Notes for Week 4/10/20 - 4/16/20

Alden Green

April 23, 2020

Let $\mathcal{M}$ be a closed, connected, smooth manifold without boundary of dimension m embedded in $\mathbb{R}^d$. We give to $\mathcal{M}$ the Riemannian structure induced by the ambient space $\mathbb{R}^d$, and denote the volume form by μ. Let P be a distribution supported on $\mathcal{M}$, with density p with respect to the volume form μ. Suppose we observe n samples $X_1, \ldots, X_n$ drawn independently from P. In addition, we observe responses

$$Y_i = f_0(X_i) + \varepsilon_i$$

where ε_i are i.i.d $N(0,1)$ noise, and $f_0 : \mathcal{M} \to \mathbb{R}$ is an unknown function we wish to estimate.

We construct a neighborhood graph $G_{n,r}$ over the data $X_1, \ldots, X_n$ as follows. For a kernel $K : \mathbb{R}^d \to [0, \infty)$ and radius $r > 0$, we let

$$K_r(X_i, X_j) = K\left(\frac{\|X_i - X_j\|_{\mathbb{R}^d}}{r}\right)$$

be the weight of the edge connecting X_i and X_j. The graph $G_{n,r}$ has an associated Laplacian matrix $L_{n,r}$, which is defined by the action

$$\mathbf{L}_{n,r} f(X_i) := \sum_{i=1}^{n} \Big(f(X_i) - f(X_j)\Big) K_r(X_i, X_j)$$

We write $(\lambda_1(G_{n,r}), v_1(G_{n,r})), \ldots, (\lambda_n(G_{n,r}), v_n(G_{n,r}))$ for the n eigenvalue/eigenvector pairs of $\mathbf{L}_{n,r}$, and adopt the usual convention of arranging the eigenvalues in ascending order, meaning $0 = \lambda_1(G_{n,r}) \leq \lambda_1(G_{n,r}) \leq \cdots \leq \lambda_1(G_{n,r})$. The Laplacian eigenmaps estimator is a nonparametric estimator which projects the data onto the span of the first κ eigenvectors

$$\widehat{f}_{\mathrm{LE}} = \sum_{k=1}^{\kappa} \langle Y, v_k(G_{n,r}) \rangle_n v_k(G_{n,r})$$

where $1 \leq \kappa \leq n$ is tuning parameter. We would like to show that $f_0 \in C^1(\mathcal{M}, Q)$ for a value of Q fixed in n, then for an appropriate choice of κ

$$\|\widehat{f}_{\mathrm{LE}} - f_0\|_n^2 \lesssim n^{-2/(2+m)}$$

with high probability.

1 Theory

Recall that the squared bias of $\widehat{f}_{\mathrm{LE}}$ is upper bounded

$$\left\|\mathbb{E}\Big[\widehat{f}_{\mathrm{LE}} \Big| \mathbf{X}\Big] - f_0\right\|_n^2 \leq \frac{f_0^T \mathbf{L}_{n,r} f_0}{n \lambda_\kappa(G_{n,r})} \tag{1}$$

(We have already supplied a sufficient bound on the variance which is independent of the dimension of $\mathcal{M}$). Under certain assumptions, we shall prove upper and lower bounds on the graph Sobolev seminorm $f_0^T \mathbf{L}_{n,r} f_0$ and the graph Laplacian eigenvalue $\lambda_\kappa(G_{n,r})$, respectively. These assumptions are as follows:

(A1) The radius r satisfies the bounds

$$(m+5)d_\infty(P,P_n) \leq r \leq \min\left\{1, \frac{i_0}{10}, \frac{1}{\sqrt{mK}}, \frac{R}{\sqrt{27m}}\right\}$$

Here, R is an upper bound on the reach of $\mathcal{M}$, $\mathcal{K}$ is an upper bound on the absolute values of the sectional curvatures of $\mathcal{M}$, i_0 is the injectivity radius of $\mathcal{M}$.

(A2) The kernel function $K : [0,\infty) \to [0,\infty)$ is supported on $[0,1]$ and is Q_K-Lipschitz continuous on its support, and satisfies $K(0) = 1$.

(A3) The density function p is a Q_p-Lipschitz continuous function on $\mathcal{M}$, and is bounded away from 0 and ∞,

$$0 < p_{\min} \leq p(x) \leq p_{\max} < \infty$$

for all $x \in \mathcal{M}$.

The bounds we prove are as follows.

