Error Estimates for Spectral Convergence of the Graph Laplacian on Random Geometric Graphs Toward the Laplace–Beltrami Operator

Nicolás García Trillos[1] · Moritz Gerlach[2] · Matthias Hein[3] · Dejan Slepčev[4]

Abstract

We study the convergence of the graph Laplacian of a random geometric graph generated by an i.i.d. sample from a m-dimensional submanifold $\mathcal{M}$ in $\mathbb{R}^d$ as the sample size n increases and the neighborhood size h tends to zero. We show that eigenvalues and eigenvectors of the graph Laplacian converge with a rate of $O\left(\left(\frac{\log n}{n}\right)^{\frac{1}{2m}}\right)$ to the eigenvalues and eigenfunctions of the weighted Laplace–Beltrami operator of $\mathcal{M}$. No information on the submanifold $\mathcal{M}$ is needed in the construction of the graph or the "out-of-sample extension" of the eigenvectors. Of independent interest is a generalization of the rate of convergence of empirical measures on submanifolds in $\mathbb{R}^d$ in infinity transportation distance.

Dedicated to the memory of Yaroslav Kurylev.

Communicated by Alan Edeiman.

✉ Dejan Slepčev
  slepcev@math.cmu.edu

  Nicolás García Trillos
  garciatrillo@wisc.edu

  Moritz Gerlach
  gerlach@cs.uni-saarland.de

  Matthias Hein
  matthias.hein@uni-tuebingen.de

[1] Department of Statistics, University of Wisconsin-Madison, Madison, USA

[2] Department of Mathematics and Computer Science, Saarland University, Saarbrücken, Germany

[3] Department of Computer Science, University of Tübingen, Tübingen, Germany

[4] Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, USA

🖄 Springer

Mathematics Subject Classification 62G20 · 65N25 · 60D05 · 58J50 · 68R10 · 05C50

List of Symbols

$\mathcal{M}$	Compact manifold without boundary embedded in $\mathbb{R}^d$. Riemann metric on $\mathcal{M}$ is the one inherited from $\mathbb{R}^d$
m	The dimension of $\mathcal{M}$
$\mathrm{Vol}(A)$	The volume of $A \subset \mathcal{M}$ according to Riemann volume form
$d(x, y)$	The geodesic distance between points $x, y \in \mathcal{M}$
$B_{\mathcal{M}}(x, r)$	Ball in $\mathcal{M}$ with respect to geodesic distance on $\mathcal{M}$
$B(r)$	Ball in $\mathbb{R}^d$ of radius r, centered at the origin
μ	Probability measure supported on $\mathcal{M}$ that describes the data distribution
p	Density of μ with respect to volume form on $\mathcal{M}$
ρ	Density of the weight measure (which allows us to consider the normalized graph Laplacian) with respect to μ
α	Constant describing the bounds on the densities p and ρ, see (1.2) and (1.9)
X	Point cloud $X = \{x_1, \ldots, x_n\} \subset \mathcal{M}$ drawn from distribution μ. Also considered as the set of vertices of the associated graph
μ_n	Empirical measure of the sample X
$\mathbf{m}$	The vector giving the values of the discrete weights used in various forms of graph Laplacian, see Sects. 1.2.1 and 1.2.2
$w_{i,j}$	Edge weight between vertices x_i and x_j
δu	Differential of function $u : X \to \mathbb{R}$. It maps edges to $\mathbb{R}$ and is defined by $\delta u_{i,j} = u(x_j) - u(x_i)$
i_0	Injectivity radius of $\mathcal{M}$. The injectivity radius at a point $p \in \mathcal{M}$ is the largest radius of a ball for which the exponential map at p is a diffeomorphism. The injectivity radius i_0 is the infimum of the injectivity radii at all points of $\mathcal{M}$
K	Maximum of the absolute value of sectional curvature of $\mathcal{M}$
R	Reach of $\mathcal{M}$, defined in (1.37)
η	Nonnegative function setting the edge weights as a function of the distance between the vertices, see (1.5)
h	Length scale such that weight between vertices is large if their distance is comparable to or less than h
σ_η	Is the kernel-dependent scaling factor relating the graph Laplacian and the continuum Laplacian; defined in (1.4)
ω_m	The volume of unit ball in $\mathbb{R}^m$
$d_\infty(\mu, \nu)$	Infinity transportation distance between measures μ, ν
ε	Upper bounds on the transportation distance between μ and μ_n
L	Lipschitz constant of various functions: p, ρ and η
P	Discretization operator defined in (1.24)
P^*	Is the adjoint of P if $\rho \equiv 1$ and an approximate adjoint otherwise
I	Interpolation operator defined in (1.24)

1 Introduction

Given an i.i.d. sample $X = \{x_1, \ldots, x_n\}$ from the data generating measure μ in Euclidean space $\mathbb{R}^d$, the goal of most tasks in machine learning and statistics is to infer properties of μ. A particularly interesting case is when μ has support on a m-dimensional compact submanifold $\mathcal{M}$ in $\mathbb{R}^d$, e.g., due to strong dependencies between the individual features. In this case, one can construct a neighborhood graph on the sample by connecting all vertices of Euclidean distance less than a certain length scale h, and in this way produce a discrete approximation of the unknown manifold $\mathcal{M}$. Laplacian Eigenmaps [2] and Diffusion Maps [8] have been proposed as tools to extract intrinsic structure of the manifold by considering the eigenvectors of the resulting unnormalized resp. normalized *graph Laplacian*; in particular, Laplacian eigenmaps are used in the first step of spectral clustering [29], one of the most popular graph-based clustering methods. In general, it is well known that the spectrum of the graph Laplacian resp. Laplace–Beltrami operator captures important structural resp. geometric properties of the graph [17] resp. manifold [7].

In this paper, we examine this question: Under what conditions, and at what rate, does the spectrum of the graph Laplacian built from i.i.d. samples on a submanifold converge to the spectrum of the (weighted) Laplace–Beltrami operator of the submanifold as the sample size $n \to \infty$ and the neighborhood radius $h \to 0$?

Graph-based approximations to the Laplace–Beltrami operator have been studied by several authors and in a variety of settings. The pointwise convergence of the graph Laplacian toward the Laplace–Beltrami operator has been proven in [4,13–15,25,28]. The spectral convergence of the graph Laplacian for fixed neighborhood size h for Euclidean domains has been established in [21,30], and error rates have been given. The spectral convergence of the graph Laplacian toward the Laplace–Beltrami operator for the uniform distribution has been discussed, for Gaussian weights, in [3], and in [26] for the more general connection Laplacian (as well as the usual graph Laplacian), without precise information on allowable scaling of neighborhood radius, h and without convergence rates. In [12], the authors establish the conditions on graph connectivity for the spectral convergence on domains in $\mathbb{R}^m$. In particular, they prove convergence when $h \to 0$ as $n \to \infty$ and

$$h \gg \frac{(\log n)^{p_m}}{n^{\frac{1}{m}}},$$

where

$$p_m = \begin{cases} \frac{3}{4} & \text{if } m = 2 \\ \frac{1}{m} & \text{if } m \geq 3. \end{cases} \tag{1.1}$$

However, no error estimates were established.

The preprint [23] establishes (in Theorem 1.1) the spectral convergence of graph Laplacians constructed from data sampled from a submanifold in $\mathbb{R}^d$ with a convergence rate of $O\left(\left(\frac{\log n}{n}\right)^{\frac{1}{4m+14}}\right)$, where m is the intrinsic dimension of the submanifold.

In this paper, we propose a general framework to analyze the rates of spectral convergence for a large family of graph Laplacians. This framework in particular allows us to improve the results in [23] and establish a convergence rate of $O\left(\left(\frac{\log n}{n}\right)^{\frac{1}{2m}}\right)$ which is a significant improvement, in particular for small dimensions m. These convergence rates hold for different reweighing schemes of the graph Laplacian found in the literature including the unnormalized Laplacian, normalized Laplacian, and the random walk Laplacian. When the intrinsic dimension of the submanifold $\mathcal{M}$ is small, our results show, to some extent, why Laplacian eigenmaps can effectively extract geometric information from the data set, even though the number of features d may be high. Moreover, similar to [12], we show that the conditions in (1.1) are sufficient for spectral convergence. This is essentially the same condition required to ensure that the constructed graph is almost surely connected [20] and thus is close to optimal. It is interesting to note that for pointwise consistency of the graph Laplacian [13,15], the required stronger condition is $\frac{nh^{m+2}}{\log n} \to \infty$.

Our framework is rather different from those in [3,23] and builds on two main ideas. First, it builds on an extension of the recent result of Burago, Ivanov und Kurylev [6], see also [10], which shows in a non-probabilistic setting how one can approximate eigenvalues and eigenfunctions of the Laplace–Beltrami operator using the eigenvalues/eigenvectors of the graph Laplacian associated with an ϵ-net of the submanifold. As in our setting the manifold $\mathcal{M}$ is unknown, we generalize the result of [6] by using a graph construction which requires no knowledge about the submanifold $\mathcal{M}$ but which achieves the same approximation guarantees for the eigenvalues. In addition, we introduce a new out-of-sample extension of the eigenvectors for the approximation of the eigenfunctions which requires no information about the submanifold without significant loss in the convergence rate compared to the corresponding construction used in [6]. Our second main result generalizes the recent work of García Trillos and Slepčev [11] to the setting of empirical measures on submanifolds $\mathcal{M} \subset \mathbb{R}^d$ and establishes their rate of convergence in ∞-optimal transportation (OT) distance; the ∞-OT distance between the empirical measure associated with a point cloud and the volume form of the submanifold can be seen to be closely related to the notion of ϵ-net used in [6]. These estimates encompass all the probabilistic computations that we need to obtain our main results, and in particular, when combined with our deterministic computations, provide all the probabilistic estimates that quantify the rate of convergence of the spectrum of graph Laplacians constructed from randomly generated data toward the spectrum of a (weighted) Laplace–Beltrami operator on $\mathcal{M}$. We believe that both the generalization of [6] as well as the generalization of [11] are of independent interest. The combination of these two ideas and a number of careful estimates lead to our main results.

In what follows, we make the setting that we consider in the sequel precise, as well as define precisely the different graph Laplacians and their continuous counterparts.

1.1 Graph Construction

Let $\mathcal{M}$ be a compact connected m-dimensional Riemannian manifold without boundary, embedded in $\mathbb{R}^d$, with $m \geq 2$. We assume that the absolute value of sectional

curvature is bounded by K, the injectivity radius is i_0 and with reach R. We write $d(x, y)$ for the distance between x and y on the manifold and $|x - y|$ for the Euclidean distance in $\mathbb{R}^d$.

Let μ be a probability measure on $\mathcal{M}$ that has a non-vanishing Lipschitz continuous density p with respect to the Riemannian volume on $\mathcal{M}$ with Lipschitz constant L_p. Compactness of $\mathcal{M}$ and continuity of p guarantee the existence of a constant $\alpha \geq 1$ such that

$$\frac{1}{\alpha} \leq p(x) \leq \alpha \quad \text{for all } x \in \mathcal{M}. \tag{1.2}$$

We let $x_1, x_2, \ldots, x_n, \ldots$ be a sequence of i.i.d. samples from μ. In order to leverage the geometry of $\mathcal{M}$ from the data, we build a graph with vertex set $X := \{x_1, \ldots, x_n\}$. In the simplest setting, for each $n \in \mathbb{N}$ we choose a neighborhood parameter $h = h_n$ and we put an edge from x_i to x_j and from x_j to x_i (and write $x_i \sim x_j$) provided that $|x_i - x_j| \leq h$; we let $E = \{(i, j) \in \{1, \ldots, n\}^2 : x_i \sim x_j\}$ be the set of such edges. More generally, we consider weighted graphs, with weights that depend on the distance between the vertices connected by them. For that purpose, let us consider a decreasing function $\eta \colon [0, \infty) \to [0, \infty)$ with support on the interval $[0, 1]$ such that the restriction of η to $[0, 1]$ is Lipschitz continuous. Normalizing η if needed allows us to assume from here on that

$$\int_{\mathbb{R}^m} \eta(|x|)\mathrm{d}x = 1. \tag{1.3}$$

For convenience, we assume that $\eta(1/2) > 0$. We denote by

$$\sigma_\eta := \int_{\mathbb{R}^m} |y_1|^2 \, \eta(|y|)\mathrm{d}y, \tag{1.4}$$

the *surface tension* of η, where y_1 represents the first coordinate of the vector $y \in \mathbb{R}^m$. To every given edge $(i, j) \in E$, we assign the weight $w_{i,j}$, where

$$w_{i,j} = \frac{1}{nh^m}\eta\left(\frac{|x_i - x_j|}{h}\right) \tag{1.5}$$

and we consider the weighted graph (X, w) with $w_{i,j}$ as in (1.5) for every (i, j). In fact, note that if the points x_i, x_j are not connected by an edge in E then $w_{i,j} = 0$.

Remark 1 The function η can be chosen as $c\mathbb{1}_{[0,1]}$ as well as a smooth function like

$$\eta(t) := c \begin{cases} \exp\left(\frac{1}{t-1}\right) & 0 \leq t < 1 \\ 0 & t \geq 1, \end{cases}$$

where c is the appropriate constant ensuring normalization or simply a truncated version of a Gaussian. Also, we note that for $\eta = \frac{1}{\omega_m}\mathbb{1}_{[0,1]}$ it follows from [6, (2.7)]

that $\sigma_\eta = \frac{1}{m+2}$, where ω_m is the volume of the unit ball in $\mathbb{R}^m$. While the definition of the weights is up to the constant σ_η and a slightly different rescaling in terms of h is similar to [6], the main difference is that we use the Euclidean metric of the ambient space $\mathbb{R}^d$ in (1.5), whereas in [6] neighborhoods are throughout defined in terms of the geodesic distance. Here, we are forced to use the metric from the ambient space as the manifold $\mathcal{M}$ is in general assumed to be unknown.

Remark 2 We have assumed that $\eta \colon [0, 1] \to \mathbb{R}$ is decreasing and that $\eta(1/2) > 0$, which would imply that $\eta(0) > 0$. Nevertheless, we remark that none of the results presented in this paper change if we modify the value of $\eta(0)$. In particular, we allow for $\eta(0) = 0$ if desired and we can simply assume that η is decreasing and Lipschitz in $(0, 1)$ (then the condition $\eta(0) > 0$ changes to $\eta(0+) > 0$). This observation is relevant in order to allow for graphs where vertices have no edges with themselves.

Remark 3 The requirement that η is compactly supported is purely a technical one. It is in principle possible to carry out the arguments of this work for noncompact kernels, like the Gaussian one. However, that would require obtaining error bounds on extra terms and would make the already involved estimates even more complicated.

1.2 Dirichlet Forms and Laplacians

In this section, we introduce the Laplacians in both discrete and continuous settings.

We use the graph structure defined in the previous section to define a Dirichlet form in the discrete setting. First, the weights $w_{i,j}$ serve as a measure on the set E and thus induce a scalar product of functions $F, G \colon E \to \mathbb{R}$ given by

$$\langle F, G \rangle := \langle F, G \rangle_{L^2(E,w)} := \frac{1}{n\sigma_\eta} \sum_{(i,j) \in E} w_{i,j} F(i, j) G(i, j).$$

Second, for functions $u, v \colon X \to \mathbb{R}$ on the vertices, we define the *discrete differential*

$$(\delta u)(i, j) := \frac{1}{h}\big(u(x_j) - u(x_i)\big) \quad \text{for } (i, j) \in E. \tag{1.6}$$

We can then define the *discrete Dirichlet form* between $u, v \colon X \to \mathbb{R}$ as

$$b(u, v) := \langle \delta u, \delta v \rangle_{L^2(E,w)}. \tag{1.7}$$

In the continuous setting, on the domain $V := H^1(\mathcal{M}, \mu)$ (the Sobolev space of functions in $L^2(\mathcal{M}, \mu)$ with distributional derivative in $L^2(\mathcal{M}, \mu)$) we define the Dirichlet form $D \colon V \times V \to \mathbb{R}$ as

$$D(f, g) := \int_{\mathcal{M}} \langle \nabla f, \nabla g \rangle_x p^2(x) \mathrm{dVol}(x), \tag{1.8}$$

where Vol stands for the Riemannian volume form of $\mathcal{M}$, ∇f and ∇g are the gradients of f and g and $\langle \cdot, \cdot \rangle$ represents the Riemannian metric induced on $\mathcal{M}$. Since p is

bounded from above, this symmetric bilinear form is continuous, i.e., $|D(f, g)| \leq C' \|f\|_V \|g\|_V$ for a suitable constant $C' > 0$ and all $f, g \in V$. For the remainder, we use $b(u)$ and $D(f)$ as shorthand for $b(u, u)$ and $D(f, f)$, respectively.

Next, we choose measures on X and on the manifold $\mathcal{M}$ and define corresponding operators associated with the forms b and D on $L^2(X)$ and $L^2(\mathcal{M})$, respectively. The idea is that by modifying the inner product in $L^2(X)$ and in $L^2(\mathcal{M})$, we obtain different realizations of Laplacian operators. The so-called unnormalized and random walk graph Laplacian (see definitions below), as well as their continuous counterparts, are instances of the general framework that we consider. Let μ_n be the empirical measure of the random sample, i.e.,

$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}.$$

On X, we consider the measure μ_n endowed with a density $\mathbf{m} = (m_1, \ldots, m_n)$, denoted by $\mathbf{m}\mu_n$. On the other hand, on $\mathcal{M}$, we consider the measure $\rho\mu$, where ρ is a Lipschitz continuous density with Lipschitz constant L_ρ with respect to μ satisfying

$$\frac{1}{\alpha} \leq \rho(x) \leq \alpha \quad \text{for all } x \in \mathcal{M}. \tag{1.9}$$

On the graph $\Gamma = \Gamma(X, \mathbf{m}\mu_n, E, w)$, we define the associated *weighted graph Laplacian* Δ_Γ as $\delta^*\delta$, i.e., as the unique operator satisfying

$$\langle \Delta_\Gamma u, v \rangle_{L^2(X, \mathbf{m}\mu_n)} = \langle \delta u, \delta v \rangle_{L^2(E, w)}$$

for all $u, v \in L^2(X)$.

At the continuum level, we define a weighted Laplacian associated with the form D and the measure $\rho\mu$ as follows. On the domain

$$\text{Dom}(\Delta) := \left\{ f \in V : \exists h \in L^2(\mathcal{M}, \rho\mu) \text{ s. t. } D(f, g) = \langle h, g \rangle_{L^2(\mathcal{M}, \rho\mu)} \ \forall g \in V \right\}$$

we set $\Delta f := h$. The operator Δ is formally defined as

$$\Delta f = -\frac{1}{\rho p} \operatorname{div}(p^2 \nabla f),$$

where div stands for the divergence operator on $\mathcal{M}$.

One of the main results of this paper is that the spectrum of Δ_Γ approximates well that of Δ. Intuitively, one of the elements needed for this to be true is that the measure $\mathbf{m}\mu_n$ approximates $\rho\mu$ as $n \to \infty$. We use

$$\|\mathbf{m} - \rho\|_\infty := \max_{i=1,\ldots,n} |m_i - \rho(x_i)| \tag{1.10}$$

to quantify this approximation.

We now describe particular forms of the graph Laplacian frequently used in the machine learning literature.

1.2.1 Unnormalized Graph Laplacian

To obtain the unnormalized graph Laplacian, we choose the density vector $\mathbf{m}$ as $(1, 1, \dots)$. Then, Δ_Γ is explicitly given by

$$(\Delta_\Gamma u)(x_i) = \frac{2}{\sigma_\eta h^2} \sum_{j:i\sim j} w_{i,j}(u(x_i) - u(x_j))$$

for all $x_i \in X$, which is, up to the factor $\frac{2}{\sigma_\eta h^2}$, known as the unnormalized graph Laplacian. In this case, $\rho \equiv 1$, since ρ is the limit of $\mathbf{m}$ as $n \to \infty$. This results in a realization of the Laplacian on $L^2(\mathcal{M}, \rho\mu)$ that satisfies

$$\int_{\mathcal{M}} \Delta f g p(x)\mathrm{d}x = \int_{\mathcal{M}} \langle \nabla f, \nabla g \rangle_x p^2(x)\mathrm{d}x = D(f, g)$$

for all $f, g \in \mathrm{Dom}(\Delta)$. In case $p \in C^1(\mathcal{M})$, this operator Δ coincides with

$$\Delta f = -p \cdot \Delta_2 f = -\frac{1}{p}\,\mathrm{div}(p^2 \nabla f)$$

from Definition 8 of [15], where it was identified as the pointwise limit of the unnormalized graph Laplacian.

1.2.2 Random Walk Graph Laplacian

In order to obtain the random walk graph Laplacian, we choose the density vector $\mathbf{m}$ as the vertex degrees, i.e.,

$$m_i := \sum_{j=1}^n w_{ij} = \frac{1}{nh^m} \sum_{j=1}^n \eta\left(\frac{|x_i - x_j|}{h}\right) \quad \text{for } i \in \{1, \dots, n\} \tag{1.11}$$

and $\rho(x) = p(x)$ for all $x \in \mathcal{M}$. Then, Δ_Γ is given by

$$(\Delta_\Gamma u)(x_i) = \frac{2}{\sigma_\eta h^2} \sum_{j:i\sim j} \frac{w_{i,j}}{m_i}\big(u(x_i) - u(x_j)\big)$$

for all $x_i \in X$ and Δ satisfies

$$\int_{\mathcal{M}} \Delta f \cdot g \cdot p^2\mathrm{dVol} = D(f, g)$$

for all $f, g \in \text{Dom}(\Delta)$. In case that $p \in C^1(\mathcal{M})$, Δ is nothing but

$$\Delta f = -\Delta_2 f = -\frac{1}{p^2} \text{div}(p^2 \nabla f)$$

from [15, Definition 8]. In the remainder, we use $\Delta_\Gamma^{\text{rw}}$ to denote the random walk graph Laplacian and Δ^{rw} for its continuous counterpart. Showing the closeness of $\mathbf{m}$ and ρ, (1.10), reduces to showing a kernel-density estimate on a manifold. In "Appendix B," we show that provided h satisfies Assumption 3, we have

$$\max_{i=1,\ldots,n} |m_i - p(\mathbf{x}_i)| \leq CL_p h + C\alpha\eta(0)m\omega_m \frac{\varepsilon}{h} + C\alpha m \left(K + \frac{1}{R^2} \right) h^2, \quad (1.12)$$

where $C > 0$ is a universal constant and ε is the ∞-OT distance between μ_n and μ (see (1.13) and Sect. 2). These estimates are proved using a simple and general approach using the transportation maps introduced in Sect. 2, in contrast to usual kernel-density estimation approaches. The estimates are not optimal, but they are on the same order of error as the approximation error of the Dirichlet form D by the discrete Dirichlet form b that we present in Lemmas 13 and 14; the bottom line is that the rates of convergence for the spectrum of the random walk graph Laplacian are unaffected by the non-optimal estimate (1.12). On the other hand, our proof of (1.12) has the advantage of reducing all probabilistic estimates in our problem to estimating the ∞-OT distance between μ_n and μ; which is done in Sect. 2.

1.2.3 Normalized Graph Laplacian

So far, we have described how one can obtain the unnormalized and random walk Laplacians as examples of the general framework introduced in this section. Let us recall another popular version of normalized Laplacian usually referred to as *symmetric* normalized graph Laplacian. For given $u: X \to \mathbb{R}$, the symmetric normalized Laplacian of u is given by

$$(\Delta_\Gamma^S u)(x_i) := \frac{2}{\sigma_\eta h^2} \sum_{j:i \sim j} \frac{w_{i,j}}{\sqrt{m_i}} \left(\frac{u(x_i)}{\sqrt{m_i}} - \frac{u(x_j)}{\sqrt{m_j}} \right)$$

with m_i defined by (1.11). We remark that Δ_Γ^S cannot be obtained by appropriately choosing the measure $\mathbf{m}\mu$ as described in this section. (In order to recover it, we would have to modify the definition of discrete differential in (1.6).) Nevertheless, we can indirectly analyze the rate convergence of its spectrum toward that of a continuous counterpart noting that Δ^S and Δ^{rw} are similar matrices. Indeed, we recall that $\Delta_\Gamma^S u = \lambda u$ if and only if $\Delta_\Gamma^{\text{rw}} v = \lambda v$, where $v(x_i) := m_i^{-1/2} u(x_i)$. Thus, $\Delta_\Gamma^{\text{rw}}$ and Δ_Γ^S share the same spectrum.

