Journal of Machine Learning Research 24 (2023) 1-85 Submitted 10/22; Revised 5/23; Published 11/23

Radial Basis Approximation of Tensor Fields on Manifolds: From Operator Estimation to Manifold Learning

John Harlim JHARLIM@PSU.EDU
Department of Mathematics
Department of Meteorology and Atmospheric Science
Institute for Computational and Data Sciences
The Pennsylvania State University
University Park, PA 16802, USA

Shixiao Willing Jiang JIANGSHX@SHANGHAITECH.EDU.CN
Institute of Mathematical Sciences
ShanghaiTech University
Shanghai 201210, China

John Wilson Peoples JWP5828@PSU.EDU
Department of Mathematics
The Pennsylvania State University
University Park, PA 16802, USA

Editor: Michael Mahoney

Abstract

In this paper, we study the Radial Basis Function (RBF) approximation to differential operators on smooth tensor fields defined on closed Riemannian submanifolds of Euclidean space, identified by randomly sampled point cloud data. The formulation in this paper leverages a fundamental fact that the covariant derivative on a submanifold is the projection of the directional derivative in the ambient Euclidean space onto the tangent space of the submanifold. To differentiate a test function (or vector field) on the submanifold with respect to the Euclidean metric, the RBF interpolation is applied to extend the function (or vector field) in the ambient Euclidean space. When the manifolds are unknown, we develop an improved second-order local SVD technique for estimating local tangent spaces on the manifold. When the classical pointwise non-symmetric RBF formulation is used to solve Laplacian eigenvalue problems, we found that while accurate estimation of the leading spectra can be obtained with large enough data, such an approximation often produces irrelevant complex-valued spectra (or pollution) as the true spectra are real-valued and positive. To avoid such an issue, we introduce a symmetric RBF discrete approximation of the Laplacians induced by a weak formulation on appropriate Hilbert spaces. Unlike the non-symmetric approximation, this formulation guarantees non-negative real-valued spectra and the orthogonality of the eigenvectors. Theoretically, we establish the convergence of the eigenpairs of both the Laplace-Beltrami operator and Bochner Laplacian for the symmetric formulation in the limit of large data with convergence rates. Numerically, we provide supporting examples for approximations of the Laplace-Beltrami operator and various vector Laplacians, including the Bochner, Hodge, and Lichnerowicz Laplacians.

Keywords: Radial Basis Functions (RBFs), Laplace-Beltrami operators, vector Laplacians, manifold learning, operator estimation, local SVD

1. Introduction

Estimation of differential operators is an important computational task in applied mathematics and engineering science. While this estimation problem has been studied since the time of Euler in the mid-18th century, numerical differential equations emerges as an important sub-field of computational mathematics in the 1940s when modern computers were starting to be developed for solving differential equations. Among many available numerical methods, finite-difference, finite-volume, and finite-element methods are considered the most reliable algorithms to produce accurate solutions whenever one can control the distribution of points (or nodes) and meshes. Specification of the nodes, however, requires some knowledge of the domain and is subjected to the curse of dimension.

On the other hand, the Radial Basis Function (RBF) method (Buhmann, 2003) has been considered as a promising alternative that can produce very accurate approximations (Narcowich and Ward, 1991; Wu and Schaback, 1993) even with randomly distributed nodes (Kanagawa et al., 2018) on high-dimensional domains. Beyond the mesh-free approximation of differential operators, the deep connection of the RBF to kernel technique has also been documented in nonparametric statistical literature (Christmann and Steinwart, 2008; Kanagawa et al., 2018) for machine learning applications. While kernel methods play a significant role in supervised machine learning (Christmann and Steinwart, 2008; Cucker and Smale, 2002), in unsupervised learning, a kernel approach usually corresponds to the construction of a graph whose spectral properties (Chung, 1997; Von Luxburg et al., 2008) can be used for clustering (Ng et al., 2002), dimensionality reduction (Belkin and Niyogi, 2003), and manifold learning (Coifman and Lafon, 2006), among others. In the context of manifold learning, given a set of data that lie on a $d-$dimensional compact Riemannian sub-manifold of Euclidean domain $\mathbb{R}^n$, the objective is to represent observable with eigensolutions of an approximate Laplacian operator. For this purpose, spectral convergence type results, concerning the convergence of the graph Laplacian matrix induced by exponentially decaying kernels to the Laplacian operator on functions, are well-documented for closed manifolds (Belkin and Niyogi, 2007; Burago et al., 2014; Trillos et al., 2020; Calder and Trillos, 2022; Dunson et al., 2021) and for manifolds with boundary (Peoples and Harlim, 2021). Beyond Laplacian on functions, graph-based approximation of connection Laplacian on vector fields (Singer and Wu, 2017) and a Galerkin-based approximation to Hodge Laplacian on 1-forms (Berry and Giannakis, 2020) have also been considered separately. Since these manifold learning methods fundamentally approximate Laplacian operators that act on smooth tensor fields defined on manifolds, it is natural to ask whether this learning problem can be solved using the RBF method. Another motivating point is that the available graph-based approaches can only approximate a limited type of differential operators whereas RBF can approximate arbitrary differential operators, including general k-Laplacians.

Indeed, RBF has been proposed to solve PDEs on 2D surfaces (Fuselier and Wright, 2009, 2012; Piret, 2012; Fuselier and Wright, 2013). In these papers, they showed that RBF solutions converge, especially when the point clouds are appropriately placed which requires some parameterization of the manifolds. When the surface parameterization is unknown, there are several approaches to characterize the manifolds, such as the closest point method (Ruuth and Merriman, 2008), the orthogonal gradient (Piret, 2012), the moving least squares (Liang and Zhao, 2013), and the local SVD method (Donoho and

Grimes, 2003; Zhang and Zha, 2004; Tyagi et al., 2013). The first two methods require laying additional grid points on the ambient space, which can be numerically expensive when the ambient dimension is high. The moving least squares locally fit multivariate quadratic functions of the local coordinates approximated by PCA (or local SVD) on the metric tensor. While fast theoretical convergence rate when the data point is well sampled (see the detailed discussion in Section 2.2 of Liang and Zhao (2013)), based on the presented numerical results in the paper, it is unclear whether the same convergence rate can be achieved when the data are randomly sampled. The local SVD method, which is also used in determining the local coordinates in the moving least-squares method, can readily approximate the local tangent space for randomly sampled training data points. This information alone readily allows one to approximate differential operators on the manifolds.

From the perspective of approximation theory, the radial basis type kernel is universal in the sense that the induced Reproducing Kernel Hilbert Space (RKHS) is dense in the space of continuous and bounded function on a compact domain, under the standard uniform norm (Steinwart, 2001; Sriperumbudur et al., 2011). While this property is very appealing, previous works on RBF suggest that the non-symmetric pointwise approximation to the Laplace-Beltrami operator can be numerically problematic (Fuselier and Wright, 2009; Piret, 2012; Fuselier and Wright, 2013). In particular, they numerically reported that when the number of training data is small the eigenvalues of the non-symmetric RBF Laplacian matrix that approximates the negative-definite Laplace-Beltrami operator are not only complex-valued, but they can also be on the positive half-plane. These papers also empirically reported that this issue, which is related to spectral pollution (Llobet et al., 1990; Davies and Plum, 2004; Lewin and Séré, 2010), can be overcome with more data points. The work in this paper is somewhat motivated by many open questions from these empirical results.

1.1 Contribution of This Paper and a Summary of Our Findings

One of the objectives of this paper is to assess the potential of the RBF method in solving the manifold learning problem, involving approximating the Laplacians acting on functions and vector fields of smooth manifolds, where the underlying manifold is identified by a set of randomly sample point cloud data. The work in this paper extends the fundamental fact that the covariant derivative on a submanifold is the projection of the directional derivative in the ambient Euclidean space onto the tangent space of the submanifold to various differential operators, including Laplacians on functions and vector fields. Since the formulation involves the ambient dimension, n, the resulting approximation will be computationally feasible for problems where the ambient dimension is much smaller than the data size, $n \ll N$. While such dimensionality scaling requires an extensive computational memory for manifold learning problems with $n = O(N)$, it can still be useful for a broad class of applications involving eigenvalue problems and PDEs where the ambient dimension is moderately high, $n = 10 - 100$, but much smaller than N. See eigenvalue problem examples in Sections 5.2 and 5.3. Also see e.g., the companion paper (Yan et al., 2023) that uses the symmetric approximation discussed below to solve elliptic PDEs on unknown manifolds.

First, we study the non-symmetric Laplacian RBF matrix, which is a pointwise approximation to the Laplacian operator, that is used in (Fuselier and Wright, 2009; Piret, 2012)

to approximate the Laplace-Beltrami operator. Through a numerical study in Section 5, we found that non-symmetric RBF (NRBF) formulation can produce a very accurate estimation of the leading spectral properties whenever the number of sample points used to approximate the RBF matrix is large enough and the local tangent space is sufficiently accurately approximated. In this paper:

1a) We propose an improved local SVD algorithm for approximating the local tangent spaces of the manifolds, where the improvement is essentially due to the additional steps designed to correct errors induced by the curvatures (see Section 3.2). We provide a theoretical error bound (see Theorem 3.2). We numerically find that this error (from the local tangent space approximation) dominates the errors induced by the non-symmetric RBF approximation of the leading eigensolutions (see Figures 7 and 10). On the upside, since this local SVD method is numerically not expensive even with very large data, applying NRBF with the accurately estimated local tangent space will give a very accurate estimation of the leading spectra with possibly expensive computational costs. Namely, solving eigenvalue problems of non-symmetric, dense discrete NRBF matrices, which can be very large, especially for the approximation of vector Laplacians. On the downside, since this pointwise estimation relies on the accuracy of the local tangent space approximation, such a high accuracy will not be attainable when the data is corrupted by noise.

1b) Through numerical verification in Sections 5.3 and 5.4, we detect another issue with the non-symmetric formulation. That is, when the training data size is not large enough, the non-symmetric RBF Laplacian matrix is subjected to spectral pollution (Llobet et al., 1990; Davies and Plum, 2004; Lewin and Séré, 2010). Specifically, the resulting matrix possesses eigenvalues that are irrelevant to the true eigenvalues. If we increase the training data, while the leading eigenvalues (that are closer to zero) are accurately estimated, the irrelevant estimates will not disappear. Their occurrence will be on higher modes. This issue can be problematic in manifold learning applications since the spectra of the underlying Laplacian to be estimated are unknown. As we have pointed out in 1a), large data size may not be numerically feasible for the non-symmetric formulation, especially in solving eigenvalue problems corresponding to Laplacians on vector fields.

To overcome the limitation of the non-symmetric formulation, we consider a symmetric discrete formulation induced by a weak approximation of the Laplacians on appropriate Hilbert spaces. Several advantages of this symmetric approximation are that the estimated eigenvalues are guaranteed to be non-negative real-valued, and the corresponding estimates for the eigenvectors (or eigenvector fields) are real-valued and orthogonal. Here, the Laplace-Beltrami operator is defined to be semi-positive definite. The price we are paying to guarantee estimators with these nice properties is that the approximation is less accurate compared to the non-symmetric formulation provided that the latter works. Particularly, the error of the symmetric RBF is dominated by the Monte-Carlo rate. Our findings are based on:

2a) A spectral convergence study with error bounds for the estimations of eigenvalues and eigenvectors (or eigenvector fields). See Theorems 4.1 and 4.2 for the approxima-

tion of eigenvalues and eigenfunctions of Laplace-Beltrami operator, respectively. See Theorems 4.3 and 4.4 for the approximation of eigenvalues and eigenvector fields of Bochner Laplacian, respectively.

2b) Numerical inspections on the estimation of Laplace-Beltrami operator. We show the empirical convergence as a function of training data size. Based on our numerical comparison on a 2D torus embedded in $\mathbb{R}^{21}$, we found that the symmetric RBF produces less accurate estimates compared to the Diffusion Maps (DM) algorithm (Coifman and Lafon, 2006) in the estimation of leading eigenvalues, but more accurate in the estimation of non-leading eigenvalues.

2c) Numerical inspections on the estimation of Bochner, Hodge, and Lichnerowicz Laplacians. We show the empirical convergence as a function of training data size.

1.2 Organization of This Paper

Section 2: We provide a detailed formulation for discrete approximation of differential operators on smooth manifolds, where we will focus on the RBF technique as an interpolant. Overall, the nature of the approximation is "exterior" in the sense that the tangential derivatives will be represented as a projection (or restriction) of appropriate ambient derivatives onto the local tangent space. To clarify this formulation, we overview the notion of the projection matrix that allows the exterior representation that will be realized through appropriate discrete tensors. We give a concrete discrete formulation for gradient, the Laplace-Beltrami operator, covariant derivative, and the Bochner Laplacian. We discuss the symmetric and non-symmetric formulations of Laplacians. We list the RBF approximation in Table 2, where we also include the estimation of the Hodge and Lichnerowicz Laplacian (which detailed derivations are reported in Appendix A).
Section 3: We present a novel algorithm to improve the characterization of the manifold from randomly sampled point cloud data, which subsequently allows us to improve the approximation of the projection matrix, which is explicitly not available when the manifold is unknown. The new local SVD method, which accounts for curvature errors, will improve the accuracy of RBF in the pointwise estimation of arbitrary differential operators.
Section 4: We deduce the spectral convergence of the symmetric estimation of the Laplace-Beltrami operator and the Bochner Laplacian. Error bounds for both the eigenvalues and eigenfunctions estimations will be given in terms of the number of training data, the smoothness, and the dimension and co-dimension of the manifolds. These results rely on a probabilistic type convergence result of the RBF interpolation error, extending the deterministic error bound of the interpolation using the RKHS induced by the Matérn kernel (Fuselier and Wright, 2012), which is reported in Appendix B. To keep the section short, we only present the proof for the spectral convergence of the Laplace-Beltrami operator. We document the proofs of the intermediate bounds needed for this proof in Appendices B, C.1 and C.2. Since the proof of the eigenvector estimation is more technical, we also present it in Appendix C.3. Since the proofs of the Bochner Laplacian approximation follow the same arguments as those for the Laplace-Beltrami operator, we document them in Appendix D.
Section 5: We present numerical examples to inspect the non-symmetric and symmetric RBF spectral approximations. The first two examples focus on the approximations of

Laplace-Beltrami on functions. In the first example (the two-dimensional general torus), we verify the effectiveness of the approximation when the low-dimensional manifold is embedded in a high-dimensional ambient space (high co-dimension). In this example, we also compare the results from a graph-based approach, diffusion maps (Coifman and Lafon, 2006). To ensure the diffusion maps result is representative, we report additional numerical results over an extensive parameter tuning in Appendix E. In the second example (four- and five-dimensional flat torus), our aim is to verify the effectiveness of the approximation when the intrinsic dimension of the manifold is higher than that in the first example. In the third example, we verify the accuracy of the spectral estimation of the Bochner, Hodge, and Lichnerowicz Laplacians on the sphere.

Section 6: We close this paper with a short summary and discussion of remaining and emerging open problems.

For reader's convenience, we provide a list of notations in Table 1.

2. Basic Formulation for Estimating Differential Operators on Manifolds

In this section, we first review the approximations of the gradient, the Laplace-Beltrami operator, and covariant derivative and then formulate the detailed discrete approximations of the connection of a vector field and the Bochner Laplacian, on smooth closed manifolds. Both the Laplace-Beltrami operator and the Bochner Laplacian have two natural discrete estimators: symmetric and non-symmetric formulations. Since each formulation has its own practical and theoretical advantages, we describe in detail both approximations. Following similar derivations, we also report the discrete approximation to other differential operators that are relevant to vector fields, e.g. the Hodge and Lichnerowicz Laplacian (see Appendix A for the detailed derivations).

In this paper, we consider estimating differential operators acting on functions (and vector fields) defined on a $d-$dimensional closed manifold M embedded in ambient space $\mathbb{R}^n$, where $d \leq n$. Each operator is estimated using an ambient space formulation followed by a projection onto the local tangent space of the manifold using a projection matrix $\mathbf{P}$. Before describing the discrete approximation of the operators, we introduce $\mathbf{P}$, discuss some of its basic properties, and quickly review the radial basis function (RBF) interpolation which is a convenient method to approximate functions and tensor fields from the point cloud training data $X = \{x_i\}_{i=1}^N$.

In the following, we periodically use the notation $\text{diag}(a_1, \ldots, a_j)$ to denote a $j \times j$ diagonal matrix with the listed entries along the diagonal. We also define $f \mapsto R_N f = \mathbf{f} = (f(x_1), f(x_2), \ldots, f(x_N))^\top$ for all $f : M \to \mathbb{R}$, which is a restriction operator to the function values on training data set $X = \{x_1, \ldots, x_N\}$. In the remainder of this paper, we use boldface to denote discrete objects (vectors, matrices) and script font to denote operators on continuous objects.

Definition 1 *For any point $x \in M$, the local parameterization $\iota : O \subseteq \mathbb{R}^d \longrightarrow M \subseteq \mathbb{R}^n$, is defined through the following map, $\mathbf{\Theta}_{\iota^{-1}(x)} \longmapsto \mathbf{X}_x$. Here, O denotes a domain that contains the point $\iota^{-1}(x)$, which we denoted as $\mathbf{\Theta}_{\iota^{-1}(x)}$ in the canonical coordinates $\left(\frac{\partial}{\partial\theta^1}\Big|_{\iota^{-1}(x)}, \ldots, \frac{\partial}{\partial\theta^d}\Big|_{\iota^{-1}(x)}\right)$ and $\mathbf{X}_x$ is the embedded point represented in the ambient coor-*

Table 1: List of notations. The notations below are used throughout the entire manuscript, and defined in Sections 2-4.

Symbol	Definition	Symbol	Definition
Sec.	**2**		
M	a manifold	$\mathbb{R}$	Euclidean space
d	intrinsic dimension of M	n	ambient space dimension
θ^i	intrinsic coordinate $i = 1, \dots, d$	X^i	ambient coordinate $i = 1, \dots, n$
$\mathbf{P}$	$P(x)$, tangential project matrix	$\mathcal{P}$	tangential projection tensor in (14)
P_{ij}	entries of projection matrix $i, j = 1, \dots, n$	$\mathbf{p}_i$	ith column of projection matrix $\mathbf{P}$ $\mathbf{p}_i = (P_{1i}(x), \dots, P_{ni}(x))$
ι	the local parameterization of the manifold	$D\iota$	the pushforward or the $n \times d$ Jacobian matrix
N	number of data points	X	$X = \{x_i\}_{i=1}^N = \{x_1, \dots, x_N\}$
R_N	restriction operator, $R_N f = \mathbf{f} = (f(x_1), \dots, f(x_N))^\top$	$\|_M$, $\|_X$	$\|_M$ is the restriction on manifold M $\|_X$ is the restriction on data set X
f	a function	$\mathbf{f}$	a function on data points, $\mathbf{f} = (f(x_1), \dots, f(x_N))^\top$
$I_{\phi_s}\mathbf{f}$	interpolation of $\mathbf{f}$ using RBF ϕ_s	s	shape parameter
$\mathbf{\Phi}$	interpolation matrix	$\mathbf{c}$	interpolation coefficient, $(c_1, \dots, c_N)^\top$
grad_g	gradient w.r.t. Riemannian metric, $\mathrm{grad}_g f = g^{ij}\frac{\partial f}{\partial \theta^i}\frac{\partial}{\partial \theta^j}$	$\overline{\mathrm{grad}}_{\mathbb{R}^n}$	gradient in Euclidean space, $\overline{\mathrm{grad}}_{\mathbb{R}^n} f = \delta^{ij}\frac{\partial f}{\partial X^i}\frac{\partial}{\partial X^j}$
$\mathcal{G}_i$	$\mathcal{G}_i = \mathbf{p}_i \cdot \mathrm{grad}_{\mathbb{R}^n}$	$\mathbf{G}_i$	$N \times N$ matrix that approximates $\mathcal{G}_i$
$\mathbf{T}, \boldsymbol{\tau}_i$	d orthonormal tangent vectors, $\mathbf{T} = [\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_d]$	$\mathbf{N}, \mathbf{n}_i$	vectors that are orthogonal to tangent space, $\mathbf{N} = [\mathbf{n}_1, \dots, \mathbf{n}_{n-d}]$
∇	Levi-Civita connection on M	$\bar{\nabla}$	Euclidean connection on $\mathbb{R}^n$
Δ_M	Laplace-Beltrami	Δ_B	Bochner Laplacian
$q(x)$	sampling density	$\mathbf{Q}$	diagonal matrix with entries $q(x_i)$
$\mathcal{H}_i$	$\mathcal{H}_i U(x) = \mathbf{P}\mathrm{diag}(\mathcal{G}_i, \dots, \mathcal{G}_i)U(x)$	$\mathbf{H}_i$	$Nn \times Nn$ matrix that estimates $\mathcal{H}_i$
Sec.	**3**		
$\boldsymbol{\rho}$	geodesic normal coordinate, $\boldsymbol{\rho} = (\rho_1, \dots, \rho_d)$	ρ	geodesic distance, $\rho = \|\boldsymbol{\rho}\|$
$\tilde{\mathbf{P}}$	1st-order SVD estimate of $\mathbf{P}$	$\hat{\mathbf{P}}$	2nd-order SVD estimate of $\mathbf{P}$
K	K-nearest neighbors	K_{max}	maximum principal curvature
D	$K > D := \frac{1}{2}d(d+1)$, the minimum number for the K-nearest neighbors	$\mathbf{D}$	$\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_K]$, $\mathbf{D}_i = y_i - x$ is the ith column of $\mathbf{D} \in \mathbb{R}^{n \times K}$
$\mathbf{\Psi}$	matrix with orthonormal tangent column vectors $\mathbf{\Psi} = [\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_d]$	$\hat{\mathbf{\Psi}}$	$\hat{\mathbf{\Psi}} = [\hat{\boldsymbol{\psi}}_1, \dots, \hat{\boldsymbol{\psi}}_d] \in \mathbb{R}^{n \times d}$ is an estimator of $\mathbf{\Psi} = [\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_d]$
Sec.	**4**		
$\langle f, h\rangle_{L^2(M)}$	inner product in $L^2(M)$ defined as $\langle f, h\rangle_{L^2(M)} = \int_M fh \, d\mathrm{Vol}$	$\langle \mathbf{f}, \mathbf{h}\rangle_{L^2(\mu_N)}$	appropriate inner product defined as $\langle \mathbf{f}, \mathbf{h}\rangle_{L^2(\mu_N)} = \frac{1}{N}\mathbf{f}^\top\mathbf{h}$
λ_i	eigenvalues	$\hat{\lambda}_i$	approximate eigenvalues
m	the geometric multiplicity of eigenvalues		

dinates $\left(\frac{\partial}{\partial X^1}\Big|_x, \ldots, \frac{\partial}{\partial X^n}\Big|_x\right)$. *The pushforward* $D\iota(x) : T_{\iota^{-1}(x)}O \to T_xM$ *is an* $n \times d$ *matrix given by* $D\iota(x) = \left[\frac{\partial \mathbf{X}_x}{\partial \theta^1}, \ldots, \frac{\partial \mathbf{X}_x}{\partial \theta^d}\right]$, *a matrix whose columns form a basis of* T_xM *in* $\mathbb{R}^n$.

Definition 2 *The projection matrix* $\mathbf{P} = P(x) \in \mathbb{R}^{n\times n}$ *on* $x \in M$ *is defined with the matrix-valued function* $P : M \to \mathrm{Proj}(n, \mathbb{R}) \subset \mathbb{R}^{n\times n}$,

$$P = [P_{st}]_{s,t=1}^n := \left[\frac{\partial X^s}{\partial \theta^i} g^{ij} \frac{\partial X^t}{\partial \theta^j}\right]_{s,t=1}^n,$$

where g^{ij} *denotes the matrix entries of the inverse of the Riemannian metric tensor* g_{ij}. *Since* $[g^{ij}]_{d\times d} = (D\iota^\top D\iota)^{-1}$, *the projection matrix can be equivalently defined as*

$$P := D\iota(D\iota^\top D\iota)^{-1} D\iota^\top. \tag{1}$$

Before proving some basic properties of the projection matrix $\mathbf{P}$, we must fix a set of orthonormal vectors that span the tangent space T_xM for each $x \in M$. In particular, for any $x \in M$, let $\{\boldsymbol{\tau}_i \in \mathbb{R}^{n\times 1}\}_{i=1}^d$ be the d orthonormal vectors that span T_xM and let $\{\mathbf{n}_i \in \mathbb{R}^{n\times 1}\}_{i=1}^{n-d}$ be the $n-d$ orthonormal vectors that are orthogonal to T_xM. Here, we suppress the dependence of $\boldsymbol{\tau}_i$ and $\mathbf{n}_i$ on x to simplify the notation. Further, let $\mathbf{T} = [\boldsymbol{\tau}_1 \ \ldots \ \boldsymbol{\tau}_d] \in \mathbb{R}^{n\times d}$ and $\mathbf{N} = [\mathbf{n}_1 \ \ldots \ \mathbf{n}_{n-d}] \in \mathbb{R}^{n\times(n-d)}$. Since $\begin{pmatrix}\mathbf{T} & \mathbf{N}\end{pmatrix} \in \mathbb{R}^{n\times n}$ is an orthonormal matrix, one has the following relation,

$$\mathbf{I} = \begin{pmatrix}\mathbf{T} & \mathbf{N}\end{pmatrix} \begin{pmatrix}\mathbf{T}^\top \\ \mathbf{N}^\top\end{pmatrix} = \mathbf{T}\mathbf{T}^\top + \mathbf{N}\mathbf{N}^\top = \sum_{i=1}^d \boldsymbol{\tau}_i \boldsymbol{\tau}_i^\top + \sum_{i=1}^{n-d} \mathbf{n}_i \mathbf{n}_i^\top.$$

We have the following proposition summarizing the basic properties of $\mathbf{P}$.

Proposition 2.1 *For any* $x \in M$, *let* $\mathbf{P} = P(x) \in \mathbb{R}^{n\times n}$ *be the projection matrix defined in Definition 2 and let* $\mathbf{T} = [\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_d] \in \mathbb{R}^{n\times d}$ *be any* d *orthonormal vectors that span* T_xM. *Then*
(1) $\mathbf{P}$ *is symmetric;*
(2) $\mathbf{P}^2 = \mathbf{P}$*;*
(3) $\mathbf{P} = P(x) = D\iota(x)(D\iota(x)^\top D\iota(x))^{-1} D\iota(x)^\top = \mathbf{T}\mathbf{T}^\top$.
(4) $\sum_{i=1}^n |\mathbf{p}_i|^2 = d$, *where* $\mathbf{p}_i = (P_{1i}(x), \ldots, P_{ni}(x))^\top$ *is the* i*th column of* $\mathbf{P}$.

Proof Properties (1) and (2) are obvious from the definition of $\mathbf{P}$ in (1). Property (3) can be easily obtained by observing that both sides of the equation are orthogonal projections, and span $\left\{\frac{\partial \mathbf{X}_x}{\partial \theta^1}, \ldots, \frac{\partial \mathbf{X}_x}{\partial \theta^d}\right\} = \mathrm{span}\{\boldsymbol{\tau}_1 \ \ldots \ \boldsymbol{\tau}_d\}$. To see (4), for each point $x \in M$, write $\mathbf{P}$ as

$$\mathbf{P} = \mathbf{T}\mathbf{T}^\top = \begin{bmatrix} P_{11}(x) & \cdots & P_{1n}(x) \\ \vdots & \ddots & \vdots \\ P_{n1}(x) & \cdots & P_{nn}(x) \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1^\top \\ \vdots \\ \mathbf{p}_n^\top \end{bmatrix},$$

where $\mathbf{p}_i \in \mathbb{R}^{n\times 1}$ for $i = 1, \ldots, n$ is the ith column of $\mathbf{P}$**.** It remains to observe the following chain of equalities:

$$\sum_{i=1}^n |\mathbf{p}_i|^2 = \mathrm{tr}\left(\mathbf{P}^\top\mathbf{P}\right) = \mathrm{tr}\left(\mathbf{P}\right) = \mathrm{tr}\left(\mathbf{T}\mathbf{T}^\top\right) = \mathrm{tr}\left(\mathbf{T}^\top\mathbf{T}\right) = \mathrm{tr}\left(\mathbf{I}_{d\times d}\right) = d.$$

■

Notice that while the specification of $\mathbf{T}$ is not unique which can be different by a rotation, the projection matrix $\mathbf{P}$ is uniquely determined at each point $x \in M$.

Let $f : M \to \mathbb{R}$ be an arbitrary some smooth function. Given function values $\mathbf{f} := (f(x_1), \ldots, f(x_N))^\top$ at $X = \{x_j\}_{j=1}^N$, the radial basis function (RBF) interpolant of f at x takes the form

$$I_{\phi_s}\mathbf{f}(x) := \sum_{k=1}^{N} c_k \phi_s \left(\|x - x_k\| \right), \tag{2}$$

where ϕ_s denotes the kernel function with shape parameter s and $\|\cdot\|$ denotes the standard Euclidean distance. We should point out that by being a kernel, it is positive definite as in the standard nomenclature (see e.g., Theorem 4.16 in Christmann and Steinwart (2008)). Here, one can interpret $I_{\phi_s} : \mathbb{R}^N \to C^\alpha(\mathbb{R}^n)$, where α denotes the smoothness of ϕ_s. In practice, common choices of kernel include the Gaussian $\phi_s(r) = \exp(-(sr)^2)$ (Fasshauer and McCourt, 2012), inverse quadratic function $\phi_s(r) = (1 + (sr)^2)^{-1}$, or Matérn class kernel (Fuselier and Wright, 2013). In our numerical examples, we have tested these kernels and they do not make too much difference when the shape parameters are tuned properly. However, we will develop the theoretical analysis with the Matérn kernel in Section 4.1 as it induces a reproducing kernel Hilbert space (RKHS) with Sobolev-like norm.

The expansion coefficients $\{c_k\}_{k=1}^N$ in (2) can be determined by a collocation method, which enforces the interpolation condition $I_{\phi_s}\mathbf{f}(x_j) = f(x_j)$ for all $j = 1, \ldots, N$, or the following linear system with the interpolation matrix $\mathbf{\Phi}$:

$$\underbrace{\begin{bmatrix} \phi_s(\|x_1 - x_1\|) & \phi_s(\|x_1 - x_2\|) & \cdots & \phi_s(\|x_1 - x_N\|) \\ \phi_s(\|x_2 - x_1\|) & \phi_s(\|x_2 - x_2\|) & \cdots & \phi_s(\|x_2 - x_N\|) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_s(\|x_N - x_1\|) & \phi_s(\|x_N - x_2\|) & \cdots & \phi_s(\|x_N - x_N\|) \end{bmatrix}}_{\mathbf{\Phi}} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}}_{\mathbf{c}} = \underbrace{\begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_N) \end{bmatrix}}_{\mathbf{f}}. \tag{3}$$

In general, better accuracy is obtained for flat kernels (small s) [see e.g., Chap. 16–17 of (Fasshauer, 2007)], however, the corresponding system in (3) becomes increasingly ill-conditioned (Fornberg and Wright, 2004; Fornberg et al., 2011; Fasshauer and McCourt, 2012). In this article, we will not focus on the shape parameter issues. Numerically, we will empirically choose the shape parameter and will use pseudo-inverse to solve the linear system in (3) when it is effectively singular.

With these backgrounds, we are now ready to discuss the RBF approximation to various differential operators.

2.1 Gradient of a Function

We first review the RBF projection method proposed since Kansa (1990)'s pioneering work and following works in Fuselier and Wright (2013); Flyer and Bengt (2015); Shankar et al. (2015); Lehto et al. (2017) for approximating gradients of functions on manifolds. The projection method represents the manifold differential operators as tangential gradients, which are formulated as the projection of the appropriate derivatives in the ambient space.

Precisely, the manifold gradient on a smooth function $f : M \to \mathbb{R}$ evaluated at $x \in M$ in the Cartesian coordinates is given as,

$$\text{grad}_g f(x) := \mathbf{P}\overline{\text{grad}}_{\mathbb{R}^n} f(x) = (\sum_{i=1}^{d} \boldsymbol{\tau}_i \boldsymbol{\tau}_i^\top)\overline{\text{grad}}_{\mathbb{R}^n} f(x),$$

where the subscript g is to associate the differential operator to the Riemannian metric g induced by M and $\overline{\text{grad}}_{\mathbb{R}^n} = [\partial_{X^1}, \cdots, \partial_{X^n}]^\top$ is the usual Euclidean gradient operator in $\mathbb{R}^n$. Let $\mathbf{e}_\ell, \ell = 1, ..., n$ be the standard orthonormal vectors in X^ℓ direction in $\mathbb{R}^n$, we can rewrite above expression in component form as

$$\text{grad}_g f(x) := \begin{bmatrix} (\mathbf{e}_1 \cdot \mathbf{P})\,\overline{\text{grad}}_{\mathbb{R}^n} f(x) \\ \vdots \\ (\mathbf{e}_n \cdot \mathbf{P})\,\overline{\text{grad}}_{\mathbb{R}^n} f(x) \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1 \cdot \overline{\text{grad}}_{\mathbb{R}^n} f(x) \\ \vdots \\ \mathbf{p}_n \cdot \overline{\text{grad}}_{\mathbb{R}^n} f(x) \end{bmatrix} := \begin{bmatrix} \mathcal{G}_1 f(x) \\ \vdots \\ \mathcal{G}_n f(x) \end{bmatrix} \tag{4}$$

where $\mathbf{p}_\ell$ is the $\ell-$th column of the projection matrix $\mathbf{P}$.

One can now consider estimating the gradient operator from the available training data set $X = \{x_1, \ldots, x_N\}$. In such a case, one considers the RBF interpolant $I_{\phi_s}\mathbf{f} \in C^\alpha(\mathbb{R}^n)$ in (2) which interpolates f on the available training data set $X = \{x_1, \ldots, x_N\}$ by solving (3). Using the interpolant (2) and denoting $r_k := \|x - x_k\|$, one can evaluate the tangential derivative in the X^i direction at each node $\{x_j \in X\}_{j=1}^N$ as,

$$\begin{aligned} \mathcal{G}_i I_{\phi_s}\mathbf{f}(x)|_{x=x_j} &= \mathbf{p}_i^\top \cdot \overline{\text{grad}}_{\mathbb{R}^n} I_{\phi_s}\mathbf{f}(x)|_{x=x_j} = \mathbf{p}_i^\top \cdot \overline{\text{grad}}_{\mathbb{R}^n} \sum_{k=1}^{N} c_k \phi_s(r_k(x))|_{x=x_j} \\ &= \sum_{k=1}^{N} c_k \mathbf{p}_i^\top \cdot \overline{\text{grad}}_{\mathbb{R}^n} \phi_s(r_k(x))|_{x=x_j} \\ &= \sum_{k=1}^{N} c_k \mathbf{p}_i^\top \cdot (x - x_k)\,\frac{\phi_s'(r_k(x))}{r_k(x)}|_{x=x_j} := \sum_{k=1}^{N} J_{jk}^{[i]} c_k. \end{aligned}$$

Let $\mathbf{J}_i = [J_{ik}^{[i]}]_{j,k=1}^N$, and also let $\mathbf{c} = (c_1, \ldots, c_N)^\top$ and $\mathbf{f} = f(x)|_X = (f(x_1), \ldots, f(x_N))$ as defined in (3). Above equation can be written in matrix form for $\mathcal{G}_i I_{\phi_s} f$ at all nodes X,

$$(\mathcal{G}_i I_{\phi_s}\mathbf{f})\,|_X = \mathbf{J}_i\mathbf{c} = \mathbf{J}_i\boldsymbol{\Phi}^{-1}\mathbf{f} := \mathbf{G}_i\mathbf{f}, \tag{5}$$

for $i = 1, \ldots, n$. Thereafter, we define

Definition 3 *Let $\mathcal{G}_i I_{\phi_s} : \mathbb{R}^N \to C^\alpha(\mathbb{R}^n)$ and $R_N : C(\mathbb{R}^n) \to \mathbb{R}^N$ be the restriction operator defined as $R_N f = \mathbf{f} = (f(x_1), \ldots, f(x_N))^\top$ for any $f \in C(M)$ and $I_{\phi_s}\mathbf{f} \in C^\alpha(\mathbb{R}^n)$ be the RBF interpolant as defined in (2). We define a linear map $\mathbf{G}_i : \mathbb{R}^N \to \mathbb{R}^N$*

$$\mathbf{G}_i\mathbf{f} = R_N \mathcal{G}_i I_{\phi_s}\mathbf{f}. \tag{6}$$

as a discrete estimator of the differential operator $\mathcal{G}_i$, restricted on the training data X. On the right-hand-side, we understood $R_N\mathcal{G}_i I_{\phi_s}\mathbf{f} = \mathcal{G}_i I_{\phi_s}\mathbf{f}|_X \in \mathbb{R}^N$.

To conclude, we define a linear map $\mathbf{G} : \mathbb{R}^N \to \mathbb{R}^{nN}$ with,

$$\mathbf{Gf} = (\mathbf{G}_1\mathbf{f}, \mathbf{G}_2\mathbf{f}, \ldots, \mathbf{G}_n\mathbf{f})^\top, \tag{7}$$

as a discrete estimator to the gradient restricted on the training data set X, in the sense that

$$\mathbf{Gf} = \mathbf{G}R_N f = R_N\left(\mathrm{grad}_g I_{\phi_s}\mathbf{f}\right) = (\mathrm{grad}_g I_{\phi_s}\mathbf{f})|_X, \tag{8}$$

where the first and third equalities follow from the definition of R_N, and the second equality follows from (4), (6), (7).

2.2 The Laplace-Beltrami Operator

The Laplace-Beltrami operator on a smooth function $f \in C^\infty(M)$ is defined as $\Delta_M f = -\mathrm{div}_g(\mathrm{grad}_g f)$ which is semi-positive definite. Using the previous ambient space formulation for gradient, one can equivalently write

$$\begin{aligned}\Delta_M f(x) &:= -\mathrm{div}_g(\mathrm{grad}_g f(x)) = -(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}) \cdot \left(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}\right) f(x) \\ &= -(\mathcal{G}_1\mathcal{G}_1 + \mathcal{G}_2\mathcal{G}_2 + \cdots + \mathcal{G}_n\mathcal{G}_n) f(x),\end{aligned}$$

for any $x \in M$. This identity yields a pointwise estimate of Δ_M by composing the discrete estimators for div_g and grad_g. Particularly, a **non-symmetric estimator** of the Laplace-Beltrami operator is a map $\mathbb{R}^N \to \mathbb{R}^N$ given by

$$\mathbf{f} \mapsto -\left(\mathbf{G}_1\mathbf{G}_1 + \cdots + \mathbf{G}_n\mathbf{G}_n\right)\mathbf{f}. \tag{9}$$

Remark 4 *The above discrete version of Laplace-Beltrami is not new. In particular, it has been well studied in the deterministic setting. See Dziuk and Elliott (2007); Wardetzky (2007); Dziuk and Elliott (2013) for a detailed review of using the ambient space for estimation of Laplace-Beltrami, as well as Fuselier and Wright (2013). In Section 5, we will numerically study the spectral convergence of this non-symmetric discretization in the setting where the data are sampled randomly from an unknown manifold. We will remark on the advantages and disadvantages of such a non-symmetric formulation for manifold learning tasks.*

In the weak form, for M without boundary, the Laplace-Beltrami operator can be written

$$\int_M h\Delta_M f d\mathrm{Vol} = \int_M \langle \mathrm{grad}_g h, \mathrm{grad}_g f\rangle d\mathrm{Vol}, \quad \forall f, h \in C^\infty(M),$$

where $\langle\cdot,\cdot\rangle$ denotes the Riemannian inner product of vector fields. Using the estimators from previous subsections, it is natural to estimate the Laplace-Beltrami operator in this setting. Based on the weak formulation, we can estimate the gradient with the matrix $\mathbf{G} : \mathbb{R}^N \to \mathbb{R}^{Nn}$, then compose with the matrix adjoint of $\mathbf{G}$, where the domain and range of $\mathbf{G}$ are equipped with appropriate inner products approximating the corresponding inner products defined on the manifold. It turns out that the adjoint of $\mathbf{G}$ following this procedure

is just the standard matrix transpose. In particular, we have that $\mathbf{G}^\top\mathbf{G}:\mathbb{R}^N\to\mathbb{R}^N$ given by

$$\mathbf{G}^\top\mathbf{G}\,\mathbf{f}:=\left(\mathbf{G}_1^\top\mathbf{G}_1+\mathbf{G}_2^\top\mathbf{G}_2+\ldots+\mathbf{G}_n^\top\mathbf{G}_n\right)\mathbf{f}$$

is a **symmetric estimator** of the Laplace-Beltrami operator.

Remark 5 *The above symmetric formulation makes use of the discrete approximation of continuous inner products, and only holds when the data is sampled uniformly. For data sampled from a non-uniform density q, however, we can perform the standard technique of correcting for non-uniform data by dividing by the sampling density. For example, the symmetric approximation to the eigenvalue problem corresponds to solving $(\lambda,\mathbf{f})$ that satisfies*

$$\mathbf{f}^\top\mathbf{G}^\top\mathbf{Q}^{-1}\mathbf{G}\,\mathbf{f}:=\mathbf{f}^\top\left(\mathbf{G}_1^\top\mathbf{Q}^{-1}\mathbf{G}_1+\mathbf{G}_2^\top\mathbf{Q}^{-1}\mathbf{G}_2+\ldots+\mathbf{G}_n^\top\mathbf{Q}^{-1}\mathbf{G}_n\right)\mathbf{f}=\lambda\mathbf{f}^\top\mathbf{Q}^{-1}\mathbf{f},\quad(10)$$

where $\mathbf{Q}\in\mathbb{R}^{N\times N}$ is a diagonal matrix with diagonal entries of sampling density $\{q(x_i)\}_{i=1,\ldots,N}$. This weighted Monte-Carlo provides an estimate for the $L^2(M)$ inner product in the weak formulation. When q is unknown, one can approximate q using standard density estimation techniques, such as Kernel Density Estimation methods (Parzen, 1962). In our numerical simulations, we use the MATLAB built-in function `mvksdensity.m`.

2.3 Covariant Derivative

The basic idea here follows from the tangential connection on a submanifold of $\mathbb{R}^n$ in Example 4.9 of Lee (2018). For smooth vector fields $u,y\in\mathfrak{X}(M)$, the tangential connection can be defined as

$$\nabla_u y=\mathcal{P}\left(\bar{\nabla}_U Y|_M\right),\quad(11)$$

where ∇ is the Levi-Civita connection on M, $\bar{\nabla}:\mathfrak{X}(\mathbb{R}^n)\times\mathfrak{X}(\mathbb{R}^n)\to\mathfrak{X}(\mathbb{R}^n)$ is the Euclidean connection on $\mathbb{R}^n$ mapping (U,Y) to $\bar{\nabla}_U Y$ (Example 4.8 of Lee (2018)), $\mathcal{P}:T\mathbb{R}^n\to TM$ is the orthogonal projection onto TM, and U and Y are smooth extensions of u and y to an open subset in $\mathbb{R}^n$ satisfying $U|_M=u$ and $Y|_M=y$, respectively. Such extensions exist by Exercise A.23 and the identity result does not depend on the chosen extension by Proposition 4.5 in Lee (2018). The identity (11) holds true based on the properties and uniqueness of Levi-Civita connection (see Lee (2018)). More geometric intuition and detailed results can also be found in (Do Carmo and Flaherty Francis, 1992; Morita, 2001; Crane et al., 2010; Azencot et al., 2015) and references therein. The key observation for identity (11) is that the covariant derivative can be written in terms of the tangential projection and Euclidean derivative. In the remainder of this section, we review the covariant derivative from a geometric viewpoint and then formulate the tangential projection identity (11) from a computational viewpoint.

Let $u\in\mathfrak{X}(M)$ be a vector field on M, and let $U\in\mathfrak{X}(\mathbb{R}^n)$ be an extension of u to an open set $O\subseteq M$. Then U is related to u via the local parameterization as follows: $U(x)=D\iota(x)\,u(x)$, where $u(x)=(u^1(x),\ldots,u^d(x))$ is the coordinate representation of the vector field u w.r.t. the basis $\left\{\frac{\partial}{\partial\theta^r}\right\}$. Using $[D\iota(x)]_{sr}=\frac{\partial X^s}{\partial\theta^r}$, we have the following equation relating the components of the vector fields:

$$U^s=u^r\frac{\partial X^s}{\partial\theta^r}.\quad(12)$$

Using this identity, we can derive

Proposition 2.2 *Let $U = U^i \frac{\partial}{\partial X^i} \in \mathfrak{X}(\mathbb{R}^n)$ be a vector field such that $U|_M = u \in \mathfrak{X}(M)$. Using the notation in Definition 1, we have*

$$\sum_r g^{ij} \frac{\partial X^r}{\partial \theta^j} \frac{\partial U^r}{\partial \theta^k} = \frac{\partial u^i}{\partial \theta^k} + u^p \Gamma^i_{pk}. \tag{13}$$

The proof is relegated in appendix A. As in the previous section, the projection matrix $\mathbf{P}$ in Definition 2 has been used for approximating operators acting on functions as matrix-vector multiplication. In the following, we first introduce a tangential projection tensor in order to derive identities for operators acting on tensor fields with extensions. Geometrically, the tangential projection tensor projects a vector field of $\mathfrak{X}(\mathbb{R}^n)$ onto a vector field of $\mathfrak{X}(M)$.

Definition 6 *The tangential projection tensor $\mathcal{P} : \mathfrak{X}(\mathbb{R}^n) \to \mathfrak{X}(M)$ is defined as*

$$\mathcal{P} = \delta_{sr} \frac{\partial X^s}{\partial \theta^i} g^{ij} \frac{\partial X^t}{\partial \theta^j} \mathrm{d}X^r \otimes \frac{\partial}{\partial X^t}, \tag{14}$$

where δ_{sr} is the Kronecker delta function. In particular, for a vector field $Y = Y^k \frac{\partial}{\partial X^k} \in \mathfrak{X}(\mathbb{R}^n)$, one has

$$\begin{aligned} \mathcal{P}\left(Y|_M\right) &= \delta_{sr} Y^k \frac{\partial X^s}{\partial \theta^i} g^{ij} \frac{\partial X^t}{\partial \theta^j} \mathrm{d}X^r \left(\frac{\partial}{\partial X^k}\right) \otimes \frac{\partial}{\partial X^t} = \delta_{sr} Y^r \frac{\partial X^s}{\partial \theta^i} g^{ij} \frac{\partial X^t}{\partial \theta^j} \frac{\partial}{\partial X^t} \\ &= \delta_{sr} Y^r \frac{\partial X^s}{\partial \theta^i} g^{ij} \frac{\partial}{\partial \theta^j} \in \mathfrak{X}(M). \end{aligned}$$

For convenience, we simplify our notation as $\mathcal{P}Y := \mathcal{P}\left(Y|_M\right)$ in the rest of the paper since we only concern about the points restricted on manifold M. Obviously, for any vector field $v \in \mathfrak{X}(M)$, we have $\mathcal{P}v = v \in \mathfrak{X}(M)$.

Using Definition in (14) and Proposition in (13), we can examine the identity $\nabla_u y = \mathcal{P}\bar{\nabla}_U Y$ via a direct calculation (see appendix A).

2.4 Gradient of a Vector Field

There are several frameworks for the discretization of vector Laplacians on manifolds such as Discrete exterior calculus (DEC) (Hirani, 2003), finite element exterior calculus (FEC) (Arnold et al., 2006, 2010; Gillette et al., 2017), Generalized Moving Least Squares (GMLS) (Gross et al., 2020) and spectral exterior calculus (SEC) (Berry, 2018; Berry and Giannakis, 2020). All these methods provide pointwise consistent discrete estimates for vector field operators such as curl, gradient, and Hodge Laplacian. Both DEC and FEC make strong use of a simplicial complex in their formulation which helps to achieve high-order accuracy in PDE problems but is not realistic for many data science applications. GMLS is a mesh-free method that is applied to solving vector PDEs on manifolds in (Gross et al., 2020). SEC can be used for processing raw data as a mesh-free tool which is appropriate for manifold learning applications. Here, we will only focus on Bochner Laplacian and derive the tangential projection identities for various vector field differential operators in terms of

the tangential projection and Euclidean derivative acting on vector fields with extensions. The formulation for the Hodge and Lichnerowicz Laplacians is similar (see Appendix A).

The gradient of a vector field $u \in \mathfrak{X}(M)$ is defined by $\mathrm{grad}_g u = \sharp \nabla u$, where $\sharp$ is the standard musical isomorphism notation to raise the index, in this case from a $(1,1)$ tensor to a $(2,0)$ tensor. In local intrinsic coordinates, one can calculate

$$\mathrm{grad}_g u = g^{kj}\left(\frac{\partial u^i}{\partial \theta^k} + u^p \Gamma^i_{pk}\right)\frac{\partial}{\partial \theta^i}\otimes\frac{\partial}{\partial \theta^j}.$$

Using (13), we can rewrite the above as

$$\begin{aligned}
\mathrm{grad}_g u &= g^{km}\left(\delta_{rs} g^{ij}\frac{\partial X^r}{\partial \theta^j}\frac{\partial U^s}{\partial \theta^k}\right)\frac{\partial}{\partial \theta^i}\otimes\frac{\partial}{\partial \theta^m}, \\
&= g^{km}\delta_{rs} g^{ij}\frac{\partial X^r}{\partial \theta^j}\left(\frac{\partial U^s}{\partial X^a}\frac{\partial X^a}{\partial \theta^k}\right)\left(\frac{\partial X^b}{\partial \theta^i}\frac{\partial}{\partial X^b}\right)\otimes\left(\frac{\partial}{\partial X^c}\frac{\partial X^c}{\partial \theta^m}\right), \\
&= \delta^{ea}\left(\delta_{ep}\frac{\partial X^c}{\partial \theta^k} g^{km}\frac{\partial X^p}{\partial \theta^m}\right)\left(\delta_{rs}\frac{\partial X^b}{\partial \theta^i} g^{ij}\frac{\partial X^r}{\partial \theta^j}\right)\frac{\partial U^s}{\partial X^a}\frac{\partial}{\partial X^b}\otimes\frac{\partial}{\partial X^c}, \\
&= \mathcal{P}_1\mathcal{P}_2\left(\delta^{ea}\frac{\partial U^r}{\partial X^a}\frac{\partial}{\partial X^r}\otimes\frac{\partial}{\partial X^e}\right) \equiv \mathcal{P}_1\mathcal{P}_2\overline{\mathrm{grad}}_{\mathbb{R}^n} U, \qquad (15)
\end{aligned}$$

where $U = U^s\frac{\partial}{\partial X^s} = u^p\frac{\partial X^s}{\partial \theta^p}\frac{\partial}{\partial X^s}$, and $\mathcal{P}_1 = \delta_{ts}\frac{\partial X^t}{\partial \theta^k} g^{km}\frac{\partial X^b}{\partial \theta^m}\mathrm{d}X^s\otimes\frac{\partial}{\partial X^b}$ acting on the first tensor component $\frac{\partial}{\partial X^r}$, and $\mathcal{P}_2 = \delta_{rq}\frac{\partial X^c}{\partial \theta^i} g^{ij}\frac{\partial X^r}{\partial \theta^j}\mathrm{d}X^q\otimes\frac{\partial}{\partial X^c}$ acting on the second component $\frac{\partial}{\partial X^e}$. Evaluating at each $x \in M$ and using the notation in (1), Eq. (15) can be written in a matrix form as,

$$\mathrm{grad}_g u(x) = \mathbf{P}\left(\overline{\mathrm{grad}}_{\mathbb{R}^n} U(x)\right)\mathbf{P} = \begin{bmatrix} \mathcal{G}_1 U^1(x) & \cdots & \mathcal{G}_1 U^n(x) \\ \vdots & \ddots & \vdots \\ \mathcal{G}_n U^1(x) & \cdots & \mathcal{G}_n U^n(x) \end{bmatrix}\begin{bmatrix} P_{11}(x) & \cdots & P_{1n}(x) \\ \vdots & \ddots & \vdots \\ P_{n1}(x) & \cdots & P_{nn}(x) \end{bmatrix}.$$

Interpreting $U(x) = (U^1(x), \ldots, U^n(x))^\top \in \mathbb{R}^{n\times 1}$ as a vector, one sees immediately by taking the transpose of the above formula that

$$\mathrm{grad}_g u(x) = \left(\mathcal{H}_1 U(x), \ldots, \mathcal{H}_n U(x)\right), \qquad (16)$$

where each component can be rewritten as,

$$\mathcal{H}_i U(x) := \mathbf{P}\mathrm{diag}\left(\mathcal{G}_i, \ldots, \mathcal{G}_i\right) U(x),$$

owing to the symmetry of $\mathbf{P}$.

To write the discrete approximation on the training data set $X = \{x_1, \ldots, x_n\}$, we define $\mathbf{U}^i = (U^i(x_1), \ldots, U^i(x_N))^\top \in \mathbb{R}^{N\times 1}$ and concatenate these $\mathbf{U}^i$ to form $\mathbf{U} = ((\mathbf{U}^1)^\top, \ldots, (\mathbf{U}^n)^\top)^\top \in \mathbb{R}^{nN\times 1}$. Consider now the map $\mathbf{H}_i : \mathbb{R}^{nN} \to \mathbb{R}^{nN}$ defined by

$$\mathbf{H}_i\mathbf{U} := R_N \mathcal{H}_i I_{\phi_s}\mathbf{U}, \qquad (17)$$

where the interpolation is defined on each $\mathbf{U}^i \in \mathbb{R}^N$ such that $I_{\phi_s}\mathbf{U}^i \in C^\alpha(\mathbb{R}^n)$ and the restriction R_N is applied on each row. Relating to $\mathbf{G}_i$ in Definition 3, one can write

$$\mathbf{H}_i\mathbf{U} = \mathbf{P}^{\otimes}\begin{bmatrix} \mathbf{G}_i & & \\ & \ddots & \\ & & \mathbf{G}_i \end{bmatrix}_{Nn\times Nn}\begin{bmatrix} \mathbf{U}^1 \\ \vdots \\ \mathbf{U}^n \end{bmatrix}_{Nn\times 1},$$

where tensor projection matrix $\mathbf{P}^{\otimes} \in \mathbb{R}^{Nn\times Nn}$ is given by

$$\mathbf{P}^{\otimes} = \sum_{k=1}^{N} \begin{bmatrix} P_{11}(x_k) & \cdots & P_{1n}(x_k) \\ \vdots & \ddots & \vdots \\ P_{n1}(x_k) & \cdots & P_{nn}(x_k) \end{bmatrix}_{n\times n} \otimes [\delta_{kk}]_{N\times N} = \begin{bmatrix} \mathrm{diag}(\mathbf{p}_{11}) & \cdots & \mathrm{diag}(\mathbf{p}_{1n}) \\ \vdots & \ddots & \vdots \\ \mathrm{diag}(\mathbf{p}_{n1}) & \cdots & \mathrm{diag}(\mathbf{p}_{nn}) \end{bmatrix}_{Nn\times Nn},$$

where $\otimes$ is the Kronecker product between two matrices, δ_{kk} has value 1 for the entry in kth row and kth column and has 0 values elsewhere, and $\mathbf{p}_{ij} = (P_{ij}(x_1), \ldots, P_{ij}(x_N)) \in \mathbb{R}^{N\times 1}$. Finally, consider the linear map $\mathbf{H} : \mathbb{R}^{nN} \to \mathbb{R}^{nN\times n}$ given by

$$\mathbf{HU} := [\mathbf{H}_1\mathbf{U}, \ldots, \mathbf{H}_n\mathbf{U}] = R_N \mathrm{grad}_g I_{\phi_s}\mathbf{U}, \tag{18}$$

as an estimator of the gradient of any vector field U restricted on the training data set X, where on the right-hand-side, the restriction is done to each function entry-wise which results in an $N \times 1$ column vector. In the last equality, we have used the identity in (17) and the representation in (16) for the gradient of the interpolated vector field $I_{\phi_s}\mathbf{U}$ whose components are functions in $C^{\alpha}(\mathbb{R}^n)$.

2.5 Divergence of a (2,0) Tensor Field

Let v be a $(2,0)$ tensor field of $v = v^{jk}\frac{\partial}{\partial\theta^j} \otimes \frac{\partial}{\partial\theta^k}$ and V be the corresponding extension in ambient space. The divergence of v is defined as

$$\mathrm{div}_1^1(v) = C_1^1(\nabla v), \tag{19}$$

where C_1^1 denotes the contraction operator. Following a similar layout as before, we obtain an ambient space formulation of the divergence of a $(2,0)$ tensor field (see Appendix A for the detailed derivations). Interpreting $V(x)$ as an $n \times n$ matrix with columns $\bar{V}_i(x) = (V_{i1}, \ldots, V_{in})^{\top}$, the divergence of a $(2,0)$ tensor field evaluated at any $x \in M$ can be written as

$$\mathrm{div}_1^1(v(x)) = \mathbf{P}\mathrm{tr}_1^1\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}(V(x))\right) = \sum_i \mathbf{P}\mathrm{diag}(\mathcal{G}_i, \ldots, \mathcal{G}_i)\bar{V}_i(x) = \sum_i \mathcal{H}_i\bar{V}_i(x). \tag{20}$$

Using the same procedure as before, employing div_1^1 on the RBF interpolant (2,0) tensor field $I_{\phi_s}\bar{\mathbf{V}}_i$, where $\bar{\mathbf{V}}_i = \bar{V}_i|_X \in \mathbb{R}^{nN\times 1}$ denotes the restriction of $\bar{V}_i$ on the training data, we arrive at an estimator of the divergence div_1^1 of a $(2,0)$ tensor field. Namely, replacing each $\mathcal{H}_i$ with the discrete version $\mathbf{H}_i$ as defined in (17), we obtain a map $\mathbb{R}^{nN\times n} \to \mathbb{R}^{nN}$ given by

$$[\bar{\mathbf{V}}_1, \ldots, \bar{\mathbf{V}}_n] \mapsto \mathbf{H}_1\bar{\mathbf{V}}_1 + \cdots + \mathbf{H}_n\bar{\mathbf{V}}_n.$$

2.6 Bochner Laplacian

The Bochner Laplacian $\Delta_B : \mathfrak{X}(M) \to \mathfrak{X}(M)$ is defined by

$$\Delta_B u = -\mathrm{div}_1^1\left(\mathrm{grad}_g u\right),$$

where div_1^1 is in fact the formal adjoint of grad_g acting on vector fields. With an extension to Euclidean space, the Bochner Laplacian can be formulated as:

$$\bar{\Delta}_B U = -(\mathcal{H}_1\mathcal{H}_1 + \cdots + \mathcal{H}_n\mathcal{H}_n)U,$$

where (16) and (20) have been used. A natural way to estimate Δ_B is to compose the discrete estimators for div_1^1 and grad_g. In particular, a **non-symmetric estimator** of the Bochner Laplacian is a map $\mathbb{R}^{nN} \to \mathbb{R}^{nN}$ given by

$$\mathbf{U} \mapsto -(\mathbf{H}_1\mathbf{H}_1 + \cdots + \mathbf{H}_n\mathbf{H}_n)\mathbf{U},$$

where $\mathbf{U} = U|_X \in \mathbb{R}^{nN\times 1}$ and $\mathbf{H}_i$ are as defined in (17).

The second formulation relies on the fact that div_1^1 is indeed the formal adjoint of grad_g acting on a vector field such that,

$$\int_M \langle \Delta_B u, v\rangle_x d\text{Vol}(x) = \int_M \langle \text{grad}_g u, \text{grad}_g v\rangle_x d\text{Vol}(x), \qquad \forall u, v \in \mathfrak{X}(M),$$

where the inner product on the left is the Riemannian inner product of vector fields, and the inner product on the right is the Riemannian inner product of $(2,0)$ tensor fields. Similar to the symmetric discrete estimator of the Laplace-Beltrami operator, we take advantages of the ambient space formulations in previous two subsections to approximate the inner product with appropriate normalized inner products in Euclidean space.

First, we notice that the transpose of the map $\mathbf{H} : \mathbf{U} \mapsto \big[\mathbf{H}_1\mathbf{U}, \mathbf{H}_2\mathbf{U}, \ldots, \mathbf{H}_n\mathbf{U}\big]$ is given by the standard transpose. Due to the possibility, however, that $\mathbf{H}^\top$ may produce a vector corresponding to a vector field with components normal to the manifold which are nonzero, there is a need to compose such an estimator with the projection matrix. With this consideration, we have $\mathbf{P}^\otimes\mathbf{H}^\top\mathbf{H}\mathbf{P}^\otimes : \mathbb{R}^{nN} \to \mathbb{R}^{nN}$ given by

$$\mathbf{P}^\otimes\mathbf{H}^\top\mathbf{H}\mathbf{P}^\otimes\mathbf{U} = \mathbf{P}^\otimes\mathbf{H}_1^\top\mathbf{H}_1\mathbf{P}^\otimes\mathbf{U} + \mathbf{P}^\otimes\mathbf{H}_2^\top\mathbf{H}_2\mathbf{P}^\otimes\mathbf{U} + \cdots + \mathbf{P}^\otimes\mathbf{H}_n^\top\mathbf{H}_n\mathbf{P}^\otimes\mathbf{U}$$

as a **symmetric estimator** of the Bochner Laplacian on vector fields.

Remark 7 *Again, we note that this symmetric formulation makes use of approximating continuous inner products, and hence obviously holds only for uniform data. For data sampled from a non-uniform density q, we perform the same trick mentioned in Remark 5.*

We conclude this section with a list of RBF discrete formulation in Table 2. One can see the detailed derivation for the non-symmetric approximations of Hodge and Lichnerowicz Laplacians in Appendix A. We neglect the derivations for the symmetric approximations of the Hodge and Lichnerowicz Laplacians as they are analogous to that for the Bochner Laplacian but involve more terms.

2.7 Numerical Verification for Operator Approximation

We now show the non-symmetric RBF (NRBF) estimates for vector Laplacians and covariant derivative. The manifold is a one-dimensional full ellipse,

$$x = (x^1, x^2) = (\cos\theta, a\sin\theta), \tag{21}$$

defined with the Riemannian metric $g = \sin^2\theta + a^2\cos^2\theta$ for $0 \le \theta < 2\pi$,where $a = 2 > 1$. The $N = 400$ data points are randomly distributed on the ellipse. The Gaussian kernel with the shape parameter $s = 1.5$ was used.

Table 2: RBF formulation for functions and vector fields from Riemannian geometry. Here, *non-symmetric* and *symmetric* correspond to the non-symmetric and symmetric approximations to the differential operator. The asterisk * is the formal adjoint of the differential operator. $\mathcal{S}_i$ and $\mathbf{S}_i$ are defined around (51) in Appendix A.

Object	Continuous operator	Discrete matrix
gradient	$\operatorname{grad}_g : C^\infty(M) \to \mathfrak{X}(M)$	$\operatorname{grad}_g : \mathbb{R}^N \to \mathbb{R}^{N\times n}$
functions	$f \mapsto \ [\mathcal{G}_1 f, \dots, \mathcal{G}_n f]$	$\mathbf{f} \mapsto (\mathbf{G}_1\mathbf{f}, \dots, \mathbf{G}_n\mathbf{f})$
divergence	$\operatorname{div}_g : \mathfrak{X}(M) \to C^\infty(M)$	$\operatorname{div}_g : \mathbb{R}^{N\times n} \to \mathbb{R}^{N\times 1}$
vector fields	$U \mapsto \mathcal{G}_1 U^1 + \cdots + \mathcal{G}_n U^n$	$\mathbf{U} \mapsto \mathbf{G}_1\mathbf{U}^1 + \cdots + \mathbf{G}_n\mathbf{U}^n$
Laplace-Beltrami	$\Delta_g : C^\infty(M) \to C^\infty(M)$	$\Delta_g : \mathbb{R}^{N\times 1} \to \mathbb{R}^{N\times 1}$
non-symmetric	$f \mapsto -(\mathcal{G}_1\mathcal{G}_1 + \cdots + \mathcal{G}_n\mathcal{G}_n) f$	$\mathbf{f} \mapsto -(\mathbf{G}_1\mathbf{G}_1 + \cdots + \mathbf{G}_n\mathbf{G}_n)\mathbf{f}$
symmetric	$f \mapsto (\mathcal{G}_1^*\mathcal{G}_1 + \cdots + \mathcal{G}_n^*\mathcal{G}_n) f$	$\mathbf{f} \mapsto (\mathbf{G}_1^\top\mathbf{G}_1 + \cdots + \mathbf{G}_n^\top\mathbf{G}_n)\mathbf{f}$
gradient	$\operatorname{grad}_g : \mathfrak{X}(M) \to \mathfrak{X}(M)\times\mathfrak{X}(M)$	$\operatorname{grad}_g : \mathbb{R}^{Nn\times 1} \to \mathbb{R}^{Nn\times n}$
vector fields	$U \mapsto [\mathcal{H}_1 U, \dots, \mathcal{H}_n U]$	$\mathbf{U} \mapsto [\mathbf{H}_1\mathbf{U}, \dots, \mathbf{H}_n\mathbf{U}]$
divergence	$\operatorname{div}_1^1 : \mathfrak{X}(M)\times\mathfrak{X}(M) \to \mathfrak{X}(M)$	$\operatorname{div}_1^1 : \mathbb{R}^{Nn\times n} \to \mathbb{R}^{Nn\times 1}$
(2,0) tensor fields	$V \mapsto \mathcal{H}_1\bar{V}_1 + \cdots + \mathcal{H}_n\bar{V}_n$	$[\bar{\mathbf{V}}_1, \dots, \bar{\mathbf{V}}_n] \mapsto \mathbf{H}_1\bar{\mathbf{V}}_1 + \cdots + \mathbf{H}_n\bar{\mathbf{V}}_n$
	$\bar{V}_i$ is the ith row of V	$\bar{\mathbf{V}}_i = [\mathbf{V}_{i1}, \dots, \mathbf{V}_{in}] \in \mathbb{R}^{Nn\times 1}$
Bochner Laplacian	$\Delta_B : \mathfrak{X}(M) \to \mathfrak{X}(M)$	$\Delta_B : \mathbb{R}^{Nn\times 1} \to \mathbb{R}^{Nn\times 1}$
non-symmetric	$U \mapsto -(\mathcal{H}_1\mathcal{H}_1 + \cdots + \mathcal{H}_n\mathcal{H}_n)U$	$\mathbf{U} \mapsto -(\mathbf{H}_1\mathbf{H}_1 + \cdots + \mathbf{H}_n\mathbf{H}_n)\mathbf{U}$
symmetric	$U \mapsto (\mathcal{H}_1^*\mathcal{H}_1 + \cdots + \mathcal{H}_n^*\mathcal{H}_n)U$	$\mathbf{U} \mapsto (\mathbf{P}^\otimes\mathbf{H}_1^\top\mathbf{H}_1\mathbf{P}^\otimes + \cdots + \mathbf{P}^\otimes\mathbf{H}_n^\top\mathbf{H}_n\mathbf{P}^\otimes)\mathbf{U}$
Hodge Laplacian	$\Delta_H : \mathfrak{X}(M) \to \mathfrak{X}(M)$	$\Delta_H : \mathbb{R}^{Nn\times 1} \to \mathbb{R}^{Nn\times 1}$
vector fields	$U \mapsto -\begin{bmatrix}\mathcal{H}_1\\ \vdots\\ \mathcal{H}_n\end{bmatrix} \cdot \operatorname{Ant} \begin{bmatrix}\mathcal{H}_1 U\\ \vdots\\ \mathcal{H}_n U\end{bmatrix}$	$\mathbf{U} \mapsto -\begin{bmatrix}\mathbf{H}_1\\ \vdots\\ \mathbf{H}_n\end{bmatrix} \cdot \operatorname{Ant} \begin{bmatrix}\mathbf{H}_1\mathbf{U}\\ \vdots\\ \mathbf{H}_n\mathbf{U}\end{bmatrix}$
Ant is the anti-symmetric part	$-\begin{bmatrix}\mathcal{G}_1\\ \vdots\\ \mathcal{G}_n\end{bmatrix} \left(\sum_{k=1}^n \mathcal{G}_k U^k\right)$	$-\begin{bmatrix}\mathbf{G}_1\\ \vdots\\ \mathbf{G}_n\end{bmatrix} \left(\sum_{k=1}^n \mathbf{G}_k\mathbf{U}^k\right)$
non-symmetric	$= -\sum_{i=1}^n \mathcal{H}_i(\mathcal{H}_i - \mathcal{S}_i)U$	$= -\sum_{i=1}^n \mathbf{H}_i(\mathbf{H}_i - \mathbf{S}_i)\mathbf{U}$
	$-[\mathcal{G}_j\mathcal{G}_k]_{j,k=1}^n U$	$-[\mathbf{G}_j\mathbf{G}_k]_{j,k=1}^n\mathbf{U}$
symmetric	$U \mapsto \frac{1}{2}\sum_{i=1}^n (\mathcal{H}_i - \mathcal{S}_i)^*(\mathcal{H}_i - \mathcal{S}_i)U$	$\mathbf{U} \mapsto \frac{1}{2}\sum_{i=1}^n \mathbf{P}^\otimes(\mathbf{H}_i - \mathbf{S}_i)^\top(\mathbf{H}_i - \mathbf{S}_i)\mathbf{P}^\otimes\mathbf{U}$
	$+[\mathcal{G}_j^*\mathcal{G}_k]_{j,k=1}^n U$	$+\mathbf{P}^\otimes[\mathbf{G}_j^\top\mathbf{G}_k]_{j,k=1}^n\mathbf{P}^\otimes\mathbf{U}$
Lichnerowicz Lap.	$\Delta_L : \mathfrak{X}(M) \to \mathfrak{X}(M)$	$\Delta_L : \mathbb{R}^{Nn\times 1} \to \mathbb{R}^{Nn\times 1}$
Sym is the symmetric part	$U \mapsto -\begin{bmatrix}\mathcal{H}_1\\ \vdots\\ \mathcal{H}_n\end{bmatrix} \cdot \operatorname{Sym} \begin{bmatrix}\mathcal{H}_1 U\\ \vdots\\ \mathcal{H}_n U\end{bmatrix}$	$\mathbf{U} \mapsto -\begin{bmatrix}\mathbf{H}_1\\ \vdots\\ \mathbf{H}_n\end{bmatrix} \cdot \operatorname{Sym} \begin{bmatrix}\mathbf{H}_1\mathbf{U}\\ \vdots\\ \mathbf{H}_n\mathbf{U}\end{bmatrix}$
non-symmetric	$= -\sum_{i=1}^n \mathcal{H}_i(\mathcal{H}_i + \mathcal{S}_i)U$	$= -\sum_{i=1}^n \mathbf{H}_i(\mathbf{H}_i + \mathbf{S}_i)\mathbf{U}$
symmetric	$U \mapsto \frac{1}{2}\sum_{i=1}^n (\mathcal{H}_i + \mathcal{S}_i)^*(\mathcal{H}_i + \mathcal{S}_i)U$	$\mathbf{U} \mapsto \frac{1}{2}\sum_{i=1}^n \mathbf{P}^\otimes(\mathbf{H}_i + \mathbf{S}_i)^\top(\mathbf{H}_i + \mathbf{S}_i)\mathbf{P}^\otimes\mathbf{U}$
covariant derivative	$\nabla : \mathfrak{X}(M)\times\mathfrak{X}(M) \to \mathfrak{X}(M)$	$\nabla : \mathbb{R}^{Nn\times 1}\times\mathbb{R}^{Nn\times 1} \to \mathbb{R}^{Nn\times 1}$
	$(U, Y) \mapsto \mathcal{P}\bar{\nabla}_U Y$	$(\mathbf{U}, \mathbf{Y}) \mapsto \mathbf{P}^\otimes\bar{\nabla}_{\mathbf{U}}\mathbf{Y}$

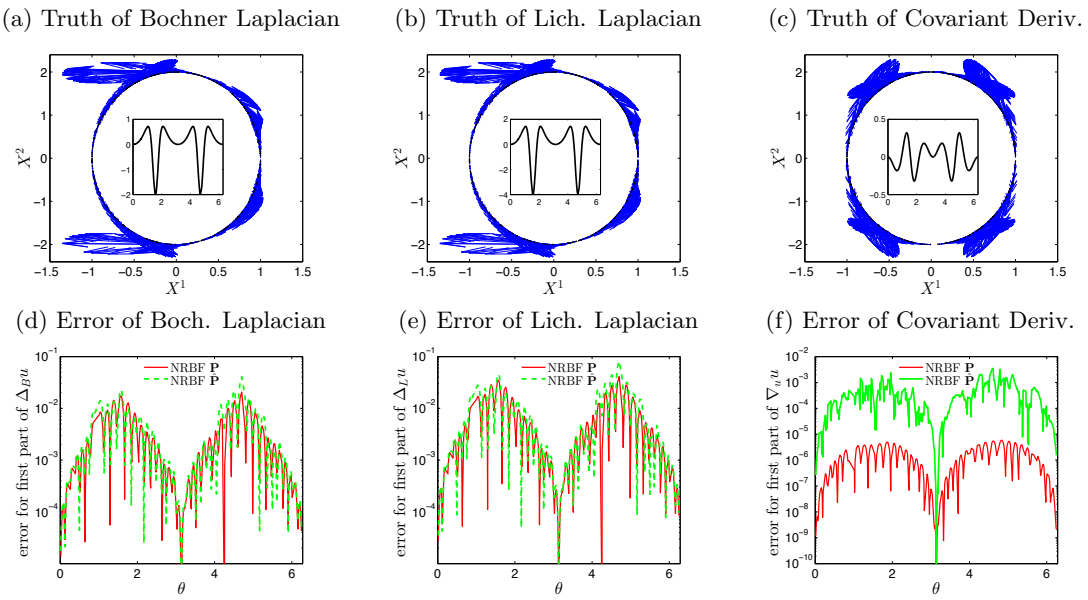

Figure 1: **1D ellipse in $\mathbb{R}^2$.** The upper panels display the truth of (a) Bochner Laplacian, (b) Lichnerowicz Laplacian, and (c) covariant derivative of a vector field. The insets of upper panels display the first components of these operator approximations. The bottom panels display the errors of NRBF approximations using analytic $\mathbf{P}$ (red curve) and approximated $\hat{\mathbf{P}}$ (green curve) for (d) Bochner Laplacian, (e) Lichnerowicz Laplacian, and (f) covariant derivative. The Gaussian kernel with shape parameter $s = 1.5$ was used. The $N = 400$ data points are randomly distributed.

We first approximate the vector Laplacians. We take a vector field $u = u^1 \frac{\partial}{\partial \theta}$ with $u^1(x) \equiv u^1(x(\theta)) = \sin\theta$. The Bochner Laplacian acting on u can be calculated as $\Delta_B u = g^{-1} u^1_{,11} \frac{\partial}{\partial \theta}$, where $u^1_{,11} = \frac{\partial^2 u^1}{\partial \theta^2} + \frac{\partial u^1}{\partial \theta}\Gamma^1_{11} + u^1 \frac{\partial \Gamma^1_{11}}{\partial \theta}$ with $\Gamma^1_{11} = \frac{1}{2} g^{-1} \frac{\partial g}{\partial \theta}$. Numerically, Fig. 1(a) shows the true Bochner Laplacian $\Delta_B u = \Delta_B U$ pointwisely, which is a 2×1 vector lying in the tangent space of each given point. The inset of Fig. 1(a) displays the first vector component of the Bochner Laplacian as a function of the intrinsic coordinate θ. Figure 1(d) displays the error for the first vector component of Bochner Laplacian as a function of θ. Here, we show the errors of using the analytic $\mathbf{P}$ and an approximated $\hat{\mathbf{P}}$. Here (and in the remainder of this paper), we used the notation $\hat{\mathbf{P}}$ to denote the approximated projection matrix obtained from a second-order method discussed in Section 3. One can clearly see that the errors for NRBF using both analytic $\mathbf{P}$ and approximated $\hat{\mathbf{P}}$ are small about 0.01. Since Hodge Laplacian is identical to Bochner Laplacian on a 1D manifold, the results for Hodge Laplacian are almost the same as those for Bochner [not shown here]. The Lichnerowicz Laplacian is the double of Bochner Laplacian, $\Delta_L u = 2 g^{-1} u^1_{,11} \frac{\partial}{\partial \theta}$, as shown in Fig. 1(b). The error for the first vector component of Lichnerowicz Laplacian is also nearly doubled as shown in Fig. 1(e).

We next approximate the covariant derivative. The covariant derivative can be calculated as $\nabla_u u = u^1(\frac{\partial u^1}{\partial \theta} + u^1 \Gamma^1_{11}) \frac{\partial}{\partial \theta}$ as shown in Fig. 1(c). Figure 1(f) displays the error of NRBF approximation for the first vector component of covariant derivative $\nabla_u u$. One can see from Fig. 1(f) that the error for analytic $\mathbf{P}$ (red) is very small about 10^{-6} and the error for approximated $\hat{\mathbf{P}}$ (green) is about 10^{-3}.

3. Estimation of the Projection Matrix

When the manifold M is unknown and identified only by a point cloud data $X = \{x_1, \ldots, x_N\}$, where $x_i \in M$, we do not immediately have access to the matrix-valued function P. In this section, we first give a quick overview of the existing first-order local SVD method for estimating $\mathbf{P} = P(x)$ on each $x \in M$. Subsequently, we present a novel second-order method (which is a local SVD method that corrects the estimation error induced by the curvature) under the assumption that the data set X lies on a C^3 d-dimensional Riemannian manifold M embedded in $\mathbb{R}^n$.

Let $x, y \in X \subset M$ such that $|y - x| = O(\rho)$. Define γ to be a geodesic, connecting x and y. The curve is parametrized by the arc-length,

$$\rho = \int_0^\rho \left|\gamma'(t)\right| dt,$$

where $\gamma(0) = x, \gamma(\rho) = y$. Taking derivative with respect to ρ, we obtain constant velocity, $1 = |\gamma'(t)|$ for all $0 \le t \le \rho$. Let $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_d)$ be the geodesic normal coordinate of y defined by an exponential map $\exp_x : T_x M \to M$. Then $\boldsymbol{\rho}$ satisfies

$$\rho \gamma'(0) = \boldsymbol{\rho} = \exp_x^{-1}(y),$$

where

$$\rho^2 = \rho^2 \left|\gamma'(0)\right|^2 = |\boldsymbol{\rho}|^2 = \sum_{i=1}^d \rho_i^2.$$

For any point $x, y \in M$, let ι be the local parametrization of manifold such that $\iota(\boldsymbol{\rho}) = y$ and $\iota(\mathbf{0}) = x$. Consider the Taylor expansion of $\iota(\boldsymbol{\rho})$ centered at $\mathbf{0}$,

$$\iota(\boldsymbol{\rho}) = \iota(\mathbf{0}) + \sum_{i=1}^{d} \rho_i \frac{\partial \iota(\mathbf{0})}{\partial \rho_i} + \frac{1}{2} \sum_{i,j=1}^{d} \rho_i \rho_j \frac{\partial^2 \iota(\mathbf{0})}{\partial \rho_i \partial \rho_j} + O(\rho^3). \tag{22}$$

Since the Riemannian metric tensor at the based point $\mathbf{0}$ is an identity matrix, $\boldsymbol{\tau}_i = \frac{\partial \iota(\mathbf{0})}{\partial \rho_i} \Big/ \left| \frac{\partial \iota(\mathbf{0})}{\partial \rho_i} \right| = \frac{\partial \iota(\mathbf{0})}{\partial \rho_i}$ are d orthonormal tangent vectors that span $T_x M$.

3.1 First-order Local SVD Method

The classical local SVD method (Donoho and Grimes, 2003; Zhang and Zha, 2004; Tyagi et al., 2013) uses the difference vector $y - x = \iota(\boldsymbol{\rho}) - \iota(\mathbf{0})$ to estimate $\mathbf{T} = (\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_d)$ (up to an orthogonal rotation) and subsequently use it to approximate $\mathbf{P} = \mathbf{T}\mathbf{T}^\top$. The same technique has also been proposed to estimate the intrinsic dimension of the manifold given noisy data (Little et al., 2009). Numerically, the first-order local SVD proceeds as follows:

1. For each $x \in X$, let $\{y_1, \ldots, y_K\} \subset X$ be the K-nearest neighbor (one can also use a radius neighbor) of x. Construct the distance matrix $\mathbf{D} := [\mathbf{D}_1, \ldots, \mathbf{D}_K] \in \mathbb{R}^{n \times K}$, where $K > d$ and $\mathbf{D}_i := y_i - x$.

2. Take a singular value decomposition of $\mathbf{D} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$. Then the leading $d-$columns of $\mathbf{U}$ consists of $\tilde{\mathbf{T}}$ which approximates a span of column vectors of $\mathbf{T}$, which forms a basis of $T_x M$.

3. Approximate $\mathbf{P}$ with $\tilde{\mathbf{P}} = \tilde{\mathbf{T}}\tilde{\mathbf{T}}^\top$.

Based on the Taylor's expansion in (22), intuitively, such an approximation can only provide an estimate with accuracy $\|\tilde{\mathbf{P}} - \mathbf{P}\|_F = O(\rho)$, which is an order-one scheme, where the constant in the big-oh notation, $O(\rho)$, depends on the base point x through the curvature, number of nearest neighbors K, intrinsic dimension d, and extrinsic dimension n, as we discuss next. Here $\|\cdot\|_F$ denotes the Frobenius matrix norm. For uniformly sampled data, we state the following definition and probabilistic type convergence result (Theorem 2 of (Tyagi et al., 2013)) for this local SVD method, which will be useful in our convergence study.

Definition 8 *For each point $x \in M$, where M is a d-dimensional smooth manifold embedded in $\mathbb{R}^n$, where $d+1 \le n$. We define $N_\epsilon(x) = M \cap B_x(\sqrt{\epsilon})$, where $B_x(\sqrt{\epsilon})$ denotes the Euclidean ball (in $\mathbb{R}^n$) centered at x with radius $\sqrt{\epsilon}$. If M has a positive injectivity radius $\sqrt{\epsilon} > 0$ at $x \in M$, then there is a diffeomorphism between $N_\epsilon(x)$ and $T_x M$. In such a case, there exists a local one-to-one map $T_x M \ni \boldsymbol{\rho} \mapsto y = \exp_x(\boldsymbol{\rho}) := (\boldsymