi_4 : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ is identity. By the same approach as in Section 5.3, we sample 3500 points $\{(s_i, r_i) : s_i = r_1 \in \mathbb{S}^1\}$ following non-constant $p^{(2)}$ and choose 1000 landmark points following different landmark distributions. We label these four different $p_{\mathcal{Z}}^{(2)}$ cases as “Case 1” to “Case 4.” One can find the sampling density functions we used in top row of Figure 4. The embedding manifolds under these different landmark distributions as α approaches 1 are shown in Figure SM2. We see that as α approaches 1, the embeddings by LAD are more similar, but these embeddings are clearly different from embeddings by AD.

SM3. Revisit Examples for Section 5.3. Under the same conditions in Section 5.3, including manifolds, sampling density functions, we increase the number of sampling points to 10,000, set the size of landmark set to 1,000, and present revised Figures 3, 4, 5 as shown in Figures SM3, SM4, SM5.

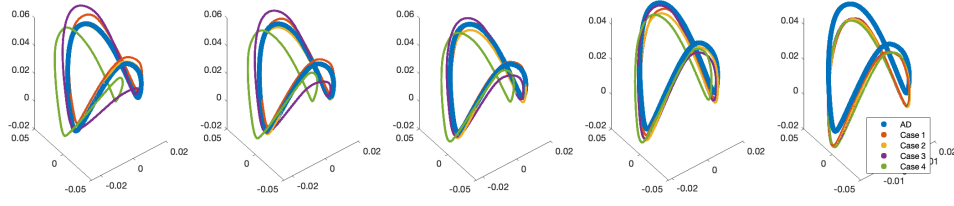


Figure SM2. The results of α -LAD with $\alpha = 0, 1/4, 1/2, 3/4$, and 1, respectively. The blue curve is the result of AD, which is the same from left to right. “Case 1” to “Case 4” are α -LAD with different $p_Z^{(2)}$ introduced in Figure 4.

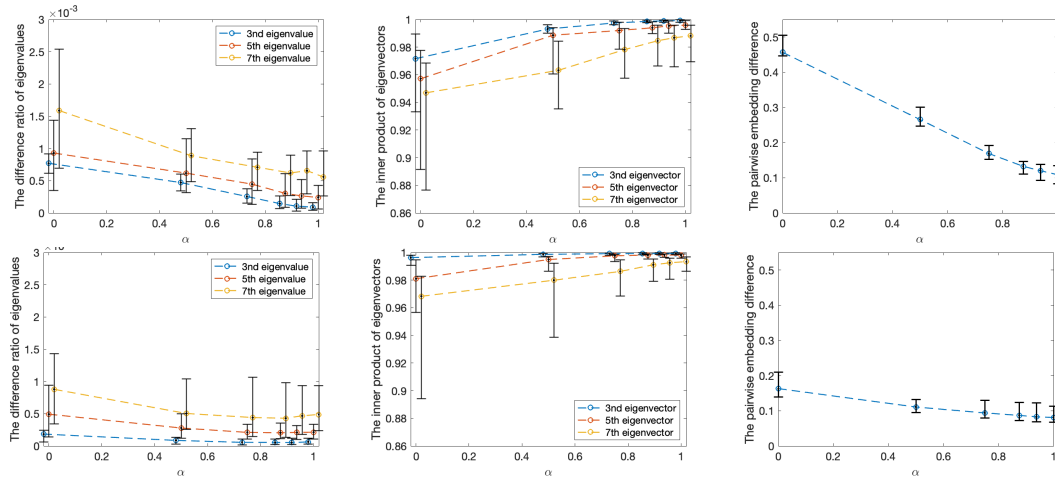


Figure SM3. Updated Figure 3 with 10,000 sample points and 1,000 landmark points. Top row, from left to right: α -LAD with various α , the difference ratio of the 2nd, 4th, 6th non-trivial eigenvalues between AD and α -LAD, the inner product between the 2nd, 4th, 6th non-trivial eigenvectors of AD and α -LAD and the pairwise embedding difference between AD and α -LAD, where the landmark distribution is non-uniform. Bottom row shows the same thing but when the landmark distribution is uniform.