1.3 Main Results

1.3.1 Convergence of Eigenvalues and Transportation Estimates

Our first main result is the following.

Theorem 1 *Let $x_1, \ldots, x_n$ be i.i.d. samples from a distribution μ supported on $\mathcal{M}$, with density p satisfying (1.2). Consider $\mathbf{m}$ and ρ as in Sect. 1.2.1 or Sect. 1.2.2. For $k \geq 2$, let $\lambda_k(\Gamma)$ be the kth eigenvalue of the graph Laplacian Δ_Γ defined in Sect. 1.2 with*

$$h := \sqrt{\frac{\log(n)^{p_m}}{n^{1/m}}},$$

where $p_m = 3/4$ if $m = 2$ and $p_m = 1/m$ if $m \geq 3$. Let $\lambda_k(\mathcal{M})$ be the kth eigenvalue of the Laplacian Δ defined in Sect. 1.2. Then,

$$\frac{|\lambda_k(\Gamma) - \lambda_k(\mathcal{M})|}{\lambda_k(\mathcal{M})} = O\left(\sqrt{\frac{\log(n)^{p_m}}{n^{1/m}}}\right), \quad \text{almost surely.}$$

The actual choice of h in the previous theorem is explained by the more general and detailed result stated in Theorem 4, together with the estimates for the ∞-OT distance between μ and μ_n in Theorem 2. Indeed, we have taken h to scale like $\sqrt{\varepsilon}$, where ε is the ∞-OT distance between μ and μ_n. More precisely,

$$\varepsilon = d_\infty(\mu, \mu_n) := \min_{T: T_\sharp \mu = \mu_n} \operatorname{esssup}_{x \in \mathcal{M}} d(x, T(x)). \tag{1.13}$$

where $T_\sharp \mu = \mu_n$ means that $\mu(T^{-1}(U)) = \mu_n(U)$ for every Borel subset U of $\mathcal{M}$. Such mappings T are called transport maps from μ to μ_n. One of the key ingredients needed to establish Theorem 1 is the probabilistic estimate on ∞-OT distance contained in our next theorem.

Theorem 2 *Let $\mathcal{M}$ be a smooth, connected, compact manifold with dimension m. Let $p: \mathcal{M} \to \mathbb{R}$ be a probability density satisfying (1.2) and consider the measure $d\mu = p\, d\text{Vol}$. Let $x_1, \ldots, x_n$ be an i.i.d sample of μ. Then, for any $\beta > 1$ and every $n \in \mathbb{N}$ there exists a transportation map $T_n: \mathcal{M} \to X$ and a constant A such that*

$$\sup_{x \in \mathcal{M}} d(x, T_n(x)) \leq \ell := A \begin{cases} \frac{\log(n)^{3/4}}{n^{1/2}}, & \text{if } m = 2, \\ \frac{(\log n)^{1/m}}{n^{1/m}}, & \text{if } m \geq 3, \end{cases} \tag{1.14}$$

holds with probability at least $1 - C_{K, \text{Vol}(\mathcal{M}), m, i_0} \cdot n^{-\beta}$, where A depends only on K, i_0, m, $\text{Vol}(\mathcal{M})$, α and β.

The exact dependency of A in (1.14) on the geometry of $\mathcal{M}$ is given in Lemma 1. We remark that the scaling on n on the right-hand side is optimal, even in the Euclidean case [11].

With the estimates in Theorem 2 at hand, Theorem 1 follows from the more general Theorem 4 below (more precisely from its corollaries). Indeed, convergence rates for the spectrum of graph Laplacians can be written in terms of h and ε as long as $0 < \varepsilon \ll h \ll 1$. Throughout this paper, we assume that h, ε, $\frac{\varepsilon}{h}$ and $\|\mathbf{m} - \rho\|_\infty$ are sufficiently small. In particular, we make the following assumptions.

Assumption 3 Assume that

$$h < \min\left\{1, \frac{i_0}{10}, \frac{1}{\sqrt{mK}}, \frac{R}{\sqrt{27m}}\right\} \quad \text{and} \quad (m+5)\varepsilon < h,$$

where i_0 is the injectivity radius of the manifold $\mathcal{M}$, K is a global upper bound on the absolute value of sectional curvatures of $\mathcal{M}$, m is the dimension of $\mathcal{M}$, and R is the reach of $\mathcal{M}$ (seen as a submanifold embedded in $\mathbb{R}^d$).

Theorem 4 *For $k \in \mathbb{N}$, let $\lambda_k(\Gamma)$ be the kth eigenvalue of the graph Laplacian Δ_Γ defined in Sect. 1.2 using the weights $\mathbf{m}$, and let $\lambda_k(\mathcal{M})$ be the kth eigenvalue of the Laplacian Δ defined in Sect. 1.2 using the weight function ρ. Finally let ε be the ∞-OT distance between μ_n and μ and assume that $h > 0$ satisfies Assumptions 3. Then,*

1. *(Upper bound) If ε and $\|\mathbf{m} - \rho\|_\infty$ are such that*

$$\sqrt{\lambda_k(\mathcal{M})}\,\varepsilon + \|\mathbf{m} - \rho\|_\infty < c, \tag{1.15}$$

for a positive constant c that depends only on m, α, L_ρ, L_p and η, then,

$$\frac{\lambda_k(\Gamma) - \lambda_k(\mathcal{M})}{\lambda_k(\mathcal{M})} \leq \tilde{C}\left(L_p h + \frac{\varepsilon}{h} + \sqrt{\lambda_k(\mathcal{M})}\varepsilon + K h^2 + \frac{h^2}{R^2} + \|\mathbf{m} - \rho\|_\infty\right) \tag{1.16}$$

where $\tilde{C}$ only depends on m, α, L_ρ, L_p, and η. We recall that L_p and L_ρ stand for the Lipschitz constants of p and ρ, respectively.

2. *(Lower bound) If h and $\|\mathbf{m} - \rho\|_\infty$ are such that*

$$\sqrt{\lambda_k(\mathcal{M})}h + \|\mathbf{m} - \rho\|_\infty < c, \tag{1.17}$$

for a positive constant c that depends only on m, α, L_ρ, L_p, and η, then,

$$\frac{\lambda_k(\Gamma) - \lambda_k(\mathcal{M})}{\lambda_k(\mathcal{M})} \geq -\tilde{C}\left(L_p h + \frac{\varepsilon}{h} + \sqrt{\lambda_k(\mathcal{M})}h + K h^2 + \|\mathbf{m} - \rho\|_\infty\right) \tag{1.18}$$

where $\tilde{C}$ only depends on m, α, L_ρ, L_p, and η.

Remark 4 Note that the lower bound does not depend on the reach R. This is due to the one-sided inequality

$$|x - y| \leq d(x, y), \quad \forall x, y \in \mathcal{M}.$$

In contrast, for the upper bound one must use a reverse inequality with an additional higher-order correction term that depends on R. See Proposition 2.

It is also worth pointing out that the presence of the term $\sqrt{\lambda_k(\mathcal{M})}\varepsilon$ in the upper bound ultimately comes from the estimate on how far is the map P in (1.22) from being an isometry when restricted to the first k eigenspaces of Δ; the relevant length scale for this estimate is the size of transport cells, i.e., ε. On the other hand, the term $\sqrt{\lambda_k(\mathcal{M})}h$ in the lower bound comes from the estimate on how far is the map I in (1.24) from being an isometry when restricted to the first k eigenspaces of Δ_Γ; the relevant length scale for this estimate is h, which is of the same order as the bandwidth for the kernel used to define the map I. This can be seen from Lemmas 13 and 14, respectively.

Remark 5 From the estimates (1.16) and (1.18), we see that curvature of $\mathcal{M}$ only introduces a second-order correction to the rate of convergence of $\lambda_k(\Gamma)$ toward $\lambda_k(\mathcal{M})$.

The estimates on ε from Theorem 2 combined with Theorem 4 imply that $\lambda_k(\Gamma)$ converges toward $\lambda_k(\mathcal{M})$ with probability one whenever $\|\mathbf{m} - \rho\|_\infty \to 0$, $h \to 0$, $\frac{\varepsilon}{h} \to 0$. We can specialize Theorem 4 to the examples from Sect. 1.2, where in particular we provide estimates on $\|\mathbf{m} - \rho\|_\infty$ in terms of n.

Corollary 1 (Convergence of eigenvalues unnormalized graph Laplacian) *In the context of Theorem 4 suppose that the weights are taken to be $\mathbf{m} \equiv 1$ and $\rho \equiv 1$. If h is small enough for*

$$(\sqrt{\lambda_k(\mathcal{M})} + 1)h \leq c,$$

to hold for a positive constant c that depends only on m, α, L_p, and η, then

$$\frac{|\lambda_k(\Gamma) - \lambda_k(\mathcal{M})|}{\lambda_k(\mathcal{M})} \leq \tilde{C}\left(\frac{\varepsilon}{h} + (1 + \sqrt{\lambda_k(\mathcal{M})})h + \left(K + \frac{1}{R^2}\right)h^2\right), \quad (1.19)$$

where $\tilde{C}$ only depends on m, α, L_p, and η.

Proof The result follows directly from Theorem 4 after noticing that in this case $\|\mathbf{m} - \rho\|_\infty = 0$ and $L_\rho = 0$. □

Corollary 2 (Convergence of eigenvalues random walk graph Laplacian) *In the context of Theorem 4 suppose that the weights $\mathbf{m}$ are as in (1.11) and $\rho \equiv p$. If h and ε/h are such that*

$$(\sqrt{\lambda_k(\mathcal{M})} + 1)h + \frac{\varepsilon}{h} \leq c,$$

for a positive constant c that depends only on m, α, L_p, and η, then,

$$\frac{|\lambda_k(\Gamma) - \lambda_k(\mathcal{M})|}{\lambda_k(\mathcal{M})} \leq \tilde{C}\left(\frac{\varepsilon}{h} + (1 + \sqrt{\lambda_k(\mathcal{M})})h + \left(K + \frac{1}{R^2}\right)h^2\right), \quad (1.20)$$

where $\tilde{C}$ only depends on m, α, L_p, and η.

Proof The result follows directly from Theorem 4 after using (1.12). Indeed, notice that the term $\|\mathbf{m} - \rho\|_\infty$ can be absorbed in the h, $\frac{\varepsilon}{h}$ and h^2 terms by enlarging constants if necessary. □

Remark 6 Notice that the estimates in the previous results provide a lower bound on the mode at which the spectrum of the graph Laplacian stops being informative about the spectrum of the Laplace–Beltrami operator. Namely, notice that the right-hand sides of (1.19) and (1.20) are small when $h\sqrt{\lambda_k(\mathcal{M})}$ is small. Using Weyl's law for the growth of eigenvalues of the Laplace–Beltrami operator, we know that

$$\sqrt{\lambda_k(\mathcal{M})} \sim k^{1/m},$$

and thus, the relative error of approximating $\lambda_k(\mathcal{M})$ with $\lambda_k(\Gamma)$ is small when $k \lesssim \frac{1}{h^m}$ and $\varepsilon \ll h$. In particular, if h is taken to scale like $h = \sqrt{\varepsilon}$ (as is the case in Theorem 1) then $\lambda_k(\mathcal{M})$ is approximated by $\lambda_k(\Gamma)$ if $k \lesssim \sqrt{\frac{n}{\log(n)}}$ for $m \geq 3$ and $k \lesssim \sqrt{\frac{n}{\log(n)^{3/2}}}$ for $m = 2$.

Remark 7 We would like to remark that one of the main advantages of writing all our estimates in Theorem 4 in terms of the quantity ε (which is the only one where randomness is involved) is that we can transfer probabilistic estimates for ε into probabilistic estimates for the error of approximation of $\lambda_k(\Gamma)$. In particular, when combined with Theorem 2, Corollary 1 and Corollary 2 can be read as follows: Suppose that $\frac{\log(n)^{p_m}}{n^{1/m}} \ll h \ll 1$. Let $k := k_n$ be such that $k_n \ll \frac{1}{h^m}$. Let $\beta > 1$. Then, with probability at least $1 - C_{K,\mathrm{Vol}(\mathcal{M}),m,i_0} n^{-\beta}$,

$$\frac{|\lambda_j(\Gamma) - \lambda_j(\mathcal{M})|}{\lambda_j(\mathcal{M})} \leq \tilde{C}_\beta \left(\frac{\log(n)^{p_m}}{hn^{1/m}} + (1 + \sqrt{\lambda_j(\mathcal{M})})h + \left(K + \frac{1}{R^2} \right) h^2 \right)$$

for all $j = 1, \ldots, k_n$.

Remark 8 Moreover, writing all our estimates in Theorem 4 in terms of the quantity ε is also convenient because the conclusion of the theorem holds even when the points $x_1, \ldots, x_n$ are not i.i.d. samples from the measure μ. That is, one only needs to ensure that the assumption (1.15) is satisfied. In other words, whenever one has an estimate on the ∞-OT distance between the point cloud and the measure μ and a kernel-density estimate to ensure that (1.15) holds, the theorem provides an error estimate on the eigenvalues. We note that the kernel-density estimate in terms of the ∞-OT distance we provide in Lemma 18 implies that one in fact only needs an estimate on the ∞-OT distance between the point cloud and the measure μ.

1.3.2 Convergence of Eigenfunctions

We prove that eigenvectors of Δ_Γ converge toward eigenfunctions of Δ and provide quantitative error estimates. To make the statements precise, we need to make sense of how to compare functions defined on the graph/sample X with functions defined on the manifold $\mathcal{M}$. In this paper, we consider two different ways of doing this.

The first approach involves an interpolation step by composing with the optimal transportation map $T : \mathcal{M} \to X$ from (1.13) followed by a smoothening step. Both of these steps require the knowledge of $\mathcal{M}$. The map T induces a partition $U_1, \ldots, U_n$ of $\mathcal{M}$, where

$$U_i := T^{-1}(\{x_i\}). \tag{1.21}$$

We note that $\mu(U_i) = \frac{1}{n}$ for all $i = 1, \ldots, n$. We define the *contractive discretization* map $P \colon L^2(\mathcal{M}, \rho\mu) \to L^2(X, \mathbf{m}\mu_n)$ by

$$(Pf)(x_i) := n \cdot \int_{U_i} f(x) \mathrm{d}\mu(x), \quad f \in L^2(\mathcal{M}, \rho\mu), \tag{1.22}$$

and the *extension* map $P^* \colon L^2(X, \mathbf{m}\mu_n) \to L^2(\mathcal{M}, \rho\mu)$ by

$$(P^*u)(x) := \sum_{i=1}^{n} u(x_i) \mathbb{1}_{U_i}(x), \quad u \in L^2(X, \mathbf{m}\mu_n). \tag{1.23}$$

We note that P^*u can be written as $P^*u = u \circ T$. We then consider the interpolation operator $I \colon L^2(X, \mathbf{m}\mu_n) \to \mathrm{Lip}(\mathcal{M})$

$$Iu := \Lambda_{h-2\varepsilon} P^* u \tag{1.24}$$

where $\Lambda_{h-2\varepsilon}$ is defined in (3.4) and is simply a convolution operator using some particularly chosen kernel; see Sect. 1.4 for a discussion on why we need to consider a specific kernel.

Theorem 5 *Let Δ_Γ be the graph Laplacian defined in Sect. 1.2 using the weights $\mathbf{m}$, and let Δ be the Laplacian defined in Sect. 1.2 using the weight function ρ. Let ε be the ∞-OT distance between μ_n and μ and assume that $h > 0$ satisfies Assumptions 3. Finally, assume that h and $\|\mathbf{m} - \rho\|_\infty$ are small enough so that*

$$(1 + \sqrt{\lambda_k(\mathcal{M})})h + \|\mathbf{m} - \rho\|_\infty \leq c,$$

for a constant c that depends only on $m, \alpha, L_p, L_\rho, \eta$.

Then, for every $u \in L^2(X, \mathbf{m}\mu_n)$ normalized eigenfunction of Δ_Γ with eigenvalue $\lambda_k(\Gamma)$, there exists a normalized eigenfunction $f \in L^2(\mathcal{M}, \rho\mu)$ of Δ with eigenvalue $\lambda_k(\mathcal{M})$ such that

$$\|Iu - f\|_{L^2(\mathcal{M},\rho\mu)} \leq \frac{\tilde{C}}{\gamma_{k,\rho\mu}} \left(\frac{\varepsilon}{h} + (1 + \sqrt{\lambda_k(\mathcal{M})})h + Kh^2 + \frac{h^2}{R^2} + \|\mathbf{m} - \rho\|_\infty \right),$$

where $\tilde{C}$ is a constant that only depends on $m, \eta, \alpha, L_\rho, L_p$ and where $\gamma_{k,\rho\mu}$ is the spectral gap, that is

$$\gamma_{k,\rho\mu} = \min\{|\lambda_j - \lambda_k| \, : \, j \in \mathbb{N} \text{ and } \lambda_j \neq \lambda_k\}$$

In particular, if we take

$$h := \sqrt{\frac{\log(n)^{p_m}}{n^{1/m}}},$$

where $p_m = 3/4$ for $m = 2$ and $p_m = 1/m$ for $m \geq 3$, then,

$$\|Iu - f\|_{L^2(\mathcal{M}, \rho\mu)} = O\left(\sqrt{\frac{\log(n)^{p_m}}{n^{1/m}}}\right), \quad \text{almost surely.}$$

Remark 9 As in Remark 7, we would like to emphasize that the probabilistic estimates for ε translate directly into probabilistic estimates for the convergence of eigenfunctions in Theorem 5. Likewise, we would like to point out that Theorem 5 can be made concrete in the context of Sects. 1.2.1 and 1.2.2 using the corresponding estimates for $\|\mathbf{m} - \rho\|_\infty$ in terms of ε and h.

The second approach to compare eigenvectors of Δ_Γ with eigenfunctions of Δ is to extrapolate the values of discrete eigenvectors to the Euclidean Voronoi cells induced by the points $\{x_1, \ldots, x_n\}$. That is, for an arbitrary function $u : X \to \mathbb{R}$ we assign to each point $x \in \mathcal{M}$ the value $u(x_i)$ where x_i is the nearest neighbor of x in X with respect to the Euclidean distance. More formally, for $i \in \{1, \ldots, n\}$ we consider the Voronoi cells

$$V_i := \{x \in \mathcal{M} : |x - x_i| = \min_{j=1,\ldots,n} |x - x_j|\}, \tag{1.25}$$

and define the function $\bar{u} \in L^2(\mathcal{M}, \rho\mu)$ by

$$\bar{u}(x) := \sum_{i=1}^{n} u(x_i) \mathbb{1}_{V_i}(x) \quad \text{for } x \in \mathcal{M}. \tag{1.26}$$

We notice that the Voronoi cells $V_1, \ldots, V_n$ form a partition of $\mathcal{M}$, up to a set of ambiguity of μ-measure zero. Besides being a computationally simple interpolation, the Voronoi extension can be constructed exclusively from the data and no information on $\mathcal{M}$ is needed. We show that the interpolation $\bar{u}$ of a discrete eigenvector u approximates the corresponding eigenfunction f on $\mathcal{M}$ with almost the same rate as in Theorem 5. In order to obtain convergence of the Voronoi extensions $\bar{u}$, we require $h = h_n$ to satisfy

$$\lim_{n\to\infty} \log^{mp_m}(n) \cdot \left(h + \frac{\varepsilon}{h}\right) = 0 \tag{1.27}$$

This condition holds, for instance, when h is chosen as $\sqrt{\varepsilon}$, which minimizes the error in the following result.

Theorem 6 *Fix $\beta > 1$. Let Δ_Γ be the graph Laplacian defined in Sect. 1.2 using the weights $\mathbf{m}$, and let Δ be the Laplacian defined in Sect. 1.2 using the weight function*

ρ. Let ε be the ∞-OT distance between μ_n and μ and assume that $h > 0$ satisfies Assumptions 3. Finally, assume that h and $\|\mathbf{m} - \rho\|_\infty$ are small enough so that in particular

$$(1 + \sqrt{\lambda_k(\mathcal{M})})h + \|\mathbf{m} - \rho\|_\infty \leq c,$$

for a constant c that depends only on $m, \alpha, L_p, L_\rho, \eta$.

Then, with probability at least $1 - C_{m,K,Vol(\mathcal{M}),i_0} \cdot n^{-\beta}$, for every $u \in L^2(X, \mathbf{m}\mu_n)$ normalized eigenfunction of Δ_Γ corresponding to the eigenvalue $\lambda_k(\mathcal{M})$, it holds

$$
\begin{aligned}
\|\bar{u} - f\|_{L^2(\mathcal{M}, \rho\mu)} &\leq \frac{\tilde{C}_\beta \sqrt{\log(n)^{mp_m}}}{\gamma_{k,\rho\mu}} \left(\frac{\varepsilon}{h} + (1 + \sqrt{\lambda_k(\mathcal{M})})h \right. \\
&\left. + Kh^2 + \frac{h^2}{R^2} + \|\mathbf{m} - \rho\|_\infty \right) + C_{\mathcal{M}} \lambda_k(\mathcal{M})^{\frac{m+1}{4}} \varepsilon,
\end{aligned}
\tag{1.28}
$$

where f and $\gamma_{k,\rho\mu}$ are as in Theorem 5, $\bar{u}$ is as in (1.26), and $C_{\mathcal{M}}$ is a constant that depends on the manifold $\mathcal{M}$.

Remark 10 We remark that the first term in (1.28) is worse than the estimate in Theorem 5 by a logarithmic factor of n. This is due to our uniform estimates on the size of Voronoi cells based on transportation (see Lemma 17). On the other hand, the extra factor in (1.28) is an estimate for the difference of the averages of f over transport cells and Voronoi cells; here, we use the regularity of an eigenfunction f and in particular we use a bound for $\|\nabla f\|_\infty$ found in [22].

1.4 Outline of the Approach and Discussion

To prove our main results, we exploit well-known variational characterizations for the spectra of Δ_Γ and Δ. Our results are then deduced from a careful comparison between the objective functionals of the variational problems.

From the definition of Δ_Γ in Sect. 1.2, it clear that Δ_Γ is positive-semidefinite with respect to the inner product of $L^2(X, \mathbf{m}\mu_n)$. We denote by

$$0 = \lambda_1(\Gamma) \leq \lambda_2(\Gamma) \leq \lambda_3(\Gamma) \leq \cdots$$

the eigenvalues of Δ_Γ, repeated according to their multiplicities. By the minmax principle, we have

$$\lambda_k(\Gamma) = \min_{L_k} \max_{u \in L_k \backslash \{0\}} \frac{b(u)}{\|u\|_{L^2(X, \mathbf{m}\mu_n)}^2} \tag{1.29}$$

where the minimum is over all k-dimensional subspaces L_k of $L^2(X, \mathbf{m}\mu_n)$. At the continuum level, and given that ρ and p are bounded from below, one can show that Δ is a closed and densely defined symmetric operator with compact resolvent [1, Lemma

2.7]. Therefore, its spectrum consists of positive eigenvalues only, which we denote by

$$0 = \lambda_1(\mathcal{M}) \le \lambda_2(\mathcal{M}) \le \lambda_3(\mathcal{M}) \le \cdots,$$

where eigenvalues are repeated according to their multiplicities. Moreover, by Courant's minmax principle we have

$$\lambda_k(\mathcal{M}) = \min_{L_k} \max_{f \in L_k \setminus \{0\}} \frac{D(f)}{\|f\|^2_{L^2(\mathcal{M}, \rho\mu)}} \tag{1.30}$$

where the minimum is over all k-dimensional subspaces L_k of $L^2(\mathcal{M}, \rho\mu)$, see [18, Lemma 2.9].

The proof of our results may be split into two main parts. The first part contains all the probabilistic estimates needed in the rest of the paper and is devoted to the proof of Theorem 2. The study of the estimates for $d_\infty(\mu, \mu_n)$ goes back to [16,24,27], where the problem was considered in a simpler setting: μ is the Lebesgue measure on the unit cube $(0, 1)^m$ and the points $x_1, \ldots, x_n$ are i.i.d. uniformly distributed on $(0, 1)^d$. In that context, with very high probability,

$$d_\infty(\mu, \mu_n) \approx \frac{(\log(n))^{p_m}}{n^{1/m}},$$

where p_m is defined in (1.1). In [11], the estimates are extended to measures defined on more general domains, that is to bounded open domains in $\mathbb{R}^d$ with Lipschitz boundary, and to general densities (with positive upper and lower bounds). In this paper, we extend the results in [11] to the manifold case. In order to prove Theorem 2, we use a similar proof scheme to the one used in [11]. Indeed, we first establish Lemma 1 below which is analogous to [11, Theorem 1.2] and is of interest on its own. The result includes explicit estimates on how the distance depends on the geometry of $\mathcal{M}$.

Lemma 1 *Let ρ_1, ρ_2 be two probability densities defined on $\mathcal{M}$ with*

$$\frac{1}{\alpha} \le \rho_i(x) \le \alpha \quad \text{for all } x \in \mathcal{M} \text{ and } i \in \{1, 2\}$$

for some $\alpha \ge 1$. Then, it holds for the corresponding measures ν_1, ν_2, defined as $d\nu_1 = \rho_1 dx$ and $d\nu_2 = \rho_2 dx$,

$$d_\infty(\nu_1, \nu_2) \le A\|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})}, \tag{1.31}$$

where $A = \frac{C_{m,\alpha} \operatorname{Vol}(\mathcal{M})^3}{r^{3m-1}} \max\left\{\tilde{N}_c, \frac{\operatorname{diam}(\mathcal{M})}{r}\right\}$, $\tilde{N}_c = (1 + CmKr^2)\frac{2^m \operatorname{Vol}(\mathcal{M})}{\omega_m r^m}$ and $r = \frac{1}{5}\min\{1, i_0, \frac{1}{\sqrt{K}}\}$.

With Lemma 1 at hand, the next step is to construct a careful partition of the manifold $\mathcal{M}$ into patches in which we can use directly the results from [11]. The construction requires some geometric estimates which are obtained in Sect. 2.1. Using properties of the constructed partition of $\mathcal{M}$ and Lemma 1, we can establish Theorem 2.

The second part of the proof of our main results consists of a set of precise deterministic computations used to relate the discrete and continuum Dirichlet energies appearing in the variational characterization of the spectra of the graph and continuum Laplacians; these computations are based on ideas from [6]. Roughly speaking, the proof of our main results relies on the following upper and lower bounds. We first show the *upper bound*

$$b(Pf) \leq (1 + \text{error})E_r(f) \leq (1 + \text{error})D(f), \quad f \in L^2(\mathcal{M}, \rho\mu),$$

where E_r is the non-local kernel approximation of the continuum Dirichlet energy defined in (3.1) and r is a length scale which up to leading order is equal to h; the term *error* can be explicitly written in terms of h, ε and geometric quantities associated with the manifold $\mathcal{M}$. It is possible to interpret the first inequality as a "variance" estimate as it relates an energy constructed exclusively from the graph with an "average" energy. The second inequality on the other hand can be thought as a "bias" estimate. We would like to point out that the second inequality is a manifestation of the intuitive fact that local energies bound non-local ones. Our *lower bound* takes the form

$$D(I(u)) \leq (1 + \text{error})E_r(P^*u) \leq (1 + \text{error})b(u), \quad u \in L^2(X).$$

We remark that it is not too hard to obtain a relation of the form $D(I(u)) \leq C E_r(P^*u)$ for some constant C. Nevertheless, since our goal is to find *error estimates*, the constant C must be sharp (up to some small error). We obtain this sharp constant using the specific form of the convolution operator Λ in the definition of I (see (3.5)). Our analysis of convergence of the spectra is completed by showing that the maps P, P^* and I are almost isometries when restricted to eigenspaces (discrete or continuum).

We want to highlight the fact that in contrast to the construction in [6], our graphs and our "out-of-sample extensions" of eigenvectors are defined exclusively from the ambient space Euclidean distance. Theorem 6 is obtained a posteriori from Theorem 2 and uses Theorem 2 to bound the measure of Voronoi cells. We also use uniform estimates for the gradient of eigenfunctions of the Laplace–Beltrami operator from [22].

1.5 Outline

The rest of the paper is organized as follows. In Sect. 1.6, we recall some notation and state some results from differential geometry that we need for the sequel. Section 2 is devoted to the estimation of the ∞-transportation distance between μ_n and μ and in particular contains the proof of Theorem 2. Section 3 contains results on the kernel-based approximation of the Laplacian operator; in more precise terms, we relate the (weighted) Dirichlet energy D with the non-local Dirichlet energy (3.1).

Section 4 addresses the convergence of eigenvalues and in particular contains the proof of Theorem 4. Finally, in Sect. 5 we establish the convergence of eigenvectors of graph Laplacians, first in the sense of the interpolation map I from 1.24 (Theorem 5) and then in the sense of Voronoi extensions (Theorem 6). "Appendix B" contains the optimal transportation-based proof of kernel-density estimates on manifolds.

1.6 Some Estimates from Differential Geometry

We conclude this introduction by recalling some notation and stating a few results from differential geometry.

For a point $x \in \mathcal{M}$, we denote by $T_x\mathcal{M}$ the tangent space of $\mathcal{M}$ at x. Fix $0 < r \leq \min\{i_0, 1/\sqrt{K}\}$ and let us denote by $\exp_x \colon B(r) \subseteq T_x\mathcal{M} \to \mathcal{M}$ the Riemannian exponential map. Since $r < i_0$, the map is a diffeomorphism between the ball $B(r)$ and the geodesic ball $B_{\mathcal{M}}(x, r)$. In particular, $\exp_x^{-1}$ defines a local chart at x. Let g be the pull back of the metric of $\mathcal{M}$ by the exponential map. That is, for an orthonormal basis $e_1, \ldots, e_m$ of $T_x\mathcal{M}$ and for given $v \in B(r)$ let $g_{i,j}|_v := \langle (d\exp_x)_v(e_i), (d\exp_x)_v(e_j) \rangle$, where we have identified the tangent space of $T_x\mathcal{M}$ at v with $T_x\mathcal{M}$ itself. Then,

$$\delta_{i,j} - CK\,|v|^2 \leq g_{i,j} \leq \delta_{i,j} + CK\,|v|^2, \tag{1.32}$$

where $|v|$ is the Euclidean length of v, $\delta_{i,j}$ is 1 if $i = j$ and 0 otherwise and where C is a universal constant. Such estimates are bounds on the metric distortion by the exponential map and follow from Rauch comparison theorem ([9, Chapter 10] and [6, Section 2.2]). Similarly, since $r < 1/\sqrt{K}$, one can show that for any $v \in B(r)$ and any $w \in T_x\mathcal{M} \cong T_v(T_x\mathcal{M})$,

$$\frac{1}{2}|w|_x \leq |(d\exp_x)_v(w)|_{\exp_x(v)} \leq 2|w|_x. \tag{1.33}$$

These estimates imply the following:

Proposition 1 *Assume* $0 < r \leq \min\{i_0, 1/\sqrt{K}\}$. *Let* $p \in \mathcal{M}$ *and consider any smooth curve* $\gamma \colon [0, 1] \to B(r) \subset T_p\mathcal{M}$. *Then,*

$$\frac{1}{2}\,\mathrm{Length}(\gamma) \leq \mathrm{Length}(\exp_p \circ\gamma) \leq 2\,\mathrm{Length}(\gamma).$$

Furthermore, on $B_{\mathcal{M}}\left(p, \frac{r}{2}\right)$ *the exponential mapping* $\exp_p \colon B\left(0, \frac{r}{2}\right) \subseteq T_p\mathcal{M} \to B_{\mathcal{M}}\left(p, \frac{r}{2}\right)$ *is a bi-Lipschitz bijection with bi-Lipschitz constant 2.*

The proof of this proposition is presented in "Appendix A."

The bounds on metric distortion (1.32) imply that the Jacobian of the exponential map (i.e., the volume element) $J_x(v) := \sqrt{\det(g)}$ satisfies

$$(1 + CmK\,|v|^2)^{-1} \leq J_x(v) \leq (1 + CmK\,|v|^2). \tag{1.34}$$

A direct consequence of (1.34) is that

$$\frac{\omega_m r^m}{1 + CmKr^2} \leq \text{Vol}(B_{\mathcal{M}}(x, r)) \leq (1 + CmKr^2)\omega_m r^m, \qquad (1.35)$$

which implies that

$$\left|\text{Vol}(B_{\mathcal{M}}(x, r)) - \omega_m r^m\right| \leq CmK r^{m+2} \qquad (1.36)$$

where ω_m is the volume of the unit ball in $\mathbb{R}^m$.

Now, we want to state a relation between the intrinsic distance on the manifold and the Euclidean distance on the ambient space. For that purpose, we recall that R, the reach of the manifold $\mathcal{M}$, is defined as

$$\begin{aligned} R := \sup \big\{ t > 0 \, : \, \forall x \in \mathbb{R}^d, \, \text{dist}(x, \mathcal{M}) \leq t, \\ \exists! y \in \mathcal{M} \, \text{ s.t. } \, \text{dist}(x, \mathcal{M}) = |x - y| \big\}. \end{aligned} \qquad (1.37)$$

We note that R is an extrinsic quantity, meaning it depends on the specific embedding of $\mathcal{M}$ into $\mathbb{R}^d$. In addition, we note that the quantity $\frac{1}{R}$ is related to extrinsic curvature, as it uniformly controls the principal curvatures of $\mathcal{M}$ (see [19]). We now show that the distances $\mathcal{M}$ are locally a second-order perturbation of the Euclidean distance in $\mathbb{R}^d$ and provide explicit error bounds in terms of the reach of $\mathcal{M}$.

Proposition 2 *Let R be the reach of the manifold $\mathcal{M} \subseteq \mathbb{R}^d$. Let $x, y \in \mathcal{M}$ and suppose that $|x - y| \leq \frac{R}{2}$. Then,*

$$|x - y| \leq d(x, y) \leq |x - y| + \frac{8}{R^2} |x - y|^3.$$

The proof of this proposition is presented in "Appendix A."

2 The ∞-Transportation Distance

The main goal of this section is to prove Theorem 2. For that purpose, we use a similar proof scheme to the one used in [11]. We first establish Lemma 1, and then we construct a "nice" partition of the manifold $\mathcal{M}$ by using a Voronoi tessellation using some (fixed) appropriately chosen points; what makes the partition nice is that each of its cells is bi-Lipschitz homeomorphic (with universal bi-Lipschitz constant) to a fixed ball in $\mathbb{R}^m$ where we can apply the results from [11]. In Sect. 2.1, we present the construction of such partition and prove Theorem 2.

Throughout this section, we make use of the following construction and estimates. Let $r = \frac{1}{5} \min\{1, i_0, \frac{1}{\sqrt{K}}\}$. Let $Y = \{y_i : i \in I\}$ be a maximal subset of $\mathcal{M}$ such that $d(y_i, y_j) \geq r$ for all $i \neq j$. Note that the balls $\{B_{\mathcal{M}}(y_i, r/2)\}_{i \in I}$ do not overlap. From

([1.35](#)), we conclude that $N_c := \operatorname{card} Y$ satisfies

$$N_c(1 + CmKr^2)^{-1}\frac{\omega_m r^m}{2^m} \le \sum_{i \in I} \operatorname{Vol}\left(B_{\mathcal{M}}(y_i, r/2)\right) \le \operatorname{Vol}(\mathcal{M})$$

and hence

$$N_c \le (1 + CmKr^2)\frac{2^m \operatorname{Vol}(\mathcal{M})}{\omega_m r^m}. \tag{2.1}$$

From now on, we list the elements of Y as $y_1, \ldots, y_{N_c}$. It follows from the maximality of Y that the collection of balls $\{B_{\mathcal{M}}(y_i, r)\}_{i=1,\ldots,N_c}$ covers $\mathcal{M}$. We also claim that if $\operatorname{dist}(y_i, y_j) \le 2r$, then the balls $B_{\mathcal{M}}(y_i, 2r)$ and $B_{\mathcal{M}}(y_j, 2r)$ have a "big" overlap in the sense that

$$(1 + CmKr^2)^{-1}\omega_m r^m \le \operatorname{Vol}(B_{\mathcal{M}}(y_i, 2r) \cap B_{\mathcal{M}}(y_j, 2r)). \tag{2.2}$$

In fact, let y_{ij} be the point that is halfway from y_i to y_j on the geodesic connecting y_i and y_j. Let $y \in B_{\mathcal{M}}(y_{ij}, r)$. Then, $\operatorname{dist}(y, y_i) \le \operatorname{dist}(y, y_{ij}) + \operatorname{dist}(y_{ij}, y_i) < r + r \le 2r$. This shows that $B_{\mathcal{M}}(y_{ij}, r) \subseteq B_{\mathcal{M}}(y_i, 2r)$. Similarly, we have $B_{\mathcal{M}}(y_{ij}, r) \subseteq B_{\mathcal{M}}(y_j, 2r)$. Inequality ([2.2](#)) now follows from the fact that

$$(1 + CmKr^2)^{-1}\omega_m r^m \le \operatorname{Vol}(B_{\mathcal{M}}(y_{ij}, r))$$

We now claim that for arbitrary y_i, y_j, there is a way to start from y_i and move from ball to ball until reaching y_j in such a way that any two consecutive balls visited have big overlap, i.e., that ([2.2](#)) holds. To make this idea precise, let us consider a graph $(Y, \leftrightarrow)$ where

$$y_j \leftrightarrow y_i \quad \text{iff} \quad y_j \ne y_i \quad \text{and} \quad \operatorname{dist}(y_j, y_i) \le 2r. \tag{2.3}$$

We claim that $(Y, \leftrightarrow)$ is a connected graph; this is a consequence of the connectedness of $\mathcal{M}$. In fact, suppose for the sake of contradiction that $(Y, \leftrightarrow)$ is not connected. Then, we can find a partition of Y into two nonempty sets S_1, S_2 such that for all $y_i \in S_1$ and all $y_j \in S_2$, $y_i \not\leftrightarrow y_j$ (i.e., $d(y_i, y_j) > 2r$). Because of this, we can find $\varepsilon > 0$ such that

$$\bigcup_{y \in S_1} B_{\mathcal{M}}(y, r + \varepsilon) \cap \bigcup_{y \in S_2} B_{\mathcal{M}}(y, r + \varepsilon) = \emptyset,$$

but since

$$\mathcal{M} = \bigcup_{i=1,\ldots,N_c} B_{\mathcal{M}}(y_i, r) \subseteq \bigcup_{y \in S_1} B_{\mathcal{M}}(y, r + \varepsilon) \cup \bigcup_{y \in S_2} B_{\mathcal{M}}(y, r + \varepsilon),$$

this implies that $\mathcal{M}$ is disconnected, which is not true. Hence, we conclude that the graph $(Y, \leftrightarrow)$ is connected. We are now ready to prove Lemma [1](#).

Proof *(Lemma 1)* In order to estimate $d_\infty(\rho_1, \rho_2)$, the idea is to construct intermediate densities and estimate the distances between them using [11, Theorem 1.2]. But to use [11, Theorem 1.2], we need to map the intermediate densities to the Euclidean space. Motivated by this, we consider the balls $B_\mathcal{M}(y_1, 2r), \ldots, B_\mathcal{M}(y_{N_c}, 2r)$ constructed before. By relabeling if necessary, the connectedness of the graph $(Y, \leftrightarrow)$ implies that we can assume that for every $k = 1, \ldots, N_c$, the graph $(\{y_1, \ldots, y_k\}, \sim)$ is connected. For $k = 1, \ldots, N_c$, we define the sets

$$I_k := B_\mathcal{M}(y_k, 2r) \setminus \bigcup_{j=1}^{k-1} B_\mathcal{M}(y_j, 2r), \quad O_k := B_\mathcal{M}(y_k, 2r) \cap \bigcup_{j=1}^{k-1} B_\mathcal{M}(y_j, 2r).$$

Note that $I_1 = B(y_1, 2r)$ and $O_1 = \emptyset$. We define the functions γ_k^+, γ_k^-, $\tilde{\rho}_k$ iteratively as follows. Let us start with $k = N_c$. If $\int_{I_{N_c}} \rho_1 dx \geq \int_{I_{N_c}} \rho_2 dx$ we set $\gamma_{N_c}^+ = \rho_1$ and $\gamma_{N_c}^- = \rho_2$; if not, we reverse the roles of ρ_1 and ρ_2. We let $\tilde{\rho}_{N_c}$ be

$$\tilde{\rho}_{N_c}(x) = \begin{cases} \gamma_{N_c}^-(x) & \text{if } x \in I_{N_c}, \\ \gamma_{N_c}^+(x) + \beta_{N_c} & \text{if } x \in O_{N_c}, \\ \gamma_{N_c}^+(x) & \text{otherwise}, \end{cases}$$

where

$$\beta_{N_c} := \frac{\int_{I_{N_c}} (\gamma_{N_c}^+ - \gamma_{N_c}^-) dx}{\mathrm{Vol}(O_{N_c})}.$$

Having defined the functions $\gamma^+, \gamma^-, \tilde{\rho}$ for the iterations $N_c, N_c - 1, \ldots, k + 1$, we define the functions $\gamma_k^+, \gamma_k^-, \tilde{\rho}_k$ as follows. If $\int_{I_k} \gamma_{k+1}^- dx \geq \int_{I_k} \tilde{\rho}_{k+1} dx$, we set $\gamma_k^+ = \gamma_{k+1}^-$ and $\gamma_k^- = \tilde{\rho}_{k+1}$; if not, we reverse the roles of γ_{k+1}^- and $\tilde{\rho}_{k+1}$. The function $\tilde{\rho}_k$ is defined as

$$\tilde{\rho}_k(x) = \begin{cases} \gamma_k^-(x) & \text{if } x \in I_k, \\ \gamma_k^+(x) + \beta_k & \text{if } x \in O_k, \\ \gamma_k^+(x) & \text{otherwise}, \end{cases}$$

where

$$\beta_k := \frac{\int_{I_k} (\gamma_k^+ - \gamma_k^-) dx}{\mathrm{Vol}(O_k)}.$$

We note that $\tilde{\rho}_1 = \gamma_1^-$ and set $\beta_1 := 0$. Also, observe that for every k, $\beta_k \geq 0$ and

$$\int_\mathcal{M} \gamma_k^- dx = \int_\mathcal{M} \gamma_k^+ dx = \int_\mathcal{M} \tilde{\rho}_k dx,$$

where the second equality follows from the definition of β_k and where the first equality follows iteratively from the definitions above. Using the triangle inequality and the above definitions, we obtain

$$
\begin{aligned}
d_\infty(\rho_1, \rho_2) &= d_\infty(\gamma_{N_c}^+, \gamma_{N_c}^-) \\
&\leq d_\infty(\gamma_{N_c}^+, \tilde{\rho}_{N_c}) + d_\infty(\tilde{\rho}_{N_c}, \gamma_{N_c}^-) \\
&= d_\infty(\gamma_{N_c}^+, \tilde{\rho}_{N_c}) + d_\infty(\gamma_{N_c-1}^+, \gamma_{N_c-1}^-) \\
&\leq d_\infty(\gamma_{N_c}^+, \tilde{\rho}_{N_c}) + d_\infty(\gamma_{N_c-1}^+, \tilde{\rho}_{N_c-1}) + d_\infty(\tilde{\rho}_{N_c-1}, \gamma_{N_c-1}^-).
\end{aligned}
$$

Continuing the chain of inequalities provides, by induction,

$$
d_\infty(\rho_1, \rho_2) \leq \sum_{k=1}^{N_c} d_\infty(\gamma_k^+, \tilde{\rho}_k).
$$

Our goal is to estimate each of the terms $d_\infty(\gamma_k^+, \tilde{\rho}_k)$. From the definitions above, it is straightforward to see that γ_k^+ and $\tilde{\rho}_k$ coincide in $\mathcal{M} \setminus B_{\mathcal{M}}(y_k, 2r)$, and thus

$$
d_\infty(\rho_1, \rho_2) \leq \sum_{k=1}^{N_c} d_\infty(\gamma_k^+, \tilde{\rho}_k) \leq \sum_{k=1}^{N_c} d_\infty(\gamma_k^+|_{B_{\mathcal{M}}(y_k, 2r)}, \tilde{\rho}_k|_{B_{\mathcal{M}}(y_k, 2r)}). \quad (2.4)
$$

The last inequality is a consequence of the following observation: If two measures ν_1, ν_2 give the same total mass and we can write $\nu_1 = \nu + \tilde{\nu}_1$ and $\nu_2 = \nu + \tilde{\nu}_2$, then one possible way to transport mass from ν_1 into ν_2 is to leave the mass distributed as ν where it is and simply focus on transporting the mass distributed as $\tilde{\nu}_1$ to have it distributed as $\tilde{\nu}_2$. This observation leads to the desired inequality.

In order to obtain an estimate on $d_\infty(\gamma_k^+|_{B_{\mathcal{M}}(y_k, 2r)}, \tilde{\rho}_k|_{B_{\mathcal{M}}(y_k, 2r)})$, we first estimate $\|\gamma_k^+ - \tilde{\rho}_k\|_{L^\infty(B_{\mathcal{M}}(y_k, 2r))}$. From the definitions above, we have

$$
\|\gamma_k^+ - \tilde{\rho}_k\|_{L^\infty(B(y_k, 2r))} \leq \max\left\{\|\gamma_k^+ - \gamma_k^-\|_{L^\infty(I_k)}, \beta_k\right\}. \quad (2.5)
$$

Hence, we focus on obtaining estimates for $\|\gamma_k^+ - \gamma_k^-\|_{L^\infty(I_k)}$ and β_k.

First, we claim that for every k, the function $(\gamma_k^+ - \gamma_k^-)\mathbb{1}_{I_k}$ has the form

$$
(\gamma_k^+ - \gamma_k^-)\mathbb{1}_{I_k} = \pm(\rho_1 - \rho_2)\mathbb{1}_{I_k} + \sum_{j=k+1}^{N_c} \pm\beta_j \mathbb{1}_{I_k \cap O_j}. \quad (2.6)
$$

To see this, note that in case $k = N_c$ the result is trivial. In general, from the definitions above it follows that

$$
\begin{aligned}
(\gamma_k^+ - \gamma_k^-)\mathbb{1}_{I_k} &= \pm(\gamma_{k+1}^- - \tilde{\rho}_{k+1})\mathbb{1}_{I_k} \\
&= \pm((\gamma_{k+1}^- - \gamma_{k+1}^+)\mathbb{1}_{I_k} - \beta_{k+1}\mathbb{1}_{I_k \cap O_{k+1}}) \\
&= \pm(\gamma_{k+1}^- - \gamma_{k+1}^+)\mathbb{1}_{I_k} + \pm\beta_{k+1}\mathbb{1}_{I_k \cap O_{k+1}} \\
&= \pm(\gamma_{k+2}^- - \tilde{\rho}_{k+2})\mathbb{1}_{I_k} + \pm\beta_{k+1}\mathbb{1}_{I_k \cap O_{k+1}} \\
&= \pm(\gamma_{k+2}^- - \gamma_{k+2}^+)\mathbb{1}_{I_k} + \pm\beta_{k+2}\mathbb{1}_{I_k \cap O_{k+2}} + \beta_{k+1}\mathbb{1}_{I_k \cap O_{k+1}}.
\end{aligned}
$$

Continuing the chain of inequalities proves the claim in $N_c - k$ iterations. An immediate consequence of the previous fact is that for $k = 2, \dots, N_c$

$$
\beta_k = \pm\frac{\int_{I_k}(\rho_1 - \rho_2)\mathrm{d}x}{\mathrm{Vol}(O_k)} + \sum_{j=k+1}^{N_c} \pm\beta_j \frac{\mathrm{Vol}(I_k \cap O_j)}{\mathrm{Vol}(O_k)},
$$

and in particular

$$
\beta_k \le \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} \frac{\mathrm{Vol}(I_k)}{\mathrm{Vol}(O_k)} + \sum_{j:k<j\le N_c} \beta_j \frac{\mathrm{Vol}(I_k \cap O_j)}{\mathrm{Vol}(O_k)}, \quad \forall k = 2, \dots, N_c.
\tag{2.7}
$$

For every $k = 2, \dots, N_c$, we claim that

$$
\beta_k \le \|\rho_1 - \rho_2\|_\infty \left(\sum a_{j_1 \dots j_s} \right),
\tag{2.8}
$$

where the sum is taken over all $s \le N_c - k$ and all s-tuples $N_c \ge j_1 > j_2 > \cdots > j_{s-1} > j_s = k$, and where

$$
a_{j_1 \dots j_s} := \frac{\mathrm{Vol}(I_{j_1})}{\mathrm{Vol}(O_{j_1})} \cdot \frac{\mathrm{Vol}(I_{j_2} \cap O_{j_1})}{\mathrm{Vol}(O_{j_2})} \cdots \frac{\mathrm{Vol}(I_{j_{s-1} \cap O_{j_{s-2}}})}{\mathrm{Vol}(O_{j_{s-1}})} \cdot \frac{\mathrm{Vol}(I_{j_s} \cap O_{j_{s-1}})}{\mathrm{Vol}(O_{j_s})}.
$$

In fact, relation (2.8) is obtained inductively by using recursion (2.7) and the fact that $\beta_{N_c} \le \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} \frac{\mathrm{Vol}(I_{N_c})}{\mathrm{Vol}(O_{N_c})}$. Let us now fix s with $0 \le s \le N_c - k$ and k' with $k + s \le k' \le N_c$; set $j_1 = k'$ and $j_s = k$. Let us write $a_{j_1 \dots j_s}$ in the more convenient way:

$$
a_{j_1 \dots j_s} = \frac{\mathrm{Vol}(I_{j_1})}{\mathrm{Vol}(O_{j_s})} \cdot \frac{\mathrm{Vol}(I_{j_2} \cap O_{j_1})}{\mathrm{Vol}(O_{j_1})} \cdots \frac{\mathrm{Vol}(I_{j_{s-1} \cap O_{j_{s-2}}})}{\mathrm{Vol}(O_{j_{s-2}})} \cdot \frac{\mathrm{Vol}(I_{j_s} \cap O_{j_{s-1}})}{\mathrm{Vol}(O_{j_{s-1}})}.
$$

Note that

$$
a_{j_1 \dots j_s} \le \frac{\mathrm{Vol}(I_{j_1})}{\mathrm{Vol}(O_{j_s})} \cdot \frac{\mathrm{Vol}(I_{j_2} \cap O_{j_1})}{\mathrm{Vol}(O_{j_1})} \cdots \frac{\mathrm{Vol}(I_{j_{s-1} \cap O_{j_{s-2}}})}{\mathrm{Vol}(O_{j_{s-2}})},
$$

and therefore summing over j_{s-1} we obtain

$$
\sum_{j_{s-1}} a_{j_1 \ldots j_s} \leq \frac{\text{Vol}(I_{j_1})}{\text{Vol}(O_{j_s})} \frac{\text{Vol}(I_{j_2} \cap O_{j_1})}{\text{Vol}(O_{j_1})} \cdots \frac{\text{Vol}(I_{j_{s-2}} \cap O_{j_{s-3}})}{\text{Vol}(O_{j_{s-3}})}
$$
$$
\sum_{j_{s-1}} \frac{\text{Vol}(I_{j_{s-1}} \cap O_{j_{s-2}})}{\text{Vol}(O_{j_{s-2}})}.
$$

Observe that the sum on the right-hand side of the above expression is less than one because the sets $I_{j_{s-1}}$ are disjoint. Proceeding in this fashion adding over $j_{s-2}, \ldots, j_2$, we conclude that

$$
\sum_{j_2 \ldots j_{s-1}} a_{j_1 \ldots j_s} \leq \frac{\text{Vol}(I_{k'})}{\text{Vol}(O_k)}.
$$

Finally, first summing over all such s and then over all such k', it follows from (2.8) that

$$
\beta_k \leq \|\rho_1 - \rho_2\|_\infty \sum a_{j_1, \ldots, j_s} \leq \|\rho_1 - \rho_2\|_\infty \sum_{k < k' \leq N_c} \frac{N_c \text{Vol}(I_{k'})}{\text{Vol}(O_k)}
$$
$$
\leq \|\rho_1 - \rho_2\|_\infty \frac{N_c \text{Vol}(\mathcal{M})}{\text{Vol}(O_k)}
\tag{2.9}
$$

where in the last inequality we have used the fact that the sets $I_{k'}$ are disjoint.

Going back to (2.5), we note that from (2.6) and (2.9) it follows that for every $k = 1, \ldots, N_c$

$$
\|\gamma_k^+ - \tilde{\rho}_k\|_{L^\infty(B_\mathcal{M}(y_k, 2r))} \leq \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} + \sum_{j=k}^{N_c} \beta_j
$$
$$
\leq \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} \left(1 + N_c^2 \text{Vol}(M) \max_{j=2, \ldots, N_c} \frac{1}{\text{Vol}(O_j)} \right)
\tag{2.10}
$$
$$
\leq \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} \left(1 + \frac{C_m N_c^2 \text{Vol}(M)}{r^m} \right),
$$

where the last inequality follows from the lower bound on the size of the overlaps (2.2).

Now, we notice that from the standing assumption $\rho_1(x), \rho_2(x) \geq \frac{1}{\alpha}$ for every $x \in \mathcal{M}$, it follows that for every $k = 1, \ldots, N_c$ and every $x \in \mathcal{M}$

$$
\gamma_k^+(x), \gamma_k^-(x), \tilde{\rho}_k(x) \geq \frac{1}{\alpha} \quad \text{for all } x \in \mathcal{M}.
$$

Likewise, from the standing assumption $\rho_1(x), \rho_2(x) \leq \alpha$ for all $x \in \mathcal{M}$, it follows that for every $k = 1, \ldots, N_c$ and every $x \in \mathcal{M}$

$$\gamma_k^+(x), \gamma_k^-(x), \tilde{\rho}_k(x) \leq \alpha + \sum_{j=1}^{N_c} \beta_j \leq \alpha + \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} \frac{C_m N_c^2 \mathrm{Vol}(\mathcal{M})}{r^m}$$

Assume for a moment that $\|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})}$ is small enough so that in particular $\|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} \frac{C_m N_c^2 \mathrm{Vol}(\mathcal{M})}{r^m} \leq \alpha$. In that case, for every $k = 1, \ldots, N_c$ we would have

$$\frac{1}{\alpha} \leq \gamma_k^+, \quad \tilde{\rho}_k \leq 2\alpha. \tag{2.11}$$

Consider the exponential map $\exp_{y_k}: B(2r) \subseteq T_{y_k}\mathcal{M} \to B_\mathcal{M}(y_k, 2r) \subseteq \mathcal{M}$ and the functions $g_1, g_2: B(2r) \to (0, \infty)$ defined as

$$g_1(v) := \gamma_k^+(\exp_{y_k}(v)) J_{y_k}(v)$$

and

$$g_2(v) := \tilde{\rho}_k(\exp_{y_k}(v)) J_{y_k}(v),$$

where J_{y_k} denotes the Jacobian of the exponential map. From (2.11), (2.10) and (1.34), we conclude that

$$\frac{1}{\alpha C m(1 + K r^2)} \leq g_i(v)$$
$$\leq \alpha C m(1 + K r^2) \quad \text{for } i = 1, 2 \text{ and all } v \in B(2r) \tag{2.12}$$

and that for all $v \in B(2r)$

$$|g_1(v) - g_2(v)| \leq (1 + C m K r^2) \left| \gamma_k^+(\exp_{y_{N_c}}(v)) - \tilde{\rho}_k(\exp_{y_{N_c}}(v)) \right|$$
$$\leq \frac{C_m N_c^2 \mathrm{Vol}(\mathcal{M})}{r^m} \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} \tag{2.13}$$

We recall that our choice of r in particular guarantees that $r^2 K \leq 1$. Applying [11, Theorem 1.2] to the densities g_1 and g_2 with the bounds given by (2.12), we conclude that

$$d_\infty(g_1, g_2) \leq C_{m,\alpha} r \|g_1 - g_2\|_{L^\infty(B(2r))} \leq \frac{C_{m,\alpha} N_c^2 \mathrm{Vol}(\mathcal{M})}{r^{m-1}} \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})},$$

[1] where the last inequality follows from (2.13). From the second part of Proposition 1, it follows that

$$d_\infty(\gamma_k^+, \tilde{\rho}_k) \le 2d_\infty(g_1, g_2) \le \frac{C_{m,\alpha} N_c^2 \text{Vol}(\mathcal{M})}{r^{m-1}} \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})}.$$

Therefore, using (2.4) it follows that if

$$\|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} \frac{C_m N_c^2 \text{Vol}(\mathcal{M})}{r^m} \le \alpha,$$

then

$$d_\infty(\rho_1, \rho_2) \le \frac{C_{m,\alpha} N_c^3 \text{Vol}(\mathcal{M})}{r^{m-1}} \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})}.$$

In case $\|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})} \frac{C_m N_c^2 \text{Vol}(\mathcal{M})}{r^m} > \alpha \ge 1$, we have

$$d_\infty(\rho_1, \rho_2) \le \text{diam}(\mathcal{M})$$
$$\le \frac{C_m N_c^2 \text{Vol}(\mathcal{M}) \, \text{diam}(\mathcal{M})}{r^m} \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})},$$

where we note that the first inequality in the above expression is always true, as the maximum distance any point can travel in $\mathcal{M}$ is $\text{diam}(\mathcal{M})$. Therefore, in any case we have

$$d_\infty(\rho_1, \rho_2) \le \tilde{C} \|\rho_1 - \rho_2\|_{L^\infty(\mathcal{M})},$$

where $\tilde{C}$ can be written as

$$\tilde{C} = \frac{C_{m,\alpha} N_c^2 \text{Vol}(\mathcal{M})}{r^{m-1}} \max \left\{ N_c, \frac{\text{diam}(\mathcal{M})}{r} \right\}. \tag{2.14}$$

$\square$

2.1 Proof of Theorem 2

In the following, we consider the Voronoi tessellation induced by the set $Y = \{y_1, \ldots, y_{N_c}\}$ constructed in the beginning of Sect. 2, i.e., for each $i \in \{1, \ldots, N_c\}$ we define

$$V_\mathcal{M}(y_i) := \{x \in \mathcal{M} : d(x, y_i) \le d(x, y_j) \text{ for all } j \in \{1, \ldots, N_c\}\}.$$

---

[1] Note that as stated, our theorems give $C_{m,\alpha,r}$, but in this case $C_{m,\alpha,r} = C_{m,\alpha}r$ because we can always rescale to the unit ball.

These measurable sets form a partition of $\mathcal{M}$ up to a negligible set of ambiguity of measure zero. We make use of the following.

Proposition 3 *For each $i \in \{1, \ldots, N_c\}$, there exists a bi-Lipschitz bijection $\Psi_i : V_{\mathcal{M}}(y_i) \to \overline{B}\left(0, \frac{r}{2}\right) \subseteq \mathbb{R}^m$ with bi-Lipschitz constant at most 18.*

To prove Proposition 3, we use the sequence of lemmas that follow.

Lemma 2 *For all $i \in \{1, \ldots, N_c\}$*

$$B_{\mathcal{M}}(y_i, r/2) \subset V_{\mathcal{M}}(y_i) \subset B_{\mathcal{M}}(y_i, r).$$

Let $V(y_i) = \exp_{y_i}^{-1}(V_{\mathcal{M}}(y_i))$. Then, $B\left(0, \frac{r}{2}\right) \subset V(y_i) \subset B(r)$ and for almost every $z_0 \in \partial V(y_i)$

$$\frac{z_0}{|z_0|} \cdot n_0 \geq \frac{1}{8}, \tag{2.15}$$

where n_0 is the outward unit normal vector to $\partial V(y_i)$ at z_0.

Proof Let $y_i \in Y$. Since for every $x \in B_{\mathcal{M}}\left(y_i, \frac{r}{2}\right)$ and every $y_j \in Y$ with $j \neq i$ it holds that $d(x, y_j) \geq d(y_j, y_i) - d(y_i, x) > \frac{r}{2}$, we conclude that $B_{\mathcal{M}}\left(y_i, \frac{r}{2}\right) \subset V_{\mathcal{M}}(y_i)$. On the other hand, since Y is a maximal set with the property that $d(y_j, y_k) \geq r$ for all $j \neq k$, we conclude that for all $x \in \mathcal{M}$ there exists $y_j \in X$ such that $d(x, y_j) < r$. Therefore, $V_{\mathcal{M}}(y_i) \subset B_{\mathcal{M}}(y_i, r)$. Since $\exp_{y_i}$ maps $B(s)$ bijectively to $B_{\mathcal{M}}(y_i, s)$ for $s = \frac{r}{2}$ and for $s = r$, it follows that $B\left(0, \frac{r}{2}\right) \subset V(y_i) \subset B(r)$. This establishes the first part of the statement.

Now, let us consider the second part of the statement. For almost every $z_0 \in \partial V(y_i)$, there exists a unique $y_j \neq y_i$ such that $z_0 \in \partial \exp_{y_i}^{-1}(V_{\mathcal{M}}(y_j))$; this follows from the fact that the set of triple junctions or higher-order junctions in a Voronoi tessellation has Hausdorff dimension smaller than $m - 2$. Let us fix one such $z_0 \in V(y_i)$ and let y_j be as before. Note that $2r \geq d(y_i, y_j) \geq r$ and that $d(y_i, z) = |z_0|_{y_i} < r$. We let $z := \exp_{y_i}(z_0)$. We consider the level set $\Gamma := \{x \in \mathcal{M} : d(x, y_i) = d(x, y_j)\}$, which is a C^1-hypersurface around z by the implicit function theorem; moreover a unit normal vector to Γ at the point z is given by

$$n := \frac{\tilde{u}_i - \tilde{u}_j}{\left|\tilde{u}_i - \tilde{u}_j\right|_z} = \frac{u_i - u_j}{\left|u_i - u_j\right|_z},$$

where $\tilde{u}_i := -\frac{\exp_z^{-1}(y_i)}{d(y_i, z)}$, $u_i := -\exp_z^{-1}(y_i)$ and $u_j, \tilde{u}_j$ are defined analogously.

Let us consider the set $\Gamma_0 := \exp_{y_i}^{-1}(\Gamma \cap B_{\mathcal{M}}(y_i, 2r))$; note that around the point z_0, Γ_0 coincides with $\partial V(y_i)$, and in particular given that Γ is a C^1-hypersurface around z, $\partial V(y_i)$ is a C^1-hypersurface around z_0. Let us denote by n_0 the outward unit normal to $\partial V(y_i)$ at z_0. We write $\frac{z_0}{|z_0|_{y_i}}$ as

$$\frac{z_0}{|z_0|_{y_i}} = w_0 + c n_0,$$

where $\langle w_0, \frac{z_0}{|z_0|_{y_i}} \rangle_{y_i} = 0$ and $\langle n_0, \frac{z_0}{|z_0|_{y_i}} \rangle_{y_i} = c$. Clearly, $c \geq 0$. Now, by definition of the exponential map, $\tilde{u}_i = (d \exp_{y_i})_{z_0} \left(\frac{z_0}{|z_0|_{y_i}} \right)$, and so

$$\tilde{u}_i = w + c\tilde{n},$$

where $w := (d \exp_{y_i})_{z_0}(w_0)$ and $\tilde{n} := (d \exp_{y_i})_{z_0}(n_0)$. Then,

$$\langle \tilde{u}_i, n \rangle_z = \langle w + c\tilde{n}, n \rangle_z = c\langle \tilde{n}, n \rangle_z \leq c |\tilde{n}|_z \leq 2c |n_0|_{y_i} = 2c,$$

where the second equality follows from the fact that w is tangent to Γ (which in turn follows from the fact that w_0 is tangent to Γ_0) and where the last inequality follows from (1.33). It thus remains to show that $\langle \tilde{u}_i, n \rangle_z \geq 1/4$. To see this, simply note that the fact that $\langle \tilde{u}_i + \tilde{u}_j, \tilde{u}_i - \tilde{u}_j \rangle_z = 0$ implies

$$\langle \tilde{u}_i, n \rangle_z = \left\langle \frac{\tilde{u}_i - \tilde{u}_j}{2}, \frac{\tilde{u}_i - \tilde{u}_j}{|\tilde{u}_i - \tilde{u}_j|_z} \right\rangle_z = \frac{|\tilde{u}_i - \tilde{u}_j|_z}{2} = \frac{|u_i - u_j|_z}{2d(z, y_i)} \geq \frac{d(y_i, y_j)}{4d(z, y_i)} \geq \frac{1}{4},$$

where the second to last inequality follows from Proposition 1. □

So far, we have been able to construct a partition of $\mathcal{M}$ into cells (the Voronoi cells $V_{\mathcal{M}}(y_i)$) with the property that when each of the cells $V_{\mathcal{M}}(y_i)$ is mapped by the inverse of the exponential map, the resulting set $V_i := \exp_{y_i}^{-1}(V_{\mathcal{M}}(y_i))$ (which is contained in $\mathbb{R}^m$) is a star-shaped domain with center the origin. We notice that condition (2.15) implies that V_i is a star-shaped domain centered at the origin. Indeed, if the set was not star-shaped with center the origin, we would be able to find a point $z_0 \in \partial V_i$ for which $n_0 \cdot \frac{z_0}{|z_0|} \leq 0$ (i.e., the outer normal at that point would be pointing toward the origin). In the next lemma, we show that when in addition the unit outer normal to the boundary of a star-shaped domain does not deviate too much from the radial direction emanating from its center, the domain is bi-Lipschitz homeomorphic to a ball and more importantly the bi-Lipschitz constant can be controlled. This establishes Proposition 3.

Lemma 3 *Let V be a star-shaped subset of $\mathbb{R}^m$ with center at 0 and such that $B(R) \subset V \subset B(2R)$. Assume V has Lipschitz boundary and let n be the unit outside normal vector to ∂V. Assume there exists $\beta \in (0, 1)$ such that for a.e. $x \in \partial V$*

$$n \cdot \frac{x}{|x|} \geq \beta.$$

Let $r: S^{m-1} \to [R, 2R]$ be the function describing ∂V in radial coordinates. That is, let $r(z) = \sup\{s \in \mathbb{R} : sz \in V\}$. Consider the function $\Phi: V \to \overline{B(R)}$ given by

$$\Phi(x) = \frac{R}{r\left(\frac{x}{|x|}\right)} x \quad \text{for } x \neq 0$$

and $\Phi(0) = 0$. Then, Φ is a bi-Lipschitz bijection with bi-Lipschitz constant at most $\frac{1}{\beta} + 1$.

Proof Extend r to $\mathbb{R}^m \setminus \{0\}$ by $\tilde{r}(x) := r\left(\frac{x}{|x|}\right)$.

For $x \neq 0$

$$D\Phi(x) = -\frac{R}{\tilde{r}^2(x)} x(\nabla \tilde{r}(x))^T + \frac{R}{\tilde{r}(x)} I. \tag{2.16}$$

Consider the function $G \colon \mathbb{R}^m \setminus \{0\} \to \partial V$ given by $x \mapsto \tilde{r}(x)\frac{x}{|x|}$. Note that at $z \in S^{m-1}$

$$DG(z) = z(\nabla \tilde{r}(z))^T + r(z)\left(I - zz^T\right).$$

Since n is orthogonal to the image of G, we conclude that $(DG(z))^T n = 0$, which implies

$$(n \cdot z)\nabla \tilde{r}(z) + r(z)(n - (n \cdot z)z) = 0.$$

Since $n \cdot z \geq \beta$, we obtain

$$\beta |\nabla \tilde{r}(z)| \leq \tilde{r}(z) \quad \text{for all } z \in S^{m-1}.$$

Combining this with (2.16), we deduce that Φ is $(\frac{1}{\beta} + 1)$-Lipschitz. Analogous computations show that Φ^{-1}, which is given by $\Phi^{-1}(y) = r\left(\frac{y}{|y|}\right)y$, is also $(\frac{1}{\beta} + 1)$-Lipschitz. $\qquad \square$

Proof *(Proposition 3)* By Proposition 1, the exponential map $\exp_{y_i} \colon B(r) \to B_{\mathcal{M}}(y_i, r)$ is a bi-Lipschitz bijection with bi-Lipschitz constant at most 2. By Lemmas 2 and 3, with $R = \frac{r}{2}$ and $\beta = \frac{1}{8}$, there exists a mapping

$$\Psi_i \colon \exp_{y_1}^{-1}(V_{\mathcal{M}}(y_i)) \to \overline{B\left(0, \frac{r}{2}\right)}$$

which is a bi-Lipschitz bijection with bi-Lipschitz constant at most 9. The composition $\Psi_i \circ \exp_{y_i}^{-1}$ provides the desired mapping. $\qquad \square$

Proof *(Theorem 2)* We consider the maps $\Psi_i \colon V_{\mathcal{M}}(y_i) \to \overline{B(r/2)} \subseteq \mathbb{R}^m$ from Proposition 3. Given the sample $x_1, \ldots, x_n$ from the density p, we define a function $p_n \colon \mathcal{M} \to \mathbb{R}$ by

$$p_n(x) := p(x) + \frac{\mu_n(V_{\mathcal{M}}(y_i)) - \mu(V_{\mathcal{M}}(y_i))}{\text{Vol}(V_{\mathcal{M}}(y_i))} \quad \text{for } x \in V_{\mathcal{M}}(y_i). \tag{2.17}$$

Notice that

$$\int_{\mathcal{M}} p_n(x)\mathrm{dVol}(x) = \sum_{i=1}^{n} \int_{V_{\mathcal{M}}(y_i)} p_n(x)\mathrm{dVol}(x) = \sum_{i=1}^{n} \mu_n(V_{\mathcal{M}}(y_i)) = 1,$$

and that as shown in (2.18), with very high probability p_n is positive. In particular, with very high probability p_n is a bona fide density function.

Let us recall that Hoeffding's inequality states that for every $t > 0$,

$$\mathbb{P}\left(|\mu_n(V_{\mathcal{M}}(y_i)) - \mu(V_{\mathcal{M}}(y_i))| > t\right) \le 2e^{-2nt^2}.$$

Using the previous concentration inequality, we conclude that for every $i = 1, \ldots, N_c$

$$\|p - p_n\|_{L^{\infty}(V_{\mathcal{M}}(y_i))} \le \frac{1}{2\alpha}$$

with probability at least $1 - 2\exp\left(-n\frac{\mathrm{Vol}(V_{\mathcal{M}}(y_i))^2}{2\alpha^2}\right)$ In particular, using a union bound, we conclude that with probability at least $1 - 2N_c \exp\left(-n\frac{C_m r^{2m}}{\alpha^2}\right)$

$$\frac{1}{2\alpha} \le p_n(x) \le 2\alpha, \quad x \in \mathcal{M}. \tag{2.18}$$

Similarly, with probability at least $1 - 2N_c \exp\left(-n\frac{C_m r^{2m}}{\alpha^2}\right)$

$$\frac{1}{2}\mu(V_{\mathcal{M}}(y_i)) \le \mu_n(V_{\mathcal{M}}(y_i)) \le \frac{3}{2}\mu(V_{\mathcal{M}}(y_i)) \tag{2.19}$$

Hoeffding's inequality together with a union bound also shows that with probability at least $1 - 2N_c n^{-\beta}$,

$$\|p - p_n\|_{L^{\infty}(\mathcal{M})} \le \frac{C_m}{r^m} \sqrt{\frac{\beta \log(n)}{n}}. \tag{2.20}$$

We let A_n be the event, where (2.18), (2.19) and (2.20) hold. From the above, we know that A_n occurs with probability at least $1 - C n^{-\beta}$, where the constant C depends on $r, \alpha, \beta, m, \mathrm{Vol}(\mathcal{M})$. We denote by $\tilde{\mu}_n$ the measure $d\tilde{\mu}_n = p_n dx$. Conditioned on the event A_n, we see from Lemma 1 and from (2.20) that

$$d_{\infty}(\tilde{\mu}_n, \mu) \le \tilde{C}\|p - p_n\|_{L^{\infty}(\mathcal{M})} \le \tilde{C}\frac{C_m}{r^m}\sqrt{\frac{\beta \log(n)}{n}},$$

where $\tilde{C}$ is the constant in (2.14).

Now, we estimate $d_{\infty}(\tilde{\mu}_n, \mu_n)$ in the event A_n. Observe that

$$\tilde{\mu}_n(V_{\mathcal{M}}(y_i)) = \mu_n(V_{\mathcal{M}}(y_i)) \quad \text{for all } i = 1, \ldots, N_c$$

and hence

$$d_\infty(\mu_n, \tilde{\mu}_n) \le \max_{i=1,\ldots,N_c} d_\infty(\mu_n \llcorner V_{\mathcal{M}}(y_i), \tilde{\mu}_n \llcorner V_{\mathcal{M}}(y_i)),$$

where we denote by $\llcorner V_{\mathcal{M}}(y_i)$ the restriction of a measure to $V_{\mathcal{M}}(y_i)$. The goal is now to estimate $d_\infty(\mu_n \llcorner V_{\mathcal{M}}(y_i), \tilde{\mu}_n \llcorner V_{\mathcal{M}}(y_i))$ for every i.

Let $x_{j_1}, \ldots, x_{j_{n_i}}$ be the points in X that fall in $V_{\mathcal{M}}(y_i)$. We consider the transformed points $\Psi_i(x_{j_1}), \ldots, \Psi_i(x_{j_{n_i}})$ and the measure $\Psi_{i\sharp}(\tilde{\mu}_n \llcorner V_{\mathcal{M}}(y_i))$, which is supported on $\overline{B(r/2)}$. The fact that Ψ_i is bi-Lipschitz with constant 18 implies that the measure $\Psi_{i\sharp}(\tilde{\mu}_n \llcorner V_{\mathcal{M}}(y_i))$ has a density with respect to the Lebesgue measure, and this density is lower and upper bounded by constant multiples of the lower and upper bounds of the density p. Hence, the transformed points are almost surely samples from $\Psi_{i\sharp}(\tilde{\mu}_n \llcorner V_{\mathcal{M}}(y_i))$ restricted to the open ball $B(r/2)$. Therefore, it follows from [11, Theorem 1.1] that conditioned on the event A_n,

$$d_\infty\left(\Psi_{i\sharp}(\tilde{\mu}_n \llcorner V_{\mathcal{M}}(y_i)), \Psi_{i\sharp}(\mu_n \llcorner V_{\mathcal{M}}(y_i))\right) \le C_{m,\alpha,\beta} \, r \, \frac{\log(n_i)^{p_m}}{n_i^{1/m}}$$

holds[2] for all $i \in \{1, \ldots, N_c\}$ with probability at least $1 - CN_c n^{-\beta}$, where C is a constant that depends on β, r, α, m. Note that we have used the fact that in the event A_n, the second inequality in (2.19) is satisfied and so we can give the probability bounds in terms of n and not in terms of n_i. Moreover, from the first inequality in (2.19) it follows that

$$\frac{\log(n_i)^{p_m}}{n_i^{1/m}} \le C_m \frac{\alpha^{1/m}}{r} \frac{(\log(n))^{p_m}}{n^{1/m}}.$$

Finally, from the fact that Ψ_i^{-1} is Lipschitz with Lipschitz constant no larger than 18, it follows that

$$d_\infty(\tilde{\mu}_n \llcorner V_{\mathcal{M}}(y_i), \mu_n \llcorner V_{\mathcal{M}}(y_i)) \le 18 d_\infty\left(\Psi_{i\sharp}(\tilde{\mu}_n \llcorner V_{\mathcal{M}}(y_i)), \Psi_{i\sharp}(\mu_n \llcorner V_{\mathcal{M}}(y_i))\right).$$

From the previous discussion, we deduce that with probability at least $1 - CN_c n^{-\beta} = 1 - C_{m,\beta,\alpha,r,\mathrm{Vol}(\mathcal{M})} \cdot n^{-\beta}$,

$$d_\infty(\mu, \mu_n) \le d_\infty(\mu, \tilde{\mu}_n) + d_\infty(\tilde{\mu}_n, \mu_n)$$
$$\le C'\left(\sqrt{\frac{\log(n)}{n}} + \frac{(\log(n))^{p_m}}{n^{1/m}}\right) \le C' \frac{(\log(n))^{p_m}}{n^{1/m}}$$

for a constant C' that can be written as $C' = \frac{C_{\alpha,\beta,m}}{r^m}\tilde{C}$, where $\tilde{C}$ is as in (2.14). $\qquad\square$

---

[2] Note that as stated, Theorem 1.1 in [11] gives $C_{m,\alpha,\beta,r}$, but in this case $C_{m,\alpha,\beta,r} = C_{m,\alpha,\beta} r$ as one can simply rescale to the unit ball.

3 Kernel-Based Approximation of the Laplacian

Here, we focus on a kernel-based approximation of the continuous Dirichlet form defined in (1.8). This part does not depend on the graph obtained from the sample set X and can be seen as the bias part of the desired error estimates.

The results in this section correspond to those of Section 3 and 5 in [6] but cannot be directly inferred from them. Instead, we need to adjust most of the proofs to our setting.

For $f \in L^2(\mathcal{M})$, $0 < r < 2h$ and a Borel set $V \subseteq \mathcal{M}$ let

$$E_r(f, V) := \int_V \int_\mathcal{M} \eta\left(\frac{d(x, y)}{r}\right) |f(y) - f(x)|^2 \, d\mu(y) d\mu(x). \qquad (3.1)$$

We write $E_r(f)$ shorthand for $E_r(f, \mathcal{M})$. The main results of this section, Lemma 5 and 9, demonstrate how this functional approximates the form Δ.

Remark 11 Let $\tilde{E}_r(f, V)$ denote the functional in (3.1) when η is taken to be the kernel $\mathbb{1}_{[0,1]}$. Then, $\tilde{E}_r(f, V)$ is nothing but $E_r(f, V)$ as defined in [6, Def. 3.1]. Note that, for general η satisfying the assumptions from Sect. 1.1

$$\tilde{E}_r(f, V) \leq \frac{1}{\eta(1/2)} E_{2r}(f, V). \qquad (3.2)$$

for every $f \in L^2(\mathcal{M})$ and any Borel set $V \subseteq \mathcal{M}$.

Lemma 4 *Suppose h satisfies Assumptions 3. Then, there exists a universal constant $C > 0$ such that for every $0 < r < 2h$ and every $f \in L^2(\mathcal{M}, \mu)$*

$$E_r(f) \leq C2^m(1 + \alpha L_p)E_{r/2}(f),$$

where we recall that L_p stands for the Lipschitz constant of p.

Proof Let $0 < r < 2h$. Then, $r \leq \min\{i_0, 1/\sqrt{K}\}$ by Assumptions 3. Note that it suffices to consider f to be smooth because smooth functions are dense in $L^2(\mathcal{M}, \mu)$ and both sides of the inequality are continuous with respect to L^2-convergence; notice that for smooth functions we can talk about pointwise values. For $x, y \in \mathcal{M}$ with $d(x, y) \leq r$, let z_{xy} be the point in $\mathcal{M}$ which lies halfway along the geodesic connecting x and y, i.e., $z_{xy} = \exp_x(\frac{1}{2}\exp_x^{-1}(y))$. In particular $d(x, z_{xy}) = d(y, z_{x,y}) = \frac{1}{2}d(x, y)$. Since $|f(x) - f(y)|^2 \leq 2|f(x) - f(z_{x,y})|^2 + 2|f(y) - f(z_{x,y})|^2$, by

symmetry we obtain

$$
\begin{aligned}
E_r(f) &\leq 4 \int_{\mathcal{M}} \int_{\mathcal{M}} \eta\left(\frac{d(x,y)}{r}\right) \left|f(x) - f(z_{x,y})\right|^2 d\mu(y)d\mu(x) \\
&= 4 \int_{\mathcal{M}} \int_{B(r)} \eta\left(\frac{|v|}{r}\right) \left|f(x) - f\left(\exp_x\left(\frac{v}{2}\right)\right)\right|^2 J_x(v) p(\exp_x(v))dvd\mu(x) \\
&\leq C2^m(1+\alpha L_p) \int_{\mathcal{M}} \int_{B(\frac{r}{2})} \eta\left(\frac{2|w|}{r}\right) \left|f(x) - f(\exp_x(w))\right|^2 J_x(w) \\
&\qquad p(\exp_x(w))dwd\mu(x) \\
&= C2^m(1+\alpha L_p) E_{r/2}(f),
\end{aligned}
$$

where C is a universal constant. In the above, we used the change in variables $w = \frac{v}{2}$ (which explains the term 2^m) and we also used the inequalities:

$$
J_x(v) \leq (1 + CmKr^2)^2 J_x\left(\frac{v}{2}\right) \leq C J_x\left(\frac{v}{2}\right),
$$

(combined with Assumptions 3) and

$$
p(\exp_x(v)) \leq (1 + \alpha L_p) p(\exp_x(v/2)).
$$

$\square$

Lemma 5 (cf. [6, Lemma 3.3]) *Suppose h satisfies Assumptions 3. Then, there exists a universal constant $C > 0$ such that*

$$
E_r(f) := E_r(f, \mathcal{M}) \leq (1 + L_p \alpha r) \cdot (1 + CmKr^2)\sigma_\eta r^{m+2} D(f),
$$

for every $f \in H^1(\mathcal{M})$ and $0 < r < 2h$.

Proof Let us first consider the case in which η takes the form $\eta = \mathbb{1}_{[0,1]}$; as we will see, the general case follows easily from this special case. As in [6, Lemma 3.3], we may assume that f is smooth and we write

$$
\int_{B_{\mathcal{M}}(x,r)} |f(y) - f(x)|^2 d\mu(y) = \int_{B(r)} \left|f(\exp_x(v)) - f(x)\right|^2 p(\exp_x(v)) J_x(v)dv
$$

where J_x denotes the determinant of the Jacobian of the exponential map. We recall from (1.34) that there exists a constant $C > 0$ such that $J_x(v)$ is bounded from above by $1 + CmKr^2$ for all $v \in B(r)$. From the fundamental theorem of calculus, it follows that

$$
\left|f(\exp_x(v)) - f(x)\right|^2 \leq \int_0^1 \left|\frac{d}{dt}f(\exp_x(tv))\right|^2 dt = \int_0^1 |df(\Phi_t(x,v))|^2 dt,
$$

In the above, Φ_t denotes the time t geodesic flow, $\Phi_t(x, v) = (\gamma_{x,v}(t), \gamma'_{x,v}(t))$, where $\gamma_{x,v}(t) := \exp_x(tv)$. The expression $df(\Phi_t(x, v))$ has to be interpreted as: the form df at $\gamma_{x,v}(t)$ acting on the tangent vector $\gamma'_{x,v}(t)$. Therefore,

$$A := \int_{\mathcal{M}} \int_{B(r)} \left| f(\exp_x(v)) - f(x) \right|^2 p(\exp_x(v)) dv \, p(x) d\mathrm{Vol}(x)$$

$$\leq \int_0^1 \int_{\mathcal{M}} \int_{B(r)} |df(\Phi_t(x, v))|^2 \, p(\Phi_1(x, v)_1) p(\Phi_0(x, v)_1) dv d\mathrm{Vol}(x) dt$$

where $\xi \mapsto \xi_1$ denotes the projection of $\xi \in T\mathcal{M}$ on $\mathcal{M}$. From the Lipschitz continuity of p, it follows that $p(x) \leq (1 + L_p \alpha r) p(y)$ for all $x, y \in \mathcal{M}$ where $d(x, y) \leq r$. Using the fact that Φ_t preserves the canonical volume $\mathrm{Vol}_{T\mathcal{M}}$ on $T\mathcal{M}$ and that

$$\mathcal{B}_r := \{\xi = (x, v) \in T\mathcal{M} : |v| \leq r\}$$

is invariant under Φ_t, see [5, 1.125], we obtain after a change in variables

$$A \leq (1 + L_p \alpha r)^2 \int_0^1 \int_{\mathcal{B}_r} |df(\Phi_t(\xi))| \, p^2(\Phi_t(\xi)_1) d\mathrm{Vol}_{T\mathcal{M}}(\xi) dt$$

$$= (1 + L_p \alpha r)^2 \int_{\mathcal{B}_r} |df(\xi)|^2 \, p^2(\xi_1) d\mathrm{Vol}_{T\mathcal{M}}(\xi)$$

$$= (1 + L_p \alpha r)^2 \int_{\mathcal{M}} \frac{\omega_m}{m+2} r^{m+2} |\nabla f|^2 \, p^2(x) d\mathrm{Vol}(x).$$

Using the previous computations, (1.34) and Remark 1, we deduce that

$$E_r(f) \leq (1 + CmKr^2) \cdot A \leq (1 + CmKr^2) \cdot (1 + L_p \alpha r) \frac{\omega_m}{m+2} r^{m+2} D(f)$$

$$= (1 + CmKr^2) \cdot (1 + L_p \alpha r) \sigma_\eta r^{m+2} D(f)$$

$$(3.3)$$

for a universal constant C, which proves the claim for $\eta = \mathbb{1}_{[0,1]}$. Now, notice that one easily obtains from the previous computations that (3.3) is still valid for η of the form $\eta = \mathbb{1}_{[0,t]}$ for some $0 < t < 1$. Finally, since $E_r(f)$ and σ_η are linear in η, the statement holds if $\eta: [0, 1] \to [0, \infty)$ is a decreasing step function (and hence can be written as linear combination of functions of the form $\mathbb{1}_{[0,t]}$). By monotone convergence applied on both sides of the inequality, the assertion follows for any decreasing (and thus measurable) function η. $\qquad \square$

Remark 12 Note that in comparison with the case of constant p treated in [6, Lemma 3.3], the above estimates have the additional term $(1 + \alpha L_p r)$.

Lemma 6 (cf. [6, Lemma 3.4]) *Suppose h satisfies Assumptions 3. Let $\varepsilon < r < 2h$, $f \in L^2(\mathcal{M})$ and $V \subseteq \mathcal{M}$ a Borel set such that $\mu(V) > 0$ and $\mathrm{diam}(V) \leq 2\varepsilon$. Then,*

$$\int_V \left| f(x) - \frac{1}{\mu(V)} \int_V f \mathrm{d}\mu \right|^2 \mathrm{d}\mu(x) \leq \frac{2(1 + CmKr^2)}{\eta(1/2)\omega_m(r-\varepsilon)^m} E_{2r}(f, V).$$

Proof The proof is almost identical to the proof of [6, Lemma 3.4], replacing the volume with the measure μ and taking Remark 11 into account. □

Next, we define a smoothening operator $\Lambda \colon L^2(\mathcal{M}, \rho\mu) \to \mathrm{Lip}(\mathcal{M})$ similar to the one introduced in [6, Section 5] but adapted to the kernel η. To this end, we first define a mapping $\psi \colon [0, \infty) \to [0, \infty)$ by

$$\psi(t) := \frac{1}{\sigma_\eta} \int_t^\infty \eta(s) s \mathrm{d}s.$$

Note that, as η is supported on $[0, 1]$, $\psi(t) = 0$ for all $t \geq 1$.

Remark 13 We remark that for $\eta(t) = \mathbb{1}_{[0,1]}(t)$, the above ψ coincides with the kernel function used in [6, Section 5].

For every $r > 0$, we define the operator $\Lambda_r^0 \colon L^2(\mathcal{M}, \mathrm{Vol}) \to \mathrm{Lip}(\mathcal{M})$ by

$$(\Lambda_r^0 f)(x) := \int_{\mathcal{M}} f(y) k_r(x, y) \mathrm{dVol}(y) \tag{3.4}$$

where

$$k_r(x, y) := \frac{1}{r^m} \psi\left(\frac{d(x, y)}{r}\right).$$

As in [6, Definition 5.2], we define the smoothing operator $\Lambda_r \colon L^2(\mathcal{M}, \rho\mu) \to \mathrm{Lip}(\mathcal{M})$ by

$$\Lambda_r f(x) := (\theta(x))^{-1} \Lambda_r^0 f(x), \tag{3.5}$$

where $\theta := \Lambda_r^0 \mathbb{1}$. Note that the term θ is introduced so that Λ_r preserves constant functions.

Let us deduce some useful properties of the functions just introduced. Since $\psi'(s) = -\frac{1}{\sigma_\eta} \eta(s) s$ for all $s \geq 0$, we obtain from the mean value theorem that for any $0 \leq t \leq r$ there exists $\frac{t}{r} \leq s \leq 1$ such that

$$\frac{1}{r^m} \psi\left(\frac{t}{r}\right) = \frac{1}{\sigma_\eta r^m} \eta(s) s \left(1 - \frac{t}{r}\right).$$

Hence, by the monotonicity of η, we have

$$k_r(x, y) \leq \frac{1}{\sigma_\eta r^m} \eta\left(\frac{d(x, y)}{r}\right) \tag{3.6}$$

for every $x, y \in \mathcal{M}$. If $d(x, y) \leq r$, then the gradient of the kernel k_r can be written as

$$\begin{aligned}
\nabla k_r(\cdot, y)(x) &= \frac{1}{r^{m+1}} \psi'\left(\frac{d(x, y)}{r}\right) \frac{-\exp_x^{-1}(y)}{d(x, y)} \\
&= \frac{1}{\sigma_\eta r^{m+2}} \eta\left(\frac{d(x, y)}{r}\right) \exp_x^{-1}(y)
\end{aligned} \tag{3.7}$$

where we refer to [6, (2.6)] for the gradient of the distance function. Moreover, we have

$$\int_{\mathbb{R}^m} \psi(|x|) dx = 1. \tag{3.8}$$

To see this, first note that using polar coordinates we obtain

$$m\sigma_\eta = \sum_{i=1}^m \int_{\mathbb{R}^m} \eta(|x|) x_i^2 dx = \int_{\mathbb{R}^m} \eta(|x|) |x|^2 dx = m\omega_m \int_0^\infty \eta(r) r^{m+1} dr,$$

where ω_m is the volume of the Euclidean unit ball in $\mathbb{R}^m$. Thus, using integration by parts and polar coordinates, it follows that

$$\begin{aligned}
\int_{\mathbb{R}^m} \psi(|x|) dx &= m\omega_m \int_0^\infty \psi(r) r^{m-1} dr \\
&= -\omega_m \int_0^\infty \psi'(r) r^m dr \\
&= \frac{\omega_m}{\sigma_\eta} \int_0^\infty \eta(r) r^{m+1} dr = 1.
\end{aligned}$$

For $\theta(x) := \Lambda_r^0(\mathbb{1})$, we now obtain the following bounds.

Lemma 7 (cf. [6, Lemma 5.1]) *There exists an absolute constant $C > 0$ such that*

$$(1 + CmKr^2)^{-1} \leq \theta(x) \leq 1 + CmKr^2$$

and $|\nabla\theta(x)| \leq CmKr/\sigma_\eta$ for all $x \in \mathcal{M}$.

Proof We have

$$\theta(x) = \frac{1}{r^m} \int_{B_\mathcal{M}(x,r)} \psi\left(\frac{d(x, y)}{r}\right) d\text{Vol}(y) = \frac{1}{r^m} \int_{B(r)} \psi\left(\frac{|v|}{r}\right) J_x(v) dv.$$

Thus, the first assertion now follows from (1.34) and (3.8). Since (1.34) implies $|J_x(v) - 1| \leq CmK |v|^2$ and since

$$\int_{B(r)} \psi\left(\frac{|v|}{r}\right) v \, dv = 0$$

for symmetry reasons, the bound on the gradient of θ can be obtained from (3.7) as

$$
\begin{aligned}
|\nabla\theta(x)| &= \frac{1}{\sigma_\eta r^{m+2}} \left| \int_{B_{\mathcal{M}}(x,r)} \psi\left(\frac{d(x,y)}{r}\right) \exp_x^{-1}(y) d\text{Vol}(y) \right| \\
&= \frac{1}{\sigma_\eta r^{m+2}} \left| \int_{B(r)} \psi\left(\frac{|v|}{r}\right) v J_x(v) dv - \int_{B(r)} \psi\left(\frac{|v|}{r}\right) v \, dv \right| \\
&= \frac{1}{\sigma_\eta r^{m+2}} \left| \int_{B(r)} \psi\left(\frac{|v|}{r}\right) v (J_x(v) - 1) dv \right| \\
&\leq \frac{CmKr^3}{\sigma_\eta r^{m+2}} \int_{B(r)} \psi\left(\frac{|v|}{r}\right) dv = \frac{CmKr}{\sigma_\eta}.
\end{aligned}
$$

$\square$

In order to establish the following properties of Λ_r, we make use of the fact that the densities p and ρ are Lipschitz continuous and are bounded from below. Thus,

$$p(x) \leq (1 + L_p \alpha r) p(y) \quad \text{and} \quad \rho(x) \leq (1 + L_\rho \alpha r)\rho(y)$$

whenever $d(x, y) \leq r$.

Lemma 8 (cf. [6, Lemma 5.4]) *Suppose that h satisfies Assumptions 3. Then, there exists a universal constant $C > 0$ such that*

$$\|\Lambda_r f\|^2_{L^2(\mathcal{M},\rho\mu)} \leq (1 + \alpha L_p r)(1 + \alpha L_\rho r)(1 + CmKr^2)\|f\|^2_{L^2(\mathcal{M},\rho\mu)}$$

and

$$\|\Lambda_r f - f\|^2_{L^2(\mathcal{M},\rho\mu)} \leq \frac{C\alpha^2}{\sigma_\eta r^m} E_r(f)$$

for all $f \in L^2(\mathcal{M})$ and all $r < 2h$.

Proof The first assertion follows from Jensen's inequality,

$$
\begin{aligned}
\int_{\mathcal{M}} (\Lambda_r f(x))^2 \rho(x) d\mu(x) &\leq \int_{\mathcal{M}} \int_{\mathcal{M}} \frac{K_r(x,y)}{\theta(x)} \rho(x)(f(y))^2 d\text{Vol}(y) d\mu(x) \\
&\leq (1 + \alpha L_p r)(1 + \alpha L_\rho r)(1 + CmKr^2)\|f\|^2_{L^2(\mathcal{M},\rho\mu)},
\end{aligned}
$$

where the last inequality follows from the Lipschitz continuity of p and ρ together with the estimates from Lemma 7.

For the second assertion notice that as in the proof of [6, Lemma 5.4], we can conclude that for a.e. x

$$|\Lambda_r f(x) - f(x)|^2 \leq \frac{1}{\theta(x)} \int_{B_\mathcal{M}(x,r)} k_r(x,y) |f(y) - f(x)|^2 \, d\mathrm{Vol}(y).$$

Integrating this inequality with respect to $\rho\mu$ and using (3.6), we obtain that

$$\begin{aligned}
&\|\Lambda_r f - f\|^2_{L^2(\mathcal{M},\rho\mu)} \\
&\leq \frac{1 + CmKr^2}{\sigma_\eta r^m} \int_\mathcal{M} \int_\mathcal{M} \eta\left(\frac{d(x,y)}{r}\right) |f(x) - f(y)|^2 \, d\mathrm{Vol}(y)\rho(x)d\mu(x) \\
&\leq \frac{C}{\sigma_\eta r^m} \alpha^2 E_r(f).
\end{aligned}$$

$\square$

Lemma 9 ([cf. [6, Lemma 5.5]) *Suppose that h satisfies Assumptions* 3. *Then, there exists a universal constant* $C > 0$ *such that*

$$D(\Lambda_r f) \leq (1 + \alpha L_p r) \cdot (1 + C(1 + 1/\sigma_\eta)mKr^2)\frac{1}{\sigma_\eta r^{m+2}} E_r(f)$$

for every $f \in L^2(\mathcal{M})$ *and every* $0 < r < 2h$.

Proof We can write

$$\nabla(\Lambda_r f) = \frac{1}{\theta(x)} A_1(x) + A_2(x)$$

where

$$A_1(x) := \int_{B_\mathcal{M}(x,r)} \nabla k_r(\cdot, y)(x)(f(y) - f(x)) d\mathrm{Vol}(y)$$

and

$$A_2(x) := \nabla(\theta^{-1})(x) \int_{B_\mathcal{M}(x,r)} k_r(x,y)(f(y) - f(x)) d\mathrm{Vol}(y).$$

Regarding A_1, we have $|A_1(x)| = \langle A_1(x), w \rangle$ for some unit vector $w \in T_x\mathcal{M}$. Therefore, using (3.7),

$$\begin{aligned}
|A_1(x)| &= \langle A_1(x), w \rangle \\
&= \frac{1}{\sigma_\eta r^{m+2}} \int_{B_\mathcal{M}(x,r)} \eta\left(\frac{d(x,y)}{r}\right) (f(y) - f(x)) \langle \exp_x^{-1}(y), w \rangle d\mathrm{Vol}(y) \\
&= \frac{1}{\sigma_\eta r^{m+2}} \int_{B(r)} \eta\left(\frac{|v|}{r}\right) \varphi(v) \langle v, w \rangle J_x(v) dv.
\end{aligned}$$

where $\varphi(v) := f(\exp_x(v)) - f(x)$. By the Cauchy–Schwartz inequality,

$$|A_1(x)|^2 \leq \frac{1}{\sigma_\eta^2 r^{2(m+2)}} \int_{B(r)} |\varphi(v)|^2 J_x(v)^2 \eta\left(\frac{|v|}{r}\right) \mathrm{d}v \int_{B(r)} \langle v, w\rangle^2 \eta\left(\frac{|v|}{r}\right) \mathrm{d}v$$

$$= \frac{1}{\sigma_\eta r^{m+2}} \int_{B(r)} |\varphi(v)|^2 J_x(v)^2 \eta\left(\frac{|v|}{r}\right) \mathrm{d}v$$

where, in the last step, we used radial symmetry to conclude that

$$\int_{B(r)} \langle v, w\rangle^2 \eta\left(\frac{|v|}{r}\right) \mathrm{d}v = r^{m+2} \int_{B(1)} u_1^2 \eta(|u|)\mathrm{d}u = r^{m+2}\sigma_\eta.$$

Now, we obtain from (1.34) that

$$|A_1(x)|^2 \leq \frac{1 + CmKr^2}{\sigma_\eta r^{m+2}} \int_{B(r)} |\varphi(v)|^2 \, \eta\left(\frac{|v|}{r}\right) J_x(v)\mathrm{d}v$$

$$= \frac{1 + CmKr^2}{\sigma_\eta r^{m+2}} \int_{\mathcal{M}} \eta\left(\frac{d(x,y)}{r}\right) (f(y) - f(x))^2 \mathrm{d}\mathrm{Vol}(y)$$

Integrating this inequality with respect to the density p^2 and using the Lipschitz continuity of p, we obtain

$$\|A_1\|^2_{L^2(\mathcal{M}, p^2\mathrm{Vol})}$$

$$\leq \frac{1 + CmKr^2}{\sigma_\eta r^{m+2}} \int_{\mathcal{M}} \int_{B_{\mathcal{M}}(x,r)} \eta\left(\frac{d(x,y)}{r}\right) |f(y) - f(x)|^2 \, \mathrm{d}\mathrm{Vol}(y) p^2(x)\mathrm{d}\mathrm{Vol}(x)$$

$$\leq \frac{(1 + \alpha L_p r)(1 + CmKr^2)}{\sigma_\eta r^{m+2}} \int_{\mathcal{M}} \int_{B_{\mathcal{M}}(x,r)} \eta\left(\frac{d(x,y)}{r}\right) |f(y) - f(x)|^2 \, \mathrm{d}\mu(y)\mathrm{d}\mu(x)$$

$$\leq \frac{(1 + \alpha L_p r)(1 + CmKr^2)}{\sigma_\eta r^{m+2}} E_r(f).$$

Regarding A_2, first note that $\left|\nabla(\theta^{-1})\right| \leq CmKr/\sigma_\eta$ and $\theta \leq C$ by Lemma 7. Therefore, by the Cauchy–Schwartz inequality and (3.6), we obtain

$$|A_2(x)|^2 \leq \left|\nabla(\theta^{-1})\right|^2 \int_{\mathcal{M}} k_r(x,y)\mathrm{d}y \int_{\mathcal{M}} |f(y) - f(x)|^2 k_r(x,y)\mathrm{d}\mathrm{Vol}(y)$$

$$= \left|\nabla(\theta^{-1})\right|^2 \theta(x) \int_{\mathcal{M}} |f(y) - f(x)|^2 k_r(x,y)\mathrm{d}\mathrm{Vol}(y)$$

$$\leq \frac{Cm^2K^2r^2}{\sigma_\eta^3 r^m} \int_{\mathcal{M}} \eta\left(\frac{d(x,y)}{r}\right) |f(y) - f(x)|^2 \, \mathrm{d}\mathrm{Vol}(y)$$

Integrating this inequality with respect to the density p^2 while using the Lipschitz continuity of p shows that

$$\|A_2\|_{L^2(\mathcal{M}, p^2 \text{Vol})} \leq \frac{C(1 + \alpha L_p r) m K r^2}{\sigma_\eta} \sqrt{\frac{1}{\sigma_\eta r^{m+2}} E_r(f)}$$

for some universal constant C. By combining these estimates and the lower bound for θ from Lemma 7, we obtain that

$$D(\Lambda_r f)^{\frac{1}{2}} \leq (1 + \alpha L_p r) \cdot (1 + C(1 + 1/\sigma_\eta) m K r^2) \sqrt{\frac{1}{\sigma_\eta r^{m+2}} E_r(f)}.$$

Hence, the claim follows. □

4 Convergence of Eigenvalues

In order to prove Theorem 4, we estimate the discrete Dirichlet form (1.7) in terms of the continuous one (1.8) while we interpolate and discretize between the graph and the manifold in an almost isometric manner using the mappings P, P^* from (1.22), (1.23) and Λ_r from (3.5). We start this section with some preliminary lemmas.

Lemma 10 *Let us assume that the support of η is contained in $[0, 1]$ and that η is Lipschitz in $[0, 1]$. Then, for all $r, s > 0$ and $t \geq 0$ we have*

(i) $\eta\left(\frac{t}{r+s}\right) \leq \eta\left(\frac{(t-s)_+}{r}\right) \leq \eta\left(\frac{t}{r+s}\right) + L_\eta \frac{s}{r} \mathbb{1}_{\{t \leq r+s\}}$

(ii) $\eta\left(\frac{t+s}{r}\right) \geq \eta\left(\frac{t}{r-s}\right) - L_\eta \frac{s}{r} \mathbb{1}_{\{t \leq r-s\}}$ *provided that $s < r$.*

where $L_\eta > 0$ denotes the Lipschitz constant of η restricted to $[0, 1]$.

Proof Regarding assertion (i), first note that every term vanishes for $t > r + s$. In order to prove the first inequality in the remaining case, we need to verify that $(t - s)/r \leq t/(r + s)$ provided that $t \leq r + s$. This follows from

$$\frac{t}{r+s} - \frac{t-s}{r} = \frac{rt - (r+s)(t-s)}{r(r+s)} = \frac{s}{r} \frac{r+s-t}{r+s} = \frac{s}{r}\left(1 - \frac{t}{r+s}\right) > 0.$$

Combining this estimate with the Lipschitz continuity of η shows that

$$0 \leq \eta\left(\frac{(t-s)_+}{r}\right) - \eta\left(\frac{t}{r+s}\right) \leq L_\eta\left(\frac{t}{r+s} - \frac{(t-s)_+}{r}\right)$$

$$\leq L_\eta\left(\frac{t}{r+s} - \frac{t-s}{r}\right) = L_\eta \frac{s}{r}\left(1 - \frac{t}{r+s}\right) \leq L_\eta \frac{s}{r}$$

which implies the second inequality of assertion (i). The proof of assertion (ii) is completely analogous. □

The next results relate the operators P and P^* defined in (1.22) and (1.23). In particular, we show that P and P^* are almost adjoint to each other and that P^* is almost an isometry. In case $\mathbf{m}\mu_n = \mu_n$ and $\rho\mu = \mu$ (i.e., in case $m = (1, \ldots, 1)$ and $\rho \equiv 1$), then P and P^* are truly adjoint to each other and P^* is truly an isometry.

Lemma 11 *For all $u \in L^2(X)$ and $f \in L^2(\mathcal{M})$*

$$\left| \langle P^*u, f \rangle_{L^2(\mathcal{M}, \rho\mu)} - \langle u, Pf \rangle_{L^2(X, \mathbf{m}\mu_n)} \right| \leq \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) \langle P^* |u|, |f| \rangle_{L^2(\mathcal{M}, \rho\mu)}$$

and

$$\left| \left\| P^*u \right\|^2_{L^2(\mathcal{M}, \rho\mu)} - \|u\|^2_{L^2(X, \mathbf{m}\mu_n)} \right| \leq \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) \left\| P^*u \right\|^2_{L^2(\mathcal{M}, \rho\mu)}.$$

Moreover, if we assume that $\alpha\|\mathbf{m} - \rho\|_\infty \leq \frac{1}{2}$ then $\forall u \in L^2(X)$,

$$\left\| P^*u \right\|^2_{L^2(\mathcal{M}, \rho\mu)} \leq 2(1 + \alpha L_\rho \varepsilon) \|u\|^2_{L^2(X, \mathbf{m}\mu_n)}$$

for some universal constant $C > 0$.

Proof We infer from (1.12) that

$$\begin{aligned}
&\left| \langle u, Pf \rangle_{L^2(X, \mathbf{m}\mu_n)} - \langle P^*u, f \rangle_{L^2(\mathcal{M}, \rho\mu)} \right| \\
&= \left| \sum_{i=1}^n \frac{m_i}{n} u(x_i) \cdot n \int_{U_i} f d\mu - \int_{\mathcal{M}} \sum_{i=1}^n u(x_i) \mathbb{1}_{U_i} f \rho d\mu \right| \\
&\leq \int_{\mathcal{M}} \sum_{i=1}^n |u(x_i)| \mathbb{1}_{U_i} |f(x)| \cdot |m_i - \rho(x_i) + \rho(x_i) - \rho(x)| d\mu \\
&\leq \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) \langle P^* |u|, |f| \rangle_{L^2(\mathcal{M}, \rho\mu)}.
\end{aligned}$$

and

$$\begin{aligned}
&\left| \left\| P^*u \right\|^2_{L^2(\mathcal{M}, \rho\mu)} - \|u\|^2_{L^2(X, \mathbf{m}\mu_n)} \right| \\
&= \left| \sum_{i=1}^n \left(\int_{U_i} u^2(x_i) \rho d\mu - \int_{U_i} m_i u^2(x_i) d\mu \right) \right| \\
&\leq \sum_{i=1}^n \int_{U_i} u^2(x_i) |\rho(x) - \rho(x_i) + \rho(x_i) - m_i| d\mu \\
&\leq \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) \left\| P^*u \right\|^2_{L^2(\mathcal{M}, \rho\mu)}.
\end{aligned}$$

To prove the last part of the lemma, we notice that

$$\left\|P^*u\right\|^2_{L^2(\mathcal{M},\rho\mu)} = \sum_{i=1}^{n} u(x_i)^2 \int_{U_i} \rho(y)d\mu(y) \leq \frac{2(1+\alpha L_\rho\varepsilon)}{n} \sum_{i=1}^{n} u(x_i)^2 m_i$$

$$= 2(1+\alpha L_\rho\varepsilon)\|u\|^2_{L^2(X,\mathbf{m}\mu_n)}.$$

$\square$

The next lemma is a straightforward generalization of [6, Lemma 4.2].

Lemma 12 (cf. [6, Lemma 4.2]) *For every $f \in L^2(\mathcal{M})$, we have*

$$\left\|P^*Pf\right\|^2_{L^2(\mathcal{M},\rho\mu)} \leq (1+2\alpha L_\rho\varepsilon)\|f\|^2_{L^2(\mathcal{M},\rho\mu)}$$

In addition, there exists a universal constant $C > 0$ such that

$$\left\|f - P^*Pf\right\|_{L^2(\mathcal{M},\mu)} \leq \frac{C(1+m\alpha L_p\varepsilon)m2^{m/2}\sigma_\eta^{1/2}}{\sqrt{\eta(1/2)\omega_m}}\varepsilon D(f)^{\frac{1}{2}}$$

for all $f \in H^1(\mathcal{M})$.

Proof The first assertion follows from Jensen's inequality and the Lipschitz continuity of ρ:

$$\int_{\mathcal{M}} (P^*Pf(x))^2\rho(x)\mathrm{d}\mu(x) \leq \sum_{i=1}^{n} \int_{U_i} \int_{U_i} nf(y)^2\rho(x)\mathrm{d}\mu(y)\mathrm{d}\mu(x)$$

$$\leq (1+2\alpha L_\rho\varepsilon)\sum_{i=1}^{n} \int_{U_i} \int_{U_i} nf(y)^2\rho(y)\mathrm{d}\mu(y)\mathrm{d}\mu(x)$$

$$= (1+2\alpha L_\rho\varepsilon)\int_{\mathcal{M}} f(y)^2\rho(y)\mathrm{d}\mu(y).$$

For the second assertion, we can use Lemma 5, Lemma 6 and Assumptions 3 on h, to obtain

$$\left\|f - P^*Pf\right\|^2_{L^2(\mathcal{M},\mu)} \leq \frac{2(1+CmKr^2)}{\eta(1/2)\omega_m(r-\varepsilon)^m}E_{2r}(f)$$

$$\leq \frac{C(1+2\alpha L_pr)2^m\sigma_\eta}{\eta(1/2)\omega_m}\frac{r^m}{(r-\varepsilon)^m}r^2 D(f)$$

for any $r \in (\varepsilon, 2h)$. When choosing $r = (m+1)\varepsilon$, the quotient $\frac{r^m}{(r-\varepsilon)^m}$ is bounded by 3 and the assertion follows. $\square$

The next lemma is a generalization of [6, Lemma 4.3].

Lemma 13 (cf. [6, Lemma 4.3]) *The following assertions hold:*

(i) *For every $f \in H^1(\mathcal{M})$,*

$$\left| \|Pf\|^2_{L^2(X,\mathbf{m}\mu_n)} - \|f\|^2_{L^2(\mathcal{M},\rho\mu)} \right| \leq \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) \|f\|^2_{L^2(\mathcal{M},\rho\mu)}$$
$$+ \tilde{C}'\varepsilon \|f\|_{L^2(\mathcal{M},\rho\mu)} D(f)^{\frac{1}{2}},$$

where $\tilde{C}'$ has the form

$$\tilde{C}' = \frac{C\alpha(1 + \alpha L_\rho)(1 + m\alpha L_p)m2^{m/2}\sigma_\eta^{1/2}}{\sqrt{\eta(1/2)\omega_m}}, \tag{4.1}$$

for some universal constant $C > 0$.

(ii) *For every $f \in H^1(\mathcal{M})$,*

$$b(Pf) \leq (1 + C_1'h + C_2'\frac{\varepsilon}{h} + C_3'h^2)D(f),$$

where the constants C_1', C_2', C_3' can be written in terms of geometric quantities as

$$C_1' = C\alpha L_p, \quad C_2' = C\left(m + \frac{2^{m+1}L_\eta(1 + \alpha L_p)}{\eta(1/2)}\right), \quad C_3' = Cm\left(K + \frac{1}{R^2}\right),$$

where C is a universal constant.

Proof Since P^* is almost an isometry by Lemma 11, we have

$$\left| \|Pf\|^2_{L^2(X,\mathbf{m}\mu_n)} - \|f\|^2_{L^2(\mathcal{M},\rho\mu)} \right|$$
$$\leq \left| \|Pf\|^2_{L^2(X,\mathbf{m}\mu_n)} - \|P^*Pf\|^2_{L^2(\mathcal{M},\rho\mu)} \right|$$
$$+ \left| \|P^*Pf\|^2_{L^2(\mathcal{M},\rho\mu)} - \|f\|^2_{L^2(\mathcal{M},\rho\mu)} \right|$$
$$\leq \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) \|P^*Pf\|^2_{L^2(\mathcal{M},\rho\mu)}$$
$$+ (\|P^*Pf\|_{L^2(\mathcal{M},\rho\mu)} + \|f\|_{L^2(\mathcal{M},\rho\mu)}) \|P^*Pf - f\|_{L^2(\mathcal{M},\rho\mu)}$$
$$\leq \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho)(1 + 2\alpha L_\rho\varepsilon) \|f\|^2_{L^2(\mathcal{M},\rho d\mu)}$$
$$+ \frac{C\alpha(2 + \alpha L_\rho\varepsilon)(1 + m\alpha L_p\varepsilon)m2^{m/2}\sigma_\eta^{1/2}}{\sqrt{\eta(1/2)\omega_m}} \varepsilon \|f\|_{L^2(\mathcal{M},\rho d\mu)} D(f)^{\frac{1}{2}}$$

where the last inequality follows from Lemma 12 and from the boundedness of ρ. This proves the first assertion.

Regarding the second assertion, we follow the proof of [6, Lemma 4.3(ii)] and obtain that

$$\left| Pf(x_j) - Pf(x_i) \right|^2 \leq n^2 \int_{U_i} \int_{U_j} |f(y) - f(x)|^2 \, d\mu(y) d\mu(x).$$

Let $\hat{h} := (1 + \frac{27}{R^2} h^2) h$. Then, by Proposition 2, Lemma 10 and by the monotonicity of η we have

$$
\begin{aligned}
b(Pf) &\le \frac{1}{\sigma_\eta h^{m+2}} \sum_i \sum_j \int_{U_i} \int_{U_j} \eta\left(\frac{|x_i - x_j|}{h}\right) |f(y) - f(x)|^2 \, d\mu(y) d\mu(x) \\
&\le \frac{1}{\sigma_\eta h^{m+2}} \sum_i \sum_j \int_{U_i} \int_{U_j} \eta\left(\frac{d(x_i, x_j)}{\hat{h}}\right) |f(y) - f(x)|^2 \, d\mu(y) d\mu(x) \\
&\le \frac{1}{\sigma_\eta h^{m+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta\left(\frac{(d(x, y) - 2\varepsilon)_+}{\hat{h}}\right) |f(y) - f(x)|^2 \, d\mu(y) d\mu(x) \\
&\le \frac{1}{\sigma_\eta h^{m+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \left(\eta\left(\frac{d(x, y)}{\hat{h} + 2\varepsilon}\right) + 2L_\eta \frac{\varepsilon}{\hat{h}} \mathbb{1}_{B_{\mathcal{M}}(x, \hat{h} + 2\varepsilon)}(y)\right) \\
&\qquad |f(y) - f(x)|^2 \, d\mu(y) d\mu(x) \\
&= \frac{1}{\sigma_\eta h^{m+2}} \left(E_{\hat{h}+2\varepsilon}(f) + \frac{2L_\eta}{\eta(1/2)} \frac{\varepsilon}{h} E_{2(\hat{h}+2\varepsilon)}(f)\right),
\end{aligned}
$$

where we refer to Remark 11 to justify the last step. Due to Assumptions 3, we obtain from Lemma 5 that

$$
\begin{aligned}
\frac{1}{\sigma_\eta h^{m+2}} E_{\hat{h}+2\varepsilon}(f) &\le (1 + C\alpha L_p h)(1 + CmKh^2) \left(1 + \frac{27h^2}{R^2} + 2\frac{\varepsilon}{h}\right)^{m+2} D(f) \\
&\le (1 + C\alpha L_p h)(1 + CmKh^2) \left(1 + Cm\frac{h^2}{R^2} + Cm\frac{\varepsilon}{h}\right) D(f),
\end{aligned}
$$

where the last inequality is obtained from the fact that

$$
(1 + s)^m \le 1 + Cs, \quad \forall 0 \le s \le \frac{3}{m},
$$

for some universal constant $C > 0$. Likewise, we obtain

$$
\begin{aligned}
\frac{1}{\sigma_\eta h^{m+2}} \frac{2L_\eta}{\eta(1/2)} \frac{\varepsilon}{h} E_{2(\hat{h}+2\varepsilon)}(f) &\le \frac{2^{m+1} L_\eta}{\eta(1/2)} (1 + C\alpha L_p h)(1 + CmKh^2) \\
&\qquad \cdot \left(1 + Cm\frac{h^2}{R^2} + Cm\frac{\varepsilon}{h}\right) \frac{\varepsilon}{h} D(f).
\end{aligned}
$$

The result follows directly from the previous estimates. $\qquad\square$

We can now establish an upper bound for $\lambda_k(\Gamma)$ in terms of $\lambda_k(\mathcal{M})$.

Proof (*of upper bound of Theorem* 4) Fix $k \in \mathbb{N}$. By the minmax principle (1.29), we have

$$
\lambda_k(\Gamma) \le \sup_{u \in L \setminus \{0\}} \frac{b(u)}{\|u\|^2_{L^2(X, \mathbf{m}\mu_n)}}
$$

for every k-dimensional subspace $L \subseteq L^2(X, \mathbf{m}\mu_n)$. Following the proof of [6, Prop 4.4], we denote by $W \subset H^1(\mathcal{M})$ the span of orthonormal (with respect to the $L^2(\mathcal{M}, \rho\mu)$ inner product) eigenfunctions of Δ corresponding to $\lambda_1(\mathcal{M}), \ldots, \lambda_k(\mathcal{M})$ and we set $L := P(W)$. For every $f \in W$, we have $D(f) \leq \lambda_k(\mathcal{M}) \|f\|^2_{L^2(\mathcal{M},\rho\mu)}$. It thus follows from part (i) of Lemma 13 that

$$\|Pf\|^2_{L^2(X,\mathbf{m}\mu_n)} \geq (1 - \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) - \tilde{C}'\sqrt{\lambda_k(\mathcal{M})}\varepsilon) \|f\|^2_{L^2(\mathcal{M},\rho\mu)}. \tag{4.2}$$

Hence, provided that

$$\alpha(\|\mathbf{m} + \rho\|_\infty + \varepsilon L_\rho) + \tilde{C}'\sqrt{\lambda_k(\mathcal{M})}\varepsilon \leq \frac{1}{2},$$

we can conclude that P is injective on W, and therefore, $\dim L = k$. Moreover, in that case by applying part (ii) of Lemma 13 to $u = Pf \in L$ we obtain that

$$\frac{b(u)}{\|u\|^2_{L^2(\mathcal{M},\mathbf{m}\mu_n)}} \leq \frac{\left(1 + C'_1 h + C'_2 \frac{\varepsilon}{h} + C'_3 h^2\right)}{1 - \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) - \tilde{C}'\sqrt{\lambda_k(\mathcal{M})}\varepsilon} \lambda_k(\mathcal{M})$$
$$\leq \left(1 + C'_1 h + C'_2 \frac{\varepsilon}{h} + C'_3 h^2 + \alpha C(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) + \tilde{C}'\sqrt{\lambda_k(\mathcal{M})}\varepsilon\right) \lambda_k(\mathcal{M}).$$

Since the previous inequality holds for every $u = Pf$ with $f \in W$, the desired estimate now follows. $\qquad\square$

Lemma 14 (cf. [6, Lemma 6.2]) *Suppose that h satisfies Assumptions 3. Then,*

(i) *For every $u \in L^2(X)$,*

$$\left| \|Iu\|^2_{L^2(\mathcal{M},\rho\mu)} - \|u\|^2_{L^2(X,\mathbf{m}\mu_n)} \right| \leq \tilde{C}'' h \|u\|_{L^2(X,\mathbf{m}\mu_n)} \cdot b(u)^{\frac{1}{2}}$$
$$+ 2\alpha(1 + \alpha L_\rho) \cdot (\|\mathbf{m} - \rho\|_\infty + L_\rho \varepsilon) \|u\|^2_{L^2(X,\mathbf{m}\mu_n)},$$

where the constant $\tilde{C}''$ can be written as

$$\tilde{C}'' = C\alpha(1 + \alpha L_p) \cdot (1 + \alpha L_\rho) \cdot \left(1 + c''\right), \quad c'' = \frac{L_\eta 4^m \omega_m^2 (1 + \alpha L_p)^2}{\eta(1/2)(m+2)}.$$

(ii) *For every $u \in L^2(X)$,*

$$D(Iu) \leq (1 + C''_1 h + C''_2 \frac{\varepsilon}{h} + C''_3 h^2) b(u),$$

where the constants C''_1, C''_2, C''_3 have the form

$$C''_1 = \alpha L_p, \quad C''_2 = C(m + C'_2), \quad C''_3 = C(1 + 1/\sigma_\eta)mK.$$

Proof First, by Lemma 11,

$$
\left| \|Iu\|_{L^2(\mathcal{M},\rho\mu)}^2 - \|u\|_{L^2(X,\mathbf{m}\mu_n)}^2 \right|
$$

$$
\leq \left| \|Iu\|_{L^2(\mathcal{M},\rho\mu)}^2 - \|P^*u\|_{L^2(\mathcal{M},\rho\mu)}^2 \right| + \left| \|P^*u\|_{L^2(\mathcal{M},\rho\mu)}^2 - \|u\|_{L^2(X,\mathbf{m}\mu_n)}^2 \right|
$$

$$
\leq (\|Iu\|_{L^2(\mathcal{M},\rho\mu)} + \|P^*u\|_{L^2(\mathcal{M},\rho\mu)}) \|Iu - P^*u\|_{L^2(\mathcal{M},\rho\mu)}
$$

$$
+ \alpha(\|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho) \|P^*u\|_{L^2(\mathcal{M},\rho\mu)}^2 .
$$

$$(4.3)$$

Since $(m + 2)\varepsilon < h$ (by Assumption 3), we conclude from Lemma 8 that

$$
\|Iu - P^*u\|_{L^2(\mathcal{M},\rho\mu)}^2 = \|\Lambda_{h-2\varepsilon} P^*u - P^*u\|^2 \leq \frac{C\alpha^2}{\sigma_\eta h^m} E_{h-2\varepsilon}(P^*u).
$$

for some universal constant $C > 0$.

Let us now estimate $E_{h-2\varepsilon}$ in terms of $b(u)$. First, consider the kernel $\tilde{\eta} = \mathbb{1}_{[0,1]}$.
We use $\tilde{b}$ and $\tilde{E}$ to denote the discrete Dirichlet form and the energy E when using
the kernel $\tilde{\eta}$ and we write b_h and $\tilde{b}_h$, respectively, to specify that the forms b and $\tilde{b}$ are
being constructed using the value h. We claim that

$$
\tilde{b}_h(u) \geq \frac{m + 2}{\omega_m h^{m+2}} \tilde{E}_{h-2\varepsilon}(P^*u).
$$

$$(4.4)$$

Indeed, let T denote the transportation map introduced in Sect. 1.3 satisfying $U_i = T^{-1}(x_i)$, then

$$
\tilde{b}_h(u) = \frac{1}{\sigma_{\tilde{\eta}} h^{m+2}} \frac{1}{n^2} \sum_i \sum_j \tilde{\eta}\left(\frac{|x_i - x_j|}{h}\right) |u(x_i) - u(x_j)|^2
$$

$$
= \frac{1}{\sigma_{\tilde{\eta}} h^{m+2}} \sum_{i,j} \int_{U_i} \int_{U_j} \tilde{\eta}\left(\frac{|T(x) - T(y)|}{h}\right) |(P^*u)(x) - (P^*u)(y)|^2 \, d\mu(y) d\mu(x)
$$

$$
\geq \frac{1}{\sigma_{\tilde{\eta}} h^{m+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \tilde{\eta}\left(\frac{d(T(x), T(y))}{h}\right) |(P^*u)(x) - (P^*u)(y)|^2 \, d\mu(y) d\mu(x)
$$

$$
\geq \frac{1}{\sigma_{\tilde{\eta}} h^{m+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \tilde{\eta}\left(\frac{d(x, y)}{h - 2\varepsilon}\right) |(P^*u)(x) - (P^*u)(y)|^2 \, d\mu(y) d\mu(x)
$$

$$
= \frac{1}{\sigma_{\tilde{\eta}} h^{m+2}} \tilde{E}_{h-2\varepsilon}(P^*u),
$$

where we note that the last inequality follows from the fact that $d(T(x), T(y)) > h$
implies that $d(x, y) > h - 2\varepsilon$; we have used Remark 1 to rewrite $\sigma_{\tilde{\eta}}$. We now consider
general η. Since $\eta(t) \geq \eta(1/2) > 0$ for all $t \in [0, 1/2]$, it follows that

$$
\tilde{b}_{h/2}(u) \leq \frac{\sigma_\eta \omega_m 2^{m+2}}{\eta(1/2)(m + 2)} b_h(u).
$$

$$(4.5)$$

On the other hand, by the monotonicity of η and Lemma 10 we obtain

$$
\begin{aligned}
b_h(u) &\geq \frac{1}{\sigma_\eta h^{m+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta\left(\frac{d(T(x), T(y))}{h}\right) \left|(P^*u)(x) - (P^*u)(y)\right|^2 d\mu(y)d\mu(x) \\
&\geq \frac{1}{\sigma_\eta h^{m+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta\left(\frac{d(x, y) + 2\varepsilon}{h}\right) \left|(P^*u)(x) - (P^*u)(y)\right|^2 d\mu(y)d\mu(x) \\
&\geq \frac{1}{\sigma_\eta h^{m+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \eta\left(\frac{d(x, y)}{h - 2\varepsilon}\right) \left|(P^*u)(x) - (P^*u)(y)\right|^2 d\mu(y)d\mu(x) \\
&\quad - \frac{L_\eta}{\sigma_\eta} \frac{\varepsilon}{h} \frac{1}{h^{m+2}} \int_{\mathcal{M}} \int_{\mathcal{M}} \mathbb{1}_{\{d(x,y)\leq h-2\varepsilon\}} \left|(P^*u)(x) - (P^*u)(y)\right|^2 d\mu(y)d\mu(x) \\
&= \frac{1}{\sigma_\eta h^{m+2}} E_{h-2\varepsilon}(P^*u) - \frac{L_\eta}{\sigma_\eta} \frac{\varepsilon}{h} \frac{1}{h^{m+2}} \tilde{E}_{h-2\varepsilon}(P^*u) \\
&\geq \frac{1}{\sigma_\eta h^{m+2}} E_{h-2\varepsilon}(P^*u) - \frac{CL_\eta 4^m (1+\alpha L_p)^2}{\sigma_\eta} \frac{\varepsilon}{h} \frac{1}{h^{m+2}} \tilde{E}_{\frac{h}{2}-2\varepsilon}(P^*u),
\end{aligned}
$$

where the last inequality follows after applying Lemma 4 twice. We conclude from (4.4) that

$$
b_h(u) \geq \frac{1}{\sigma_\eta h^{m+2}} E_{h-2\varepsilon}(P^*u) - \frac{CL_\eta 2^m \omega_m (1+\alpha L_p)^2}{(m+2)\sigma_\eta} \frac{\varepsilon}{h} \tilde{b}_{\frac{h}{2}}(u).
$$

Combining this inequality with (4.5), we deduce that

$$
\left(1 + \frac{CL_\eta 4^m \omega_m^2 (1+\alpha L_p)^2}{\eta(1/2)(m+2)^2} \frac{\varepsilon}{h}\right) b_h(u) \geq \frac{1}{\sigma_\eta h^{m+2}} E_{h-2\varepsilon}(P^*u)
$$

which can be rewritten as

$$
E_{h-2\varepsilon}(P^*u) \leq \left(1 + \frac{CL_\eta 4^m \omega_m^2 (1+\alpha L_p)^2}{\eta(1/2)(m+2)^2} \frac{\varepsilon}{h}\right) \sigma_\eta h^{m+2} b(u). \tag{4.6}
$$

Hence,

$$
\left\| Iu - P^*u \right\|^2 \leq \frac{C\alpha^2}{\sigma_\eta h^m} E_{h-2\varepsilon}(P^*u) \leq C\alpha^2 \left(1 + \frac{L_\eta 4^m \omega_m^2 (1+\alpha L_p)^2}{\eta(1/2)(m+2)^2} \frac{\varepsilon}{h}\right) h^2 b(u). \tag{4.7}
$$

Finally, from Lemma 11 it follows that

$$
\left\| P^*u \right\|^2_{L^2(\mathcal{M}, \rho\mu)} \leq 2(1 + \alpha L_\rho \varepsilon) \|u\|^2_{L^2(X, \mathbf{m}\mu)}
$$

and from Lemma 8

$$\|Iu\|_{L^2(\mathcal{M},\rho\mu)} = \|\Lambda_{h-2\varepsilon} P^* u\|_{L^2(\mathcal{M},\rho\mu)}$$
$$\leq C(1+\alpha L_p h)^{1/2} \cdot (1+\alpha L_\rho h)^{1/2} \|P^* u\|_{L^2(\mathcal{M},\rho\mu)}$$
$$\leq C(1+\alpha L_p h) \cdot (1+\alpha L_\rho h) \|u\|_{L^2(X,\mathbf{m}\mu_n)}$$

Assertion (i) follows by inserting all these estimates back in (4.3).

Regarding assertion (ii), we conclude from Lemma 9 that

$$D(Iu) \leq (1+\alpha L_p h) \cdot \left(1 + C\left(1+\frac{1}{\sigma_\eta}\right) mKh^2\right) \frac{1}{\sigma_\eta(h-2\varepsilon)^{m+2}} E_{h-2\varepsilon}(P^* u)$$
$$\leq (1+\alpha L_p h) \cdot \left(1 + C\left(1+\frac{1}{\sigma_\eta}\right) mKh^2\right)\left(1 + Cm\frac{\varepsilon}{h}\right) \frac{1}{\sigma_\eta h^{m+2}} E_{h-2\varepsilon}(P^* u)$$
$$\leq \left(1 + \alpha L_p h + C\left(1+\frac{1}{\sigma_\eta}\right) mKh^2 + Cm\frac{\varepsilon}{h}\right) \frac{1}{\sigma_\eta h^{m+2}} E_{h-2\varepsilon}(P^* u).$$

Combining with (4.6), we obtain the desired estimate. $\qquad\square$

We can now establish a lower bound for $\lambda_k(\Gamma)$ in terms of $\lambda_k(\mathcal{M})$.

Proof *(of lower bound of Theorem 4)* Let $k \in \mathbb{N}$. It follows from (1.30) that for very k-dimensional subspace $L \subset H^1(\mathcal{M})$, we have

$$\lambda_k(\mathcal{M}) \leq \sup_{f \in L \setminus \{0\}} \frac{D(f)}{\|f\|_{L^2(\mathcal{M},\rho\mu)}^2}.$$

As in the proof of [6, Prop 6.3] we denote by $W \subseteq L^2(X)$ the span of orthonormal eigenvectors of Δ_Γ corresponding to $\lambda_1(\Gamma), \ldots, \lambda_k(\Gamma)$ and we set $L := I(W)$. Then, $b(u) \leq \lambda_k(\Gamma) \|u\|_{L^2(X,\mathbf{m}\mu_n)}^2$ for all $u \in W$. Using this, we conclude from Lemma 14

$$\|Iu\|_{L^2(\mathcal{M},\rho\mu)}^2 \geq \left(1 - 2\alpha(1+\alpha L_\rho)(\|\mathbf{m} - \rho\|_\infty + L_\rho\varepsilon) - \tilde{C}''\sqrt{\lambda_k(\Gamma)}h\right) \tag{4.8}$$
$$\|u\|_{L^2(X,\mathbf{m}\mu_n)}^2$$

for all $u \in W$. It follows that if

$$2\alpha(1+\alpha L_\rho)(\|\mathbf{m} - \rho\|_\infty + L_\rho\varepsilon) + \tilde{C}''\sqrt{\lambda_k(\Gamma)}h \leq \frac{1}{2}$$

then the operator I is injective on W, and thus, $\dim L = k$; notice that this inequality is satisfied under condition (1.17) thanks to the upper bound for $\lambda_k(\Gamma)$ in terms of $\lambda_k(\mathcal{M})$. It follows from part (ii) of Lemma 14 that for any $f = Iu$ with $u \in W$,

$$\frac{D(f)}{\|f\|_{L^2(\mathcal{M},\rho\mu)}^2} \leq \frac{1 + C_1''h + C_2''\frac{\varepsilon}{h} + C_3''h^2}{1 - 2\alpha(1+\alpha L_\rho)(\|\mathbf{m} - \rho\|_\infty + L_\rho\varepsilon) - \tilde{C}''\sqrt{\lambda_k(\Gamma)}h} \lambda_k(\Gamma)$$

The result now follows from the fact that the above inequality holds for arbitrary $u \in W$ and the fact that $\lambda_k(\Gamma)$ can be bounded from above by a constant multiple of $\lambda_k(\mathcal{M})$. $\qquad\square$

5 Approximation of Eigenfunctions

In this section, we are concerned with the convergence of eigenvectors of Δ_Γ. We start by showing that the discretization and interpolation operators P and I are almost inverse of one another.

Lemma 15 (cf. [6, Lemma 6.4]) *Under Assumptions 3, there exists a constant C''' only depending on $m, \alpha, \eta, L_p, L_\rho$ such that*

(i) $\|IPf - f\|_{L^2(\mathcal{M}, \rho\mu)} \le C''' h D(f)^{\frac{1}{2}}$ *for all $f \in H^1(\mathcal{M})$.*

Moreover, if

$$\alpha \|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho \le \frac{1}{2}, \ then,$$

(ii) $\|PIu - u\|_{L^2(X, \mathbf{m}\mu_n)} \le C''' h b(u)^{\frac{1}{2}}$ *for all $u \in L^2(X)$.*

Proof By definition of I, we have

$$\|IPf - f\| \le \left\| \Lambda_{h-2\varepsilon}(P^* Pf - f) \right\| + \left\| \Lambda_{h-2\varepsilon} f - f \right\|.$$

From Lemmas 8 and 12, and from Assumptions 3, we know that for a constant $C'' > 0$, depending on η, m, L_p, L_ρ and α,

$$\left\| \Lambda_{h-2\varepsilon}(P^* Pf - f) \right\|_{L^2(\mathcal{M}, \rho\mu)} \le C'' \left\| P^* Pf - f \right\|_{L^2(\mathcal{M}, \mu)} \le C'' \varepsilon D(f)^{\frac{1}{2}}.$$

Likewise, from Lemmas 8 and 5,

$$\|\Lambda_{h-2\varepsilon} f - f\|^2_{L^2(\mathcal{M}, \rho\mu)} \le \frac{C''}{(h-2\varepsilon)^m} E_{h-2\varepsilon}(f) \le C'' h^2 D(f),$$

and from this we deduce assertion (i).

Regarding assertion (ii), if we assume that $\alpha \|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho \le \frac{1}{2}$, we obtain from Lemma 11 that

$$\begin{aligned}\|PIu - u\|_{L^2(\mathcal{M}, \rho\mu)} &\le 4 \left\| P^*(PIu - u) \right\|_{L^2(\mathcal{M}, \rho\mu)} \\ &\le 4 \left\| P^* PIu - Iu \right\|_{L^2(\mathcal{M}, \rho\mu)} + 4 \left\| Iu - P^* u \right\|_{L^2(\mathcal{M}, \rho\mu)}.\end{aligned}$$

From Lemmas 12 and 14, and from Assumptions 3, we obtain that

$$\left\| P^* PIu - Iu \right\|_{L^2(\mathcal{M}, \rho\mu)} \le C' \varepsilon D(Iu)^{\frac{1}{2}} \le C' \varepsilon b(u)^{\frac{1}{2}}$$

for a constant C' depending on η, m, L_p, L_ρ and α. Moreover, by (4.7) we know there exists $C''' > 0$ (depending on η, m, L_p, L_ρ and α) such that

$$\left\| Iu - P^*u \right\|_{L^2(\mathcal{M},\rho\mu)} \le C'''hb(u)^{\frac{1}{2}}.$$

This implies assertion (ii). □

Now, we adopt some additional notation from [6, Section 7]. For a value $\lambda \in \mathbb{R}$, we denote by $H_\lambda(\mathcal{M})$ the linear span in $H^1(\mathcal{M})$ of all eigenfunctions of Δ corresponding to eigenvalues in the interval $(-\infty, \lambda)$. Similarly, we define $H_\lambda(X)$ as the linear span of eigenvectors of Δ_Γ corresponding to eigenvalues in $(-\infty, \lambda)$. We write $\mathbb{P}_\lambda$ for both, the orthogonal projection onto $H_\lambda(\mathcal{M})$ and $H_\lambda(X)$.

Lemma 16 (cf. [6, Lemma 7.1]) *Suppose that h satisfies Assumptions 3 and that*

$$\alpha \|\mathbf{m} - \rho\|_\infty + \varepsilon L_\rho \le \frac{1}{2}.$$

Then, for every $\lambda > 0$ we have

(i) $b(Pf)^{\frac{1}{2}} \ge \left(1 - (\sqrt{\lambda}C''' + C_1'')h - C_2''\frac{\varepsilon}{h} - C_3''h^2\right)D(f)^{\frac{1}{2}}.$

(ii) $D(Iu)^{\frac{1}{2}} \ge \left(1 - (\sqrt{\lambda}C''' + C_1')h - C_2'\frac{\varepsilon}{h} - C_3'h^2\right)b(u)^{\frac{1}{2}}$

for all $f \in H_\lambda(\mathcal{M})$ and $u \in H_\lambda(X)$. The constants C_1'', C_2'', C_3'' are as in Lemma 14, C_1', C_2', C_3' are as in Lemma 13, and the constant C''' is as in Lemma 15.

Proof Fix some $\lambda > 0$. First note that the projection $\mathbb{P}_\lambda$ does not increase the Dirichlet energy (neither the graph one nor the continuum one), and hence we conclude that

$$D(IPf)^{\frac{1}{2}} \ge D(\mathbb{P}_\lambda IPf)^{1/2} \ge D(f)^{1/2} - D(\mathbb{P}_\lambda IPf - f)^{1/2}.$$

From Lemma 15 (i) it follows that,

$$D(\mathbb{P}_\lambda IPf - f)^{\frac{1}{2}} = D(\mathbb{P}_\lambda(IPf - f))^{\frac{1}{2}} \le \sqrt{\lambda}\|IPf - f\|_{L^2(\mathcal{M},\rho\mu)} \le C'''\sqrt{\lambda}hD(f)^{\frac{1}{2}}$$

for all $f \in H_\lambda(\mathcal{M})$. Hence,

$$D(IPf)^{\frac{1}{2}} \ge (1 - C'''h\sqrt{\lambda})D(f)^{\frac{1}{2}}.$$

Moreover, we know from Lemma 14 (ii) that

$$D(IPf)^{\frac{1}{2}} \le \left(1 + C_1''h + C_2''\frac{\varepsilon}{h} + C_3''h^2\right)b(Pf)^{\frac{1}{2}}$$

and thus

$$b(Pf)^{\frac{1}{2}} \geq \frac{1 - C'''h\sqrt{\lambda}}{1 + C_1''h + C_2''\frac{\varepsilon}{h} + C_3''h^2} D(f)^{\frac{1}{2}}$$
$$\geq \left(1 - (\sqrt{\lambda}C''' + C_1'')h - C_2''\frac{\varepsilon}{h} - C_3''h^2\right) D(f)^{\frac{1}{2}}$$

for all $f \in H_\lambda(\mathcal{M})$ as claimed in (i). Regarding assertion (ii), we proceed similarly. First, we obtain that

$$b(PIu)^{\frac{1}{2}} \geq b(u)^{1/2} - b(\mathbb{P}_\lambda PIu - u)^{1/2}$$

for $u \in H_\lambda(X)$. Since

$$b(\mathbb{P}_\lambda(PIu - u))^{\frac{1}{2}} \leq \sqrt{\lambda}\|PIu - u\|_{L^2(X, \mathbf{m}\mu_n)} \leq \sqrt{\lambda}C'''hb(u)^{\frac{1}{2}}$$

by part (ii) of Lemma 15, we have

$$b(PIu)^{\frac{1}{2}} \geq (1 - C'''\sqrt{\lambda}h)b(u)^{\frac{1}{2}}.$$

Moreover, we know from part (ii) of Lemma 13 that

$$b(PIu)^{\frac{1}{2}} \leq \left(1 + C_1'h + C_2'\frac{\varepsilon}{h} + C_3'h^2\right) D(Iu)^{\frac{1}{2}}.$$

Therefore,

$$D(Iu)^{\frac{1}{2}} \geq \frac{1 - C'''\sqrt{\lambda}h}{1 + C_1'h + C_2'\frac{\varepsilon}{h} + C_3'h^2} b(u)^{\frac{1}{2}}$$
$$\geq \left(1 - (C'''\sqrt{\lambda} + C_1')h - C_2'\frac{\varepsilon}{h} - C_3'h^2\right) b(u)^{\frac{1}{2}},$$

which proves assertion (ii). □

Proof *(Theorem 5)* This theorem can now be proven word-for-word as [6, Theorem 4] together with the required Lemmas [6, Lemma 7.2, 7.3, 7.4] by replacing every application of Lemmas 4.3, 6.2, 7.1 and Theorem 1 therein with the previously proven Lemmas 13, 14, 16 and Theorem 4, respectively. □

We now focus on establishing Theorem 6. To simplify our computations, we set

$$\theta := \left(\frac{\varepsilon}{h} + (1 + \sqrt{\lambda_k(\mathcal{M})})h + \left(K + \frac{1}{R^2}\right)h^2 + \|\mathbf{m} - \rho\|_\infty\right).$$

In the setting of Theorem 5, we have

$$\left\| P^*u - f \right\| \leq \left\| P^*u - Iu \right\| + \left\| Iu - f \right\| \leq C'hb(u)^{1/2} + \frac{\tilde{C}}{g_{k,\rho\mu}}\theta$$

$$= C'h\sqrt{\lambda_k(\Gamma)} + \frac{\tilde{C}}{g_{k,\rho\mu}}\theta$$

where the second inequality follows from (4.7). From Theorem 4, for h small enough we have

$$\left\| P^*u - f \right\| \leq C'h\sqrt{\lambda_k(\mathcal{M})} + \frac{\tilde{C}}{g_{k,\rho\mu}}\theta, \tag{5.1}$$

Therefore, every extension of u that approximates P^*u in $L^2(\mathcal{M}, \rho\mu)$ (or equivalently in $L^2(\mathcal{M}, \mu)$) is also an approximation of the eigenfunction f.

We recall the definition of sets $U_i \subset \mathcal{M}$ in (1.21), Euclidean Voronoi cells V_i in (1.25), and of the extended vector $\bar{u}$ from (1.26). Concerning the measure of such a Voronoi cell, we obtain the following bound.

Lemma 17 *For every $\beta > 1$, there exists a constant $C > 0$ depending on m and on ℓ from (1.14) such that*

$$\frac{\mu(V_i)}{\mu(U_i)} \leq C \cdot \log^{mp_m} n =: C(n)$$

for all $i = 1, \ldots, n$ and all $n \in \mathbb{N}$ with probability at least $1 - C_{K,Vol(\mathcal{M}),m,i_0}n^{-\beta}$.

Proof We first show that $V_i \subseteq \{x \in \mathcal{M} : |x - x_i| \leq \varepsilon\}$. To this end, suppose $x \in \mathcal{M}$ such that $|x - x_i| > \varepsilon$. Then also $d(x, x_i) > \varepsilon$. Since the balls $B_{\mathcal{M}}(x_j, \varepsilon)$ cover $\mathcal{M}$ by the choice of ε, there exists x_j such that $d(x, x_j) < \varepsilon$. Therefore, $|x - x_j| < \varepsilon < |x - x_i|$ and thus $x \notin V_i$. This proves the claim.

Now, we assume that the assertion of Theorem 2 holds. For $\varepsilon \leq \frac{R}{2}$, it follows from Proposition 2 that V_i is contained in the ball $B_{\mathcal{M}}(x_i, 3\varepsilon)$. Thus, we obtain from the bounds on the distortion of metric by the exponential map (1.35) that

$$\frac{\mu(V_i)}{\mu(U_i)} \leq \frac{\mu(B_{\mathcal{M}}(x_i, 3\varepsilon))}{\mu(U_i)} \leq \frac{\alpha\omega_m(3\varepsilon)^m C}{1/n} = C\alpha\omega_m 3^m \ell^m \log^{p_m \cdot m}(n)$$

where ℓ defined in (1.14), and $C > 0$ is a universal constant. $\qquad\square$

Proof *(Theorem 6)* Let $u \in L^2(X)$ be a normalized eigenvector of Δ_Γ corresponding to $\lambda_k(\Gamma)$ and let f a normalized eigenfunction of Δ corresponding to $\lambda_k(\mathcal{M})$ as in Theorem 5 (or as in (5.1)). Let

$$V := \|\bar{u} - f\|^2_{L^2(\mathcal{M},\mu)} = \sum_{i=1}^n \int_{V_i} |u(x_i) - f(y)|^2 \, d\mu(y)$$

and

$$U := \sum_{i=1}^{n} \frac{\mu(V_i)}{\mu(U_i)} \int_{U_i} |u(x_i) - f(x)|^2 \, d\mu(x).$$

Then, by Lemma 17 and (5.1),

$$\sqrt{U} \leq \sqrt{C(n)} \cdot \|P^*u - f\|_{L^2(\mathcal{M},\mu)} \leq \sqrt{C(n)} \left(C'h\sqrt{\lambda_k(\mathcal{M})} + \frac{\tilde{C}}{g_{k,\rho\mu}} \theta \right).$$

On the other hand,

$$|V - U| \leq \sum_{i=1}^{n} \mu(V_i) \left| \int_{V_i} |u(x_i) - f(y)|^2 \frac{d\mu(y)}{\mu(V_i)} - \int_{U_i} |u(x_i) - f(x)|^2 \frac{d\mu(x)}{\mu(U_i)} \right|$$

$$= \sum_{i=1}^{n} \mu(V_i) \left| \int_{V_i} \int_{U_i} \left(|u(x_i) - f(y)|^2 - |u(x_i) - f(x)|^2 \right) \frac{d\mu(y)}{\mu(V_i)} \frac{d\mu(x)}{\mu(U_i)} \right|$$

$$= \sum_{i=1}^{n} \mu(V_i) \left| \int_{V_i} \int_{U_i} \left(2(u(x_i) - f(x))(f(x) - f(y)) + (f(x) - f(y))^2 \right) \right.$$

$$\left. \frac{d\mu(y)}{\mu(V_i)} \frac{d\mu(x)}{\mu(U_i)} \right|$$

$$\leq 8\|\nabla f\|_\infty \varepsilon \left(\sum_{i=1}^{n} \mu(V_i) \int_{V_i} \int_{U_i} |u(x_i) - f(x)| \frac{d\mu(y)}{\mu(V_i)} \frac{d\mu(x)}{\mu(U_i)} \right) + 16\|\nabla f\|_\infty^2 \varepsilon^2$$

$$= 8\|\nabla f\|_\infty \varepsilon \left(\sum_{i=1}^{n} \mu(V_i) \int_{U_i} |u(x_i) - f(x)| \frac{d\mu(x)}{\mu(U_i)} \right) + 16\|\nabla f\|_\infty^2 \varepsilon^2$$

$$\leq 8\|\nabla f\|_\infty \varepsilon \sqrt{U} + 16\|\nabla f\|_\infty^2 \varepsilon^2$$

where in the second equality we have used the fact that for all $y \in V_i$ and all $x \in U_i$, $d(x, y) \leq d(x, x_i) + d(x_i, y) \leq 3\varepsilon + \varepsilon$; the last inequality follows from Jensen's inequality.

Thus,

$$V \leq |V - U| + U \leq 16(\varepsilon\|\nabla f\|_\infty + \sqrt{U})^2,$$

and from this it follows that

$$\|\bar{u} - f\|_{L^2(\mathcal{M},\mu)} = \sqrt{V} \leq 4\varepsilon\|\nabla f\|_\infty + 4\sqrt{U}.$$

Using [22], we know that

$$\|\nabla f\|_\infty \leq C_{\mathcal{M}} \lambda_k(\mathcal{M})^{\frac{m+1}{4}} \|f\|_{L^2(\mathcal{M},\mu)} = C_{\mathcal{M}} \lambda_k(\mathcal{M})^{\frac{m+1}{4}},$$

for a constant $C_{\mathcal{M}} > 0$ that depends on the manifold $\mathcal{M}$. Putting everything together, we deduce that

$$\|\bar{u} - f\|_{L^2(\mathcal{M},\mu)} \leq C_{\mathcal{M}} \lambda_k(\mathcal{M})^{\frac{m+1}{4}} \varepsilon + \tilde{C}\sqrt{C(n)}\left(\sqrt{\lambda_k(\mathcal{M})}h + \frac{\theta}{g_{k,\rho\mu}}\right),$$

which is the desired estimate. $\qquad\square$

A Proofs of Propositions in Sect. 1.6

Proof *(of Proposition 1)* The first claim follows immediately from (1.33). To deduce the second part, let $q_1, q_2 \in B_{\mathcal{M}}(p, \frac{r}{2})$. Consider a smooth curve $\tilde{\gamma} \colon [0, 1] \to \mathcal{M}$ connecting q_1 and q_2, i.e., $\tilde{\gamma}(0) = q_1$ and $\tilde{\gamma}(i) = q_2$. We observe that if $\tilde{\gamma}$ is not contained in $B_{\mathcal{M}}(p, r)$, then

$$d(q_1, q_2) \leq d(q_1, p) + d(q_2, p) < r \leq \text{Length}(\tilde{\gamma}).$$

In fact, to deduce that $r \leq \text{Length}(\tilde{\gamma})$ let $s \in (0, 1)$ be such that $\tilde{\gamma}(s) \notin B_{\mathcal{M}}(p, r)$. It is straightforward to see that the length of the restriction of $\tilde{\gamma}$ to the interval $[0, s]$ is larger than the distance between $\tilde{\gamma}(s)$ and $\partial B_{\mathcal{M}}(p, \frac{r}{2})$, which in turn is larger than $\frac{r}{2}$. Similarly, the length of the restriction of $\tilde{\gamma}$ to the interval $[s, 1]$ is larger than $\frac{r}{2}$. Hence, $r \leq \text{Length}(\tilde{\gamma})$ as desired.

Now, let $\tilde{\gamma}$ be a smooth curve realizing the distance between q_1 and q_2 (which after appropriate normalization has to be a geodesic). From the previous observation, we see that $\tilde{\gamma}$ is contained in $B_{\mathcal{M}}(p, r)$. Consider $\gamma := \exp_p^{-1} \circ \tilde{\gamma}$, where we note that $\exp_p^{-1}$ is well defined along $\tilde{\gamma}$ given that $r \leq i_0$. From the first part of the proposition, we deduce that

$$\frac{1}{2}d(\exp_p^{-1}(q_1), \exp_p^{-1}(q_2)) \leq \frac{1}{2}\text{Length}(\gamma) \leq \text{Length}(\tilde{\gamma}) = d(q_1, q_2).$$

Finally, for an arbitrary smooth curve $\gamma \colon [0, 1] \to B(r) \subseteq T_p\mathcal{M}$ with $\gamma(0) = \exp_p^{-1}(q_1)$ and $\gamma(i) = \exp_p^{-1}(q_2)$ we have

$$d(q_1, q_2) \leq \text{Length}(\exp_p \circ \gamma) \leq 2\,\text{Length}(\gamma).$$

Taking the infimum on the right-hand side over all such curves γ, we deduce that $d(q_1, q_2) \leq 2d(\exp_p^{-1}(q_1), \exp_p^{-1}(q_2))$. This completes the proof. $\qquad\square$

Proof *(Proof of Proposition* 2*)* The inequality $|x - y| \leq d(x, y)$ is trivial. To show the other inequality, we note that since $|x - y| \leq \frac{R}{2}$, it follows from [19, Prop 6.3] that

$$d(x, y) \leq R - R\sqrt{1 - \frac{2|x - y|}{R}}.$$

Using the fact that for every $t \in [0, 1]$, $\sqrt{1 - t} \geq 1 - \frac{1}{2}t - \frac{1}{2}t^2$

$$d(x, y) \leq R - R\left(1 - \frac{|x - y|}{R} - \frac{2}{R^2}|x - y|^2\right)$$

$$= |x - y| + \frac{2}{R}|x - y|^2 \leq 2|x - y|. \tag{A.1}$$

To improve the error estimate, let $L = d(x, y)$ and let $\gamma : [0, L] \to \mathcal{M}$ be an arc-length-parameterized length-minimizing geodesic between x and y. Heuristically, γ is a "straight" line in $\mathcal{M}$, and thus, its curvature in $\mathbb{R}^d$ is bounded by the maximal principal curvature of $\mathcal{M}$ in $\mathbb{R}^d$, which is bounded by $\frac{1}{R}$. More precisely, we claim that

$$|\ddot{\gamma}(t)| \leq \frac{1}{R} \qquad \text{for all } t \in [0, L]. \tag{A.2}$$

This statement follows from [19, Prop 6.1] (and is used in the proof of Proposition 6.3 of [19]). Using translation, we can assume that $x = 0$. Furthermore, note that that $\dot{\gamma}(t) \cdot \ddot{\gamma}(t) = 0$ for all t. Thus,

$$|x - y| = |\gamma(L)| \geq \gamma(L) \cdot \dot{\gamma}(L) = \int_0^L \dot{\gamma}(s) \cdot \dot{\gamma}(L) \mathrm{d}s$$

$$= \int_0^L \left(\dot{\gamma}(L) - \int_s^L \ddot{\gamma}(r)\mathrm{d}r\right) \cdot \dot{\gamma}(L) \, \mathrm{d}s \tag{A.3}$$

$$= L - \int_0^L \int_s^L \int_r^L \ddot{\gamma}(r) \cdot \ddot{\gamma}(z)\mathrm{d}z\mathrm{d}r\mathrm{d}s \geq L - \frac{L^3}{R^2}$$

Combining with (A.1) implies $L \leq |x - y| + \frac{8}{R^2}|x - y|^3$. $\qquad \square$

B Kernel-Density Estimates via Transportation

Here, we use the estimates on infinity transportation distance established in Sect. 2 to show the kernel density estimates we need. While the estimates we prove are not optimal, they do not affect the rate of convergence of eigenvalues and eigenfunctions in our main theorems. We chose to present the proof as follows as it highlights how the optimal transportation estimates can be used to provide general kernel-density estimates in a simple and direct way.

Lemma 18 *Consider $\eta : \mathbb{R} \to \mathbb{R}$, nonincreasing, supported on $[0, 1]$, and normalized: $\int_{\mathbb{R}^m} \eta(|x|)dx = 1$. Consider $h > 0$ satisfying Assumption 3. Then, (1.12) holds. That is, there exists a universal constant $C > 0$ such that*

$$\max_{i=1,\ldots,n} |m_i - p(\mathbf{x}_i)| \leq CL_p h + C\alpha\eta(0)m\omega_m \frac{\varepsilon}{h} + C\alpha m \left(K + \frac{1}{R^2} \right) h^2, \quad \text{(B.1)}$$

where ε is the ∞-OT distance between μ_n and μ (see Sect. 2).

The weights $\mathbf{m}$ are defined by

$$m_i = \frac{1}{nh^m} \sum_{j=1}^{n} \eta\left(\frac{|x_i - x_j|}{h} \right), \quad i = 1, \ldots, n,$$

p is the density of μ with respect to $\mathcal{M}$'s volume form. We remark that we do not require η to be Lipschitz on $[0, 1]$.

Proof First, notice that for every i, j with $|x_i - x_j| \leq h$ we have $|x_i - x_j| \leq \frac{R}{2}$, and hence, Proposition 2 implies that

$$d(x_i, x_j) \leq |x_i - x_j| + \frac{8}{R^2}|x_i - x_j|^3 \leq \left(1 + \frac{8h^2}{R^2} \right) |x_i - x_j|.$$

Therefore, for every i, j and every $y \in U_j$,

$$\eta\left(\frac{|x_i - x_j|}{h} \right) \leq \eta\left(\frac{d(x_i, x_j)}{\hat{h}} \right) \leq \eta\left(\frac{(d(x_i, y) - \varepsilon)_+}{\hat{h}} \right),$$

where we recall that ε is the ∞-OT distance between μ_n and μ and where $\hat{h} := h + \frac{27h^3}{R^2}$. From this, it follows that

$$m_i = \frac{1}{nh^m} \sum_{j=1}^{n} \eta\left(\frac{|\mathbf{x}_i - \mathbf{x}_j|}{h} \right) \leq \frac{1}{h^m} \int_{\mathcal{M}} \eta\left(\frac{(d(x_i, y) - \varepsilon)_+}{\hat{h}} \right) p(y)\mathrm{dVol}(y)$$

$$\leq (p(x_i) + 10L_p h)\frac{1}{h^m}$$

$$\int_{\mathcal{M}} \eta\left(\frac{(d(x_i, y) - \varepsilon)_+}{\hat{h}} \right) \mathrm{dVol}(y),$$

$$\text{(B.2)}$$

where the last inequality follows using the Lipschitz continuity of p, the fact that $\varepsilon < h$ and the fact that $h < \frac{R}{2}$ (so that in particular $\hat{h} + \varepsilon < 10h$). Now,

$$
\begin{aligned}
\frac{1}{h^m} \int_{\mathcal{M}} \eta \left(\frac{(d(x_i, y) - \varepsilon)_+}{\hat{h}} \right) d\mathrm{Vol}(y) &= \frac{1}{h^m} \int_{B(\hat{h}+\varepsilon)} \eta \left(\frac{(|z| - \varepsilon)_+}{\hat{h}} \right) J_{x_i}(z) dz \\
&\leq \frac{1}{h^m} \int_{B(\hat{h}+\varepsilon)} \eta \left(\frac{(|z| - \varepsilon)_+}{\hat{h}} \right) J_{x_i}(z) dz \\
&\leq (1 + CmKh^2) \frac{1}{h^m} \\
&\quad \int_{B(\hat{h}+\varepsilon)} \eta \left(\frac{(|z| - \varepsilon)_+}{\hat{h}} \right) dz,
\end{aligned}
$$
(B.3)

where C is a universal constant. The last integral above can be estimated as follows

$$
\begin{aligned}
\frac{1}{h^m} \int_{\mathbb{R}^m} \eta \left(\frac{(|z| - \varepsilon)_+}{\hat{h}} \right) dz &= \eta(0) \omega_m \frac{\varepsilon^m}{h^m} + \frac{1}{h^m} \int_{B(\hat{h}+\varepsilon) \backslash B(\varepsilon)} \eta \left(\frac{|z| - \varepsilon}{\hat{h}} \right) dz \\
&= \eta(0) \omega_m \frac{\varepsilon^m}{h^m} + \frac{\hat{h}^m}{h^m} \int_0^1 m \omega_m \left(r + \frac{\varepsilon}{\hat{h}} \right)^{m-1} \eta(r) \, dr \\
&\leq \eta(0) \omega_m \frac{\varepsilon^m}{h^m} + \left(1 + \frac{16mh^2}{R^2} \right) \\
&\quad \int_0^1 m \omega_m \left(r + \frac{\varepsilon}{h} \right)^{m-1} \eta(r) \, dr
\end{aligned}
$$
(B.4)

Using the binomial theorem, we obtain

$$
\begin{aligned}
m \omega_m \int_0^1 \left(r + \frac{\varepsilon}{h} \right)^{m-1} \eta(r) dr &\leq m \omega_m \int_0^1 r^{m-1} \eta(r) dr + m \omega_m \eta(0) \\
&\quad \sum_{k=1}^{m-1} \binom{m-1}{k} \left(\frac{\varepsilon}{h} \right)^k \frac{1}{m-k} \\
&= 1 + \omega_m \eta(0) \sum_{k=1}^{m-1} \binom{m}{k} \left(\frac{\varepsilon}{h} \right)^k \\
&= 1 + \omega_m \eta(0) \left(\left(1 + \frac{\varepsilon}{h} \right)^m - 1 - \frac{\varepsilon^m}{h^m} \right) \\
&\leq 1 + 2m \eta(0) \omega_m \frac{\varepsilon}{h} - \eta(0) \omega_m \frac{\varepsilon^m}{h^m}
\end{aligned}
$$

where in the first equality we have used the fact that η was assumed to be normalized, and in the last inequality, we have used

$$(1+s)^m \le 1 + 2ms \quad \text{whenever } 0 \le s < \frac{1}{m}.$$

Combining (B.2), (B.3) and (B.4), we conclude that

$$m_i - p(x_i) \le p(x_i) + CL_p h + C\alpha\eta(0)m\omega_m \frac{\varepsilon}{h} + C\alpha m \left(K + \frac{1}{R^2}\right)h^2,$$

for a universal constant $C > 0$.

In a similar fashion, we can find an upper bound for $p(x_i) - m_i$. Indeed, observe that for every i, j and $y \in U_i$ we have

$$\eta\left(\frac{|x_i - x_j|}{h}\right) \ge \eta\left(\frac{d(x_i, x_j)}{h}\right) \ge \eta\left(\frac{d(x_i, y) + \varepsilon}{h}\right)$$

and so

$$
\begin{aligned}
m_i &\ge \frac{1}{h^m} \int_{\mathcal{M}} \eta\left(\frac{d(x_i, y) + \varepsilon}{h}\right) p(y)\mathrm{dVol}(y) \\
&\ge \frac{1}{h^m} \int_{\mathcal{M}} \eta\left(\frac{d(x_i, y) + \varepsilon}{h}\right) (p(x_i) - L_p d(x_i, y))\mathrm{dVol}(y) \qquad \text{(B.5)} \\
&\ge (p(x_i) - L_p h)\frac{1}{h^m} \int_{\mathcal{M}} \eta\left(\frac{d(x_i, y) + \varepsilon}{h}\right) \mathrm{dVol}(y).
\end{aligned}
$$

The above integral can be estimated from below by

$$
\begin{aligned}
\frac{1}{h^m} \int_{\mathcal{M}} \eta\left(\frac{d(x_i, y) + \varepsilon}{h}\right) \mathrm{dVol}(y) &= \frac{1}{h^m} \int_{B(h-\varepsilon)} \eta\left(\frac{|z| + \varepsilon}{h}\right) J_{x_i}(z)\mathrm{d}z \\
&\ge (1 - CmKh^2)\frac{1}{h^m} \int_{B(h-\varepsilon)} \eta\left(\frac{|z| + \varepsilon}{h}\right) \mathrm{d}z \\
&= (1 - CmKh^2) \int_{\varepsilon/h}^{1} m\omega_m \eta(r)(r - \frac{\varepsilon}{h})^{m-1}\mathrm{d}r
\end{aligned}
$$
$$\text{(B.6)}$$

where the second equality follows using polar coordinates and a change in variables; the last inequality follows from the fact that η is assumed to be normalized. In turn,

$$
\int_{\varepsilon/h}^{1} m\omega_m \eta(r) \left(r - \frac{\varepsilon}{h} \right)^{m-1} \mathrm{d}r \geq \int_{\varepsilon/h}^{1} m\omega_m \eta(r) r^{m-1} \mathrm{d}r
$$
$$
- m\omega_m \frac{\varepsilon}{h} \int_{\varepsilon/h}^{1} (m-1)\eta(r) r^{m-2} \mathrm{d}r
$$
$$
\geq 1 - 2\eta(0)m\omega_m \frac{\varepsilon}{h},
$$

where we have used the fact that η was assumed to be normalized. Combining the above inequalities, we deduce that

$$
p(x_i) - m_i \leq L_p h + C\alpha m\omega_m K h^2 + C\alpha m\omega_m \eta(0)\frac{\varepsilon}{h}.
$$

$\square$

References

1. W. Arendt and A. F. M. ter Elst, *Sectorial forms and degenerate differential operators*, J. Operator Theory, 67 (2012), pp. 33–72.
2. M. Belkin and P. Niyogi, *Laplacian eigenmaps for dimensionality reduction and data representation*, Neural Computation, 15 (2002), pp. 1373–1396.
3. M. Belkin and P. Niyogi, *Convergence of Laplacian eigenmaps*, Advances in Neural Information Processing Systems (NIPS), 19 (2007), p. 129.
4. M. Belkin and P. Niyogi, *Towards a theoretical foundation for Laplacian-based manifold methods*, J. Comput. System Sci., 74 (2008), pp. 1289–1308.
5. A. L. Besse, *Manifolds all of whose geodesics are closed*, vol. 93 of Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas], Springer-Verlag, Berlin-New York, 1978. With appendices by D. B. A. Epstein, J.-P. Bourguignon, L. Bérard-Bergery, M. Berger and J. L. Kazdan.
6. D. Burago, S. Ivanov, and Y. Kurylev, *A graph discretization of the Laplace-Beltrami operator*, J. Spectr. Theory, 4 (2014), pp. 675–714.
7. I. Chavel, *Eigenvalues in Riemannian geometry*, Academic Press, New York, 1984.
8. R. R. Coifman and S. Lafon, *Diffusion maps*, Appl. Comput. Harmon. Anal., 21 (2006), pp. 5–30.
9. M. P. do Carmo, *Riemannian geometry*, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.
10. K. Fujiwara, *Eigenvalues of Laplacians on a closed riemannian manifold and its nets*, Proc. Amer. Math. Soc., 123 (1995), pp. 2585–2594.
11. N. García Trillos and D. Slepčev, *On the rate of convergence of empirical measures in ∞-transportation distance*, Canad. J. Math., 67 (2015), pp. 1358–1383.
12. N. García Trillos and D. Slepčev, *A variational approach to the consistency of spectral clustering*, Appl. Comput. Harmon. Anal., 45 (2018), pp. 239–281.
13. E. Giné and V. Koltchinskii, *Empirical graph Laplacian approximation of Laplace-Beltrami operators: large sample results*, in High dimensional probability, vol. 51 of IMS Lecture Notes Monogr. Ser., Inst. Math. Statist., Beachwood, OH, 2006, pp. 238–259.
14. M. Hein, *Uniform convergence of adaptive graph-based regularization*, in Proc. of the 19th Annual Conference on Learning Theory (COLT), G. Lugosi and H. U. Simon, eds., Springer, 2006, pp. 50–64.
15. M. Hein, J.-Y. Audibert, and U. v. Luxburg, *Graph Laplacians and their convergence on random neighborhood graphs*, Journal of Machine Learning Research, 8 (2007), pp. 1325–1368.

16. T. Leighton and P. Shor, *Tight bounds for minimax grid matching with applications to the average case analysis of algorithms*, Combinatorica, 9 (1989), pp. 161–187.
17. B. Mohar, *Some applications of Laplace eigenvalues of graphs*, in Graph Theory, Combinatoris and Applications, Y. Alavi, G. Chartrand, O. R. Oellermann, and A. J. Schwenk, eds., Wiley, 1991, pp. 871–898.
18. D. Mugnolo and R. Nittka, *Convergence of operator semigroups associated with generalised elliptic forms*, J. Evol. Equ., 12 (2012), pp. 593–619.
19. P. Niyogi, S. Smale, and S. Weinberger, *Finding the homology of submanifolds with high confidence from random samples*, Discrete Comput. Geom., 39 (2008), pp. 419–441.
20. M. Penrose, *Random geometric graphs*, vol. 5 of Oxford Studies in Probability, Oxford University Press, Oxford, 2003.
21. L. Rosasco, M. Belkin, and E. D. Vito, *On learning with integral operators*, Journal of Machine Learning Research, 11 (2010), pp. 905–934.
22. Y. SHI AND B. XU, *Gradient estimate of an eigenfunction on a compact Riemannian manifold without boundary*, Ann. Global Anal. Geom., 38 (2010), pp. 21–26.
23. Z. Shi, *Convergence of Laplacian spectra from random samples*. preprint, arXiv:1507.00151, 2015.
24. P. W. Shor and J. E. Yukich, *Minimax grid matching and empirical measures*, Ann. Probab., 19 (1991), pp. 1338–1348.
25. A. Singer, *From graph to manifold Laplacian: The convergence rate*, Applied and Computational Harmonic Analysis, 21 (2006), pp. 128–134.
26. 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.
27. M. Talagrand, *The generic chaining*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2005. Upper and lower bounds of stochastic processes.
28. D. Ting, L. Huang, and M. I. Jordan, *An analysis of the convergence of graph Laplacians*, in Proc. of the 27th Int. Conference on Machine Learning (ICML), 2010.
29. U. von Luxburg, *A tutorial on spectral clustering*, Statistics and computing, 17 (2007), pp. 395–416.
30. U. von Luxburg, M. Belkin, and O. Bousquet, *Consistency of spectral clustering*, Ann. Statist., 36 (2008), pp. 555–586.

Springer