(I) **Graph Sobolev semi-norm:** Assume (A1)-(A3). There exists a constant C such that the following statement holds: for any $f_0 \in C^1(\mathcal{M}, Q)$ and any $r > 0$,

$$f_0^T \mathbf{L}_{n,r} f_0 \leq \frac{C}{\delta} Q^2 n^2 r^{m+2} \tag{2}$$

with probability at least $1-\delta$.

(II) **Laplacian eigenvalue:** Assume (A1)-(A3). There exist constants $c, C, k_\star$ and β such that the following statement holds: for all $\kappa \in \mathbb{N}$ satisfying $\kappa \geq k_\star$, and any r satisfying

$$C\frac{\log(n)^{p_d}}{n^{1/d}} < r < c\kappa^{-1/m} \tag{3}$$

the graph Laplacian eigenvalue satisfies the lower bound

$$\lambda_\kappa(G_{n,r}) \geq cnr^{m+2}\kappa^{2/m}. \tag{4}$$

with probability at least $1 - Cn^{-\beta}$.

1.1 (I): Graph Sobolev semi-norm.

We shall prove the following upper bound on the expectation of the graph Sobolev semi-norm.

Lemma 1. *Suppose assumptions (A1)-(A3) are satisfied. Then there exists a constant $C_1 > 0$ such that for any $f \in C^1(\mathcal{M}, Q)$,*

$$\mathbb{E}\Big[f_0^T \mathbf{L}_{n,r} f_0\Big] \leq C_1 p_{\max} Q^2 n^2 r^{m+2} \tag{5}$$

with probability at least $1 - n\exp(c_0 n r^m p_{\min})$.

Note that (2) then follows by Markov's inequality.

1.2 (II): Graph Laplacian eigenvalue

Let Δ_p be defined for smooth functions $f : \mathcal{M} \to \mathbb{R}$ as

$$\Delta_p f = -\frac{1}{p}\mathrm{div}\big(\nabla(p^2 f)\big)$$

Δ_p has a point spectrum, and we denote its eigenvalues by $\lambda_1(\mathcal{M}) \leq \lambda_2(\mathcal{M}) \leq \cdots$.

In Lemma 2, we will prove a lower bound on the graph Laplacian eigenvalues $\lambda_k(G_{n,r})$ by the eigenvalues $\lambda_k(\mathcal{M})$. The lower bound will be a simple corollary of the developments in [Trillos et al., 2019], which we summarize in 3.1.

Lemma 2 (Lower bound on graph eigenvalue $\lambda_k(G_{n,r})$)**.** *Assume (A1)-(A3). Then there exists some $k_\star \in \mathbb{N}$ such that for any $k \geq k_\star$, the following statement holds: if*

$$8C_3 A\bigg(\frac{\log(n)^{p_d}}{n^{1/d}}\bigg) < r < \min\bigg\{\frac{C_2}{C_5 k^{1/m}}, \frac{1}{8C_3 L(p)}, \frac{1}{8C_3 k^{1/m}}, \frac{1}{8c_3\sqrt{\mathcal{K}}}\bigg\}$$

then

$$\lambda_k(G_{n,r}) \geq \frac{c_1}{2} n r^{m+2} k^{2/m}$$

with probability at least $1 - C_4 n^{-\beta}$.

1.3 (III): Higher order graph Sobolev semi-norms

We have the following bound on $f^T \mathbf{L}_{n,r}^2 f$ for functions $f \in C^2(\mathcal{M})$.

Lemma 3. *Assume (A1)-(A3). Let $f \in C^2(\mathcal{M}, Q)$. Then there exists constants C_7 and c_2 which depend only on $m, \mathcal{K}, R$ and i_0 such that*

$$f^T \mathbf{L}_{n,r}^2 f \leq 2C_7^2 Q^2 n^3 r^{2(m+2)} p_{\max}^2$$

with probability at least $1 - 2n\exp(-2cnr^{m+2})$.

A similar bound holds on $f^T \mathbf{L}_{n,r}^3 f$ for functions $f \in C^3(\mathcal{M})$.

Lemma 4. *Assume (A1)-(A3). Let $f \in C^3(\mathcal{M}, Q)$ and $p \in C^2(\mathcal{M}, Q)$. Then there exist constants C_8 and c_3 which depend only on $m, \mathcal{K}, R, i_0$ and Q_K such that*

$$f^T \mathbf{L}_{n,r}^3 f \leq \frac{C_8}{\delta} Q^2 n^4 r^{3(m+2)} p_{\max}^3 (p_{\max} + 1) \tag{6}$$

with probability at least $1 - 2n\exp(-c_2 n r^{4+m}) - n\exp(4n\nu_m r^m p_{\min}/3) - \delta$.