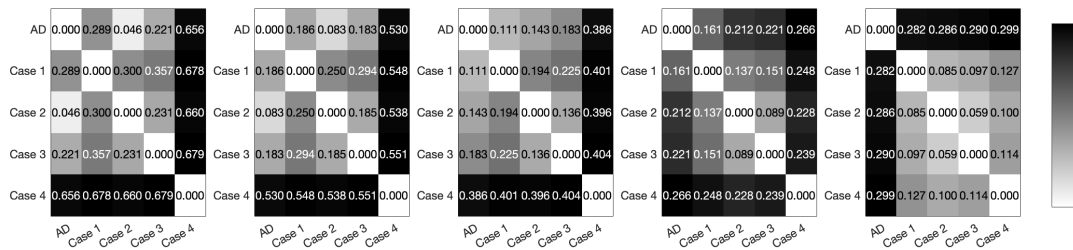


Figure SM4. Updated lower row of Figure 4 with 10,000 sample points and 1,000 landmark points. From left to right: the pairwise embedding difference between different embeddings with $\alpha = 0, 1/4, 1/2, 3/4$, and 1, respectively. “Case 1” to “Case 4” are α -LAD with different $p_Z^{(2)}$ introduced in Figure 4.

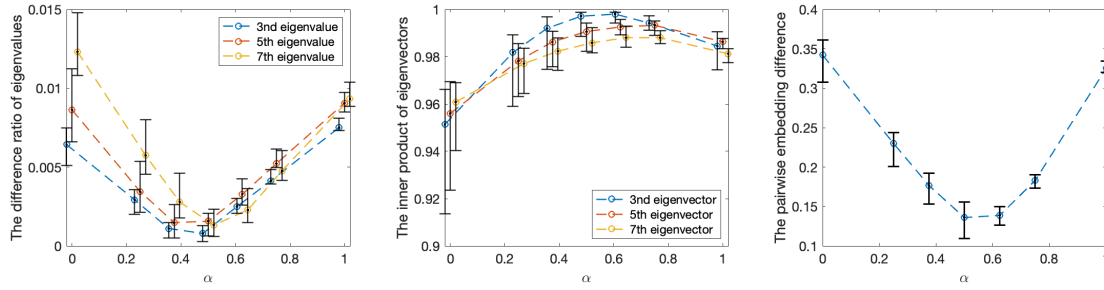


Figure SM5. Updated Figure 5 with 10,000 sample points and 1,000 landmark points. From left to right: α -LAD with various α , the difference ratio of the 2nd, 4th, 6th non-trivial eigenvalues between AD and α -LAD, the inner product between the 2nd, 4th, 6th non-trivial eigenvectors of AD and α -LAD and the pairwise embedding difference between AD and α -LAD.

REFERENCES

- [SM1] C. SHEN AND H.-T. WU, *Scalability and robustness of spectral embedding: landmark diffusion is all you need*, Information and Inference: A Journal of the IMA, 11 (2022), pp. 1527–1595, <https://doi.org/10.1093/imaiai/iaac013>.
- [SM2] A. SINGER, *From graph to manifold laplacian: The convergence rate*, Applied and Computational Harmonic Analysis, 21 (2006), pp. 128–134.
- [SM3] A. SINGER AND H.-T. WU, *Spectral convergence of the connection laplacian from random samples*, Information and Inference: A Journal of the IMA, 6 (2017), pp. 58–123.
- [SM4] R. TALMON AND H.-T. WU, *Latent common manifold learning with alternating diffusion: Analysis and applications*, Applied and Computational Harmonic Analysis, 47 (2019), pp. 848–892.