When we turn to $f^T \mathbf{L}_{n,r}^s f$ for $s \geq 4$, things become difficult, and at present I cannot obtain results analogous to 3 and 4. I detail what I find difficult—in the $s = 4$ case—in Subsection 3.5.

Finally, we note the following bound on the graph Sobolev seminorm when $f \in H^1(P, Q)$, a direct corollary of Lemma 6 in [Trillos et al., 2019] (c.f. Lemma 3.3 [Burago et al., 2014]) plus Markov's inequality.

Lemma 5. *Assume (A1)-(A3). Then there exists a universal constant $C > 0$ such that the following statement holds for any $f \in H^1(P)$:*

$$f^T \mathbf{L}_{n,r} f \leq \frac{1}{\delta}\big(1 + Q_p r\big)\big(1 + CmKr^2\big)\sigma_K r^{m+2} |f|_{H^1(P)}$$

with probability at least $1 - \delta$.

2 Proof of Theorems and Major Lemmas

We will often make use of the following geometric estimates, which hold under assumption (A1). These help us to convert from integrals over $\mathcal{M}$ to integrals over a Euclidean space. First

$$d_{\mathcal{M}}(y,x) \leq \|x-y\|_{\mathbb{R}^d} + \frac{8}{R^2}\|x-y\|_{\mathbb{R}^d}^3 \tag{7}$$

and additionally

$$|\mu(B(x,r)) - \nu_m r^m| \leq C_0 r^{m+2} \tag{8}$$

See ((3.3) and (3.2) of (Calder and Garcia-Trillos 19)).

2.1 Proof of Lemma 1

In the following manipulations we invoke first the Holder property of f, then the upper bound in 7, and finally Lemma 8 with $\delta = 2$ to get

$$\begin{aligned}
f^T \mathbf{L}_{n,r} f &= \sum_{i,j=1}^{n} \Big(f(X_i) - f(X_j)\Big)^2 K\bigg(\frac{\|X_i - X_j\|_{\mathbb{R}^d}}{r}\bigg) \\
&\leq Q^2 \sum_{i,j=1}^{n} \Big(d_{\mathcal{M}}(X_i, X_j)\Big)^2 K\bigg(\frac{\|X_i - X_j\|_{\mathbb{R}^d}}{r}\bigg) \\
&\leq 2Q^2 \sum_{i,j=1}^{n} \Big(\|X_i - X_j\|_{\mathbb{R}^d}^2 + \frac{64}{R^4}\|X_i - X_j\|_{\mathbb{R}^d}^6\Big) K\bigg(\frac{\|X_i - X_j\|_{\mathbb{R}^d}}{r}\bigg) \\
&\leq 2Q^2 \Big(r^2 + \frac{64}{R^4} r^6\Big) n \deg_{\max}(G_{n,r}) \\
&\leq 4Q^2 p_{\max} \Big(r^2 + \frac{64}{R^4} r^6\Big)\Big(\nu_m r^m + C_0 r^{m+2}\Big) n^2
\end{aligned}$$

with the last inequality holding with probability at least $1 - n\exp(-\frac{4}{3} n \nu_m r^m p_{\min})$. Lemma 1 then follows from appropriate choice of constants c_0 and C_1.

2.2 Proof of Lemma 2

Applying Theorem 2—which gives an estimate on the ∞-optimal transport distance between P and P_n—to Theorem 1—which gives a lower bound on $\lambda_k(G_{n,r})$ in terms of $\lambda_k(\mathcal{M})$ and $d_\infty(P, P_n)$— we have that if

$$2C_3 A \frac{\log(n)^{p_d}}{n^{1/d}} \leq r < \min\bigg\{\frac{C_2}{\sqrt{\lambda_k(\mathcal{M})}}, \frac{1}{8C_3 L(p)}, \frac{1}{8C_3\sqrt{\lambda_k(\mathcal{M})}}, \frac{1}{8C_3\sqrt{\mathcal{K}}}\bigg\}$$

then

$$\lambda_k(G_{n,r}) \geq \frac{1}{2} n r^{m+2} \lambda_k(\mathcal{M})$$

with probability at least $1 - C_4 n^{-\beta}$. Lemma 2 follows from replacing all instances of $\lambda_k(\mathcal{M})$ in the above expression with estimates given by Weyl's Law (Theorem 3).

2.3 Proof of Lemma 3

Note that

$$f^T \mathbf{L}_{n,r}^2 f = \sum_{i=1}^{n} \Big(\mathbf{L}_{n,r} f(X_i)\Big)^2$$

Lemma 3 follows directly from the pointwise estimate Lemma 7 upon taking a union bound over $x = X_1, \ldots, X_n$.

2.4 Proof of Lemma 4

We will consider $\mathbf{L}_{n,r}f(X_i)$ as an estimate of $\sigma_K \Delta_p f(X_i)$ (when appropriately scaled), and show that this deviation is small when r is sufficiently large. To begin, writing $\mathbf{L}_{n,r}f(X_i) = \sigma_K \Delta_p f(X_i) n r^{m+2} + \mathbf{L}_{n,r}f(X_i) - \sigma_K \Delta_p f(X_i) n r^{m+2}$ inside the definition of the 3rd-order graph Sobolev seminorm, we obtain

$$
\begin{aligned}
f^T \mathbf{L}_{n,r}^3 f &= \sum_{i,j=1}^{n} \Big(\mathbf{L}_{n,r}f(X_i) - \mathbf{L}_{n,r}f(X_j)\Big)^2 K_r(X_i,X_j) \\
&\leq 6 \sum_{i,j=1}^{n} \Big(\mathbf{L}_{n,r}f(X_i) - \sigma_K \Delta_p f(X_i) n r^{m+2}\Big)^2 K_r(X_i,X_j) + \\
&\quad 3\sigma_K^2 n^2 r^{2(m+2)} \sum_{i,j=1}^{n} \Big(\Delta_p f(X_i) - \Delta_p f(X_j)\Big)^2 K_r(X_i,X_j)
\end{aligned}
$$

Applying Theorem 4 with $\delta = r$, and Lemma 8 with $\delta = 1$, we have

$$
\begin{aligned}
\sum_{i,j=1}^{n} \Big(\mathbf{L}_{n,r}f(X_i) - \sigma_K \Delta_p f(X_i) n r^{m+2}\Big)^2 K_r(X_i,X_j) &\leq C_6^2 \big([f]_{1;\mathcal{M}} + \|f\|_{C^3(\mathcal{M})} + 1\big)^2 n^2 r^{2m+6} \sum_{i,j=1}^{n} K_r(X_i,X_j) \\
&\leq C_6^2 \big([f]_{1;\mathcal{M}} + \|f\|_{C^3(\mathcal{M})} + \|p\|_{C^2(\mathcal{M})}\big)^2 n^3 r^{2m+6} \deg_{\max}(G_{n,r}) \\
&\leq 4C_6^2 \big([f]_{1;\mathcal{M}} + \|f\|_{C^3(\mathcal{M})} + \|p\|_{C^2(\mathcal{M})}\big)^2 n^4 r^{2m+6} \Big[\nu_m r^m + C_0 r^{m+2}\Big] p_{\max} \\
&\leq C p_{\max} \big(Q + p_{\max}\big)^2 n^4 r^{3m+6},
\end{aligned}
$$

with probability at least $1 - 2n\exp(-c_2 n r^{4+m}) - n\exp(4n\nu_m r^m p_{\min}/3)$.

It remains to upper bound $(\Delta_p f)^T \mathbf{L}_{n,r}(\Delta_p f)$. Rewriting

$$
\Delta_p f = -\big(\nabla p^T \nabla f + p\Delta_{\mathcal{M}} f\big)
$$

we observe that since $f \in C^3(\mathcal{M}, Q)$ and $p \in C^2(\mathcal{M}, p_{\max})$, $\Delta_p f \in C^1(\mathcal{M}, 4Qp_{\max})$. By Lemma 1 we therefore have

$$
(\Delta_p f)^T \mathbf{L}_{n,r}(\Delta_p f) \leq \frac{16}{\delta} C_1 n^2 r^{m+2} Q^2 p_{\max}^4
$$

with probability at least $1 - \delta$. The claim of Lemma 4 then follows after an appropriate choice of constant in (6).

3 Additional Theory

3.1 Estimates on Graph Eigenvalues

Theorem 1 ((Part of) Theorem 4 of [Trillos et al., 2019])**.** *There exist constants C_2 and C_3 which depend only on $m, p_{\min}, p_{\max}, L(p)$ and K such that if*

$$
\sqrt{\lambda_k(\mathcal{M})}\, r < C_2,
$$

then,

$$
\lambda_k(G_{n,r}) \geq n r^{m+2} \lambda_k(\mathcal{M}) \left[1 - C_3\left(L(p) r + \frac{d_\infty(P,P_n)}{r} + \sqrt{\lambda_k(\mathcal{M})}\, r + \mathcal{K} r^2\right)\right]
$$

In order to make use of Theorem 1, we need an upper bound on $d_\infty(P, P_n)$. Theorem 2 provides a probabilistic estimate.

Theorem 2 (Theorem 2 of [Trillos et al., 2019])**.** *Let P satisfy (A3). Then for any $\beta > 1$ and every $n \in \mathbb{N}$, there exists a transport map $T_n : \mathcal{M} \to \mathbf{X}$ and a constant A such that*

$$\sup_{x\in\mathcal{M}} d_\infty(x, T_n(x)) \leq A \cdot \begin{cases} \frac{(\log(n))^{3/4}}{n^{1/2}}, & \text{if } m = 2, \\ \left(\frac{\log(n)}{n}\right)^{1/m}, & \text{if } m \geq 3, \end{cases}$$

with probability at least $1 - C_4 n^{-\beta}$. The constant A depends only on $\mathcal{K}, i_0, m, \mu(\mathcal{M}), \alpha$ and β, and the constant C_4 depends only on $\mathcal{K}, i_0, m, \mu(\mathcal{M})$.

3.2 Weyl's Law

Although Weyl's Law is traditionally stated with respect to the unweighted Laplace-Beltrami operator, the same asymptotics apply to Δ_p when p is bounded away from 0 and ∞ on $\mathcal{M}$. Let $N(\lambda)$ count the number of eigenvalues of $\mathcal{M}$ which are less than λ.

Theorem 3 (Weyl's Law)**.** *Assume (A3). If $\mathcal{M}$ is a compact connected oriented manifold then*

$$N(\lambda) \sim \lambda^{m/2}$$

or, equivalently,

$$\lambda_k(\mathcal{M}) \sim k^{2/m}.$$

3.3 Pointwise estimates on graph Laplacians

We give some pointwise estimates on $\mathbf{L}_{n,r} f(x)$, many of which are reproduced from [Calder and Trillos, 2019]. Define

$$\mathbf{L}_{P,r} f(x) := \frac{1}{r^{m+2}} \int_{B(x,r)} K_r(y,x)\Big(f(x) - f(y)\Big) p(y)\, d\mu(y)$$

to be a nonlocal Laplacian operator, which acts as an intermediary between $\mathbf{L}_{n,r}$ and Δ_p.

Theorem 4 (Theorem 3.3 of [Calder and Trillos, 2019])**.** *Assume (A1)-(A3). Let $f \in C^3(\mathcal{M})$ and $p \in C^2(\mathcal{M})$. There exist constants C_6 and c_2 such that for any $r \leq \delta \leq r^{-1}$,*

$$\max_{1\leq i\leq n} \left|\mathbf{L}_{n,r} f(X_i) - \sigma_K \Delta_p f(X_i) n r^{m+2}\right| \leq C_6\big([f]_{1;\mathcal{M}} + \|f\|_{C^3(\mathcal{M})} + \|p\|_{C^2(\mathcal{M})}\big)\delta n r^{m+2}$$

with probability at least $1 - 2n\exp(-c\delta^2 n r^{m+2})$.

Lemma 6 (Lemma 3.4 of [Calder and Trillos, 2019])**.** *Assume (A1)-(A3). Let $f \in C^1(\mathcal{M})$. Then for $x \in \mathcal{M}$ and $r \leq \delta \leq r^{-1}$, there exists constants C_6 and c_2 which depend only on $m, \mathcal{K}, R$ and i_0 such that*

$$\mathbb{P}\left[\left|\frac{1}{nr^{m+2}} \mathbf{L}_{n,r} f(x) - \mathbf{L}_{P,r} f(x)\right| \geq C_6 [f]_{1,B(x,2r)} \delta\right] \leq 2\exp(-c_2 \delta^2 n r^{m+2})$$

Most analysis on $\mathbf{L}_{n,r} f(x)$ provide pointwise estimates assuming $f \in C^{q+2}(\mathcal{M})$ for some $q > 0$. We shall instead assume only $f \in C^2(\mathcal{M})$, and content ourselves with merely an upper bound on $|\mathbf{L}_{n,r} f(x)|$.

Lemma 7. *Assume (A1)-(A3). Let $f \in C^2(\mathcal{M})$. Then for $x \in \mathcal{M}$ and $r \leq \delta \leq r^{-1}$, there exists constants C_7 and c_2 which depend only on $m, \mathcal{K}, R$ and i_0 such that*

$$\left|\mathbf{L}_{n,r} f(x)\right| \leq C_7 n r^{m+2}\Big([f]_{1;B(x,2r)}\delta + p_{\max}\nu_d [f]_{1,B(x,r)} r + \|\nabla f(x)\| + \|f\|_{C^2(B(x,r))}\Big)$$

with probability at least $1 - 2\exp(-c_2\delta^2 n r^{m+2})$.

Proof. We shall follow the proof of Theorem 3.3 of [Calder and Trillos, 2019], deviating when necessary to fit our differing assumptions and ultimate goal. Let

$$\mathbf{L}^{(i)}_{P,r} f(x) = \frac{1}{r^{m+2}} \int_{B(x,r)} K\left(\frac{d_{\mathcal{M}}(y,x)}{r}\right)\big(f(x) - f(y)\big)p(y)\, d\mu(y)$$

be an intrinsic analogue to $\mathbf{L}_{P,r}$. It follows from (7) and (A2) that

$$\begin{aligned}\left|\mathbf{L}^{(i)}_{P,r} f(x) - \mathbf{L}_{P,r} f(x)\right| &\le \frac{1}{r^{2+m}} \int_{B(x,r)} \left|K\Big(\frac{d_{\mathcal{M}}(x,y)}{r}\Big) - K\Big(\frac{d_{\mathcal{M}}(x,y)}{r}\Big)\right| \big|f(x) - f(y)\big| p(y)\, d\mu(y) \\ &\le \frac{p_{\max}}{r^m} \int_{B(x,r)} \big|f(x) - f(y)\big|\, d\mu(y) \qquad (9) \\ &\le p_{\max} \nu_d [f]_{1,B(x,r)} r \qquad (10)\end{aligned}$$

By assumption $r < i_0$, and therefore the exponential map $\exp_x : B(0,r) \subset \mathcal{T}_x(\mathcal{M}) \to \mathcal{M} \rfloor$ is a diffeomorphism. For $v \in B(0,r)$ let $w(v) = f(\exp_x(v))$ and $\rho(v) = p(\exp_x(v))$, i.e. express f and p in terms of normal Riemmanian coordinates, and let $J_x(v)$ be the Jacobian of $\exp_x$ at v. Then

$$\begin{aligned}\mathbf{L}^{(i)}_{P,r} f(x) &= -\frac{1}{r^{m+2}} \int_{B(0,r)} K\left(\frac{\|v\|}{r}\right)\big(w(v) - w(0)\big)\rho(v) J_x(v)\, d\mu(v) \\ &= -\frac{1}{r^2} \int_{B(0,1)} K(\|v\|)\big(w(rv) - w(0)\big)\rho(rv) J_x(rv)\, d\mu(v).\end{aligned}$$

Now we plug in the Taylor expansions $w(rv) = w(0) + \nabla w(0)\cdot v + O(\|w\|_{C^2(B(0,r))} r^2)$, $\rho(rv) = \rho(0) + O(r)$ and $J_x(rv) = 1 + O(r^2)$ to obtain

$$\begin{aligned}\mathbf{L}^{(i)}_{P,r} f(x) &= -\frac{1}{r^2} \int_{B(0,1)} K(\|v\|)\Big(r\nabla w(0)^T v + O\big(\|w\|_{C^2(B(0,r))} r^2\big)\Big)\Big(\rho(0) + O(r)\Big)\Big(1 + O(r^2)\Big)\, d\mu(v) \\ &= O\Big(\|\nabla w(0)\| + \|w\|_{C^2(B(0,r))}\Big) \qquad (11)\end{aligned}$$

The claim then follows by combining (10), (11) and Lemma 6. ☐

3.4 Degree bounds

Lemma 8 follows from the multiplicative form of Hoeffding's inequality along with (8) and (A2).

Lemma 8. *Assume (A1)-(A3). Then*

$$\deg_{\max}(G_{n,r}) \le (1+\delta)\Big[\nu_m r^m + C_0 r^{m+2}\Big] p_{\max} n$$

with probability at least $1 - n\exp(-\delta^2 n \nu_m r^m p_{\min}/3)$.

3.5 Explaining what is difficult when $s = 4$.

Let me start by giving a rough summary of the difficulty, and then go into more detail .We would like to show

$$f^T \mathbf{L}^4_{n,r} f = \sum_{i=1}^n \Big(\mathbf{L}^2_{n,r} f(X_i)\Big)^2 \lesssim n^5 r^{8(m+2)} \qquad (12)$$

for $f \in C^4(\mathcal{M})$. Note that

$$\mathbf{L}_{P,r} f(x) = \int_{B(x,r)\cap\mathcal{M}} \big(f(x) - f(y)\big)\, dP(y)$$

is (up to scaling by a factor of nr^{m+2})the expectation of $\mathbf{L}_{n,r} f(x)$. It is possible to show a sufficiently tight bound on $|\mathbf{L}_{n,r}^2 f(x) - nr^{2+m}\mathbf{L}_{P,r}^2 f(x)|$ when $r \gg n^{-1/(6+m)}$, thus handling the variance. To handle the bias, we can show that

$$\mathbf{L}_{P,r} f(x) = nr^{m+2}\sigma_K \Delta_p f(x) + \mathcal{I}f(x)$$

where Δ_p is the weighted Laplace-Beltrami operator on $\mathcal{M}$, and $\mathcal{I}f$ is an error term satisfying $\mathcal{I}f(x) = O(r^{m+3})$. This bound on the error term is sufficient for $s = 1$ up to $s = 3$, however, when $s \geq 4$ it is insufficient (check by plugging back in to (12)). In the Euclidean case, we had that $\mathcal{I}f(x)$ itself belong to C^1 and thus could argue $(\mathbf{L}_{P,r}\mathcal{I}f)(x) \ll \mathcal{I}f(x)$. In the manifold case, I no longer know how to do this.

The longer version: we can make the following progress towards upper bound $f^T \mathbf{L}_{n,r}^4 f$:

$$\begin{aligned}
f^T \mathbf{L}_{n,r}^4 f &\leq n \max_{1 \leq i \leq n} \left|\mathbf{L}_{n,r}^2 f(X_i)\right|^2 \\
&\leq n \max_{1 \leq i \leq n} \left|\mathbf{L}_{n,r}(\mathbf{L}_{n,r} - nr^{m+2}\mathbf{L}_{P,r}) f(X_i)\right|^2 + n^2 r^{m+2} \max_{1 \leq i \leq n} \left|(\mathbf{L}_{n,r} - nr^{m+2}\mathbf{L}_{P,r})\mathbf{L}_{P,r} f(X_i)\right|^2 + \\
&\quad n^4 r^{4m+8} \max_{1 \leq i \leq n} \left|\mathbf{L}_{P,r}^2 f(X_i)\right|
\end{aligned}$$

We can upper bound the first term as follows:

$$\begin{aligned}
\left|\mathbf{L}_{n,r}(\mathbf{L}_{n,r} - nr^{m+2}\mathbf{L}_{P,r}) f(x)\right|^2 &\leq \max_{1 \leq j \leq n} \left|(\mathbf{L}_{n,r} - nr^{m+2}\mathbf{L}_{P,r}) f(X_j)\right| \cdot \sum_{j=1}^{n} K\Big(\frac{\|x - X_j\|_{\mathbb{R}^d}}{r}\Big) \\
&\leq Cnr^m \max_{1 \leq j \leq n} \left|(\mathbf{L}_{n,r} - nr^{m+2}\mathbf{L}_{P,r}) f(X_j)\right| \\
&\leq Cn^2 r^{2m+2}\delta
\end{aligned}$$

where the last inequality follows from Bernstein's inequality and holds with probability at least on the order of $1 - \exp(-nr^{2+m}\delta^2)$. Choosing $\delta \asymp r^2$ gives the desired rate, assuming $n \gg r^{6+m}$.

We can use a similar argument to upper bound the second term.

The third term—$\mathbf{L}_{P,r}^2 f(X_i)$—is what poses the challenge. We want to convert the integral on $\mathcal{M}$ to an integral on a Euclidean space using exponential maps. Letting $\exp_x : T_x(\mathcal{M}) \to \mathcal{M}$ be the exponential map, letting $\widetilde{B}$ satisfy $\exp_x(\widetilde{B}) = B(x,r) \cap \mathcal{M}$, associating the tangent plane $T_x(\mathcal{M})$ with $\mathbb{R}^m$, and writing $g_x(v) = \|x - \exp_x(v)\|_{\mathbb{R}^d} - \|v\|_{\mathbb{R}^m}$ we have

$$\begin{aligned}
\mathbf{L}_{P,r} f(x) &= \int_{\mathcal{M}} (f(x) - f(y)) K\Big(\frac{\|y - x\|_{\mathbb{R}^d}}{r}\Big) p(y)\, d\mu(y) \\
&= \int_{\widetilde{B}} (w(0) - w(v)) K\Big(\frac{\|x - \exp_x(v)\|_{\mathbb{R}^d}}{r}\Big) \rho(v) J_v(x)\, dv \\
&= \int_{\widetilde{B}} (w(0) - w(v)) K\Big(\frac{\|v\|_{\mathbb{R}^m} + g_x(v)}{r}\Big) \rho(v) J_v(x)\, dv \\
&= \int_{B(0,r) \subset \mathbb{R}^m} (w(0) - w(v)) K\Big(\frac{\|v\|_{\mathbb{R}^m}}{r}\Big) \rho(v) J_v(x)\, dv + \\
&\quad \underbrace{\int_{\widetilde{B}} (w(0) - w(v)) \Big[K\Big(\frac{\|v\|_{\mathbb{R}^m} + g_x(v)}{r}\Big) - K\Big(\frac{\|v\|_{\mathbb{R}^m}}{r}\Big)\Big] \rho(v) J_v(x)\, dv}_{\mathcal{I}(x)}
\end{aligned}$$

The first of the two integrals is now an integral over a ball in a Euclidean domain, and is not a problem. The second integral is tricky. The only way I know how to analyze $\mathcal{I}(x)$ is to use the following facts: first, $|w(v) - w(0)| \lesssim r$, second that $g_x(v) \lesssim r^3$ for all $\exp_x(v) \in B(0,r) \cap \mathcal{M}$, third that

$|\mu(B(x,r) \cap \mathcal{M}) - \nu_m r^m| \lesssim r^{m+2}$, and fourth that $|J_x(v)| \lesssim 1 + r^2$. Together, these give

$$\int_{\widetilde{B}} (w(0) - w(v)) \Big[K\Big(\frac{\|v\|_{\mathbb{R}^m} + g_x(v)}{r}\Big) - K\Big(\frac{\|v\|_{\mathbb{R}^m}}{r}\Big) \Big] \rho(v) J_v(x) \, dv \lesssim r \Big(\mu\big(B(x, r + r^3) \cap \mathcal{M}\big) - \mu\big(B(x,r) \cap \mathcal{M}\big) \Big) \lesssim r^{m+3}.$$

But this only gives $\mathbf{L}_{P,r}^2 f(x) \lesssim r^{2m+3}$, which is not nearly enough. The problem is that I am not using any smoothness properties of $\mathcal{I}(x)$ in my upper bound on $\mathbf{L}_{P,r}\mathcal{I}f(x)$. I suspect it may have such properties, related to the smoothness assumptions we place on $\mathcal{M}$, but proving it seems non-trivial.

4 Notation

- For any $x, y \in \mathcal{M}$ we write $d_{\mathcal{M}}(x,y)$ for the geodesic distance between x and y, that is the length of the shortest path connecting x and y.
- For measures P and Q on $\mathcal{M}$, we write $d_\infty(P,Q)$ for the ∞-optimal transport distance between P and Q. Formally,
$$d_\infty(P,Q) = \inf \Big\{ \operatorname{esssup}_\gamma \Big\{ |x-y| : x, y \in \mathcal{M} \Big\} : \gamma \in \Gamma(P,Q) \Big\}$$
where $\Gamma(P,Q)$ is the set of joint distributions on $\mathcal{M} \times \mathcal{M}$ with marginals P and Q.

REFERENCES

Dmitri Burago, Sergei Ivanov, and Yaroslav Kurylev. A graph discretization of the laplace-beltrami operator. *Journal of Spectral Theory*, pages 675–714, 2014.

Jeff Calder and Nicolas Garcia Trillos. Improved spectral convergence rates for graph laplacians on epsilon-graphs and k-nn graphs. *arXiv preprint arXiv:1910.13476*, 2019.

Nicolás García Trillos, Moritz Gerlach, Matthias Hein, and Dejan Slepčev. Error estimates for spectral convergence of the graph laplacian on random geometric graphs toward the laplace–beltrami operator. *Foundations of Computational Mathematics*, pages 1–61, 2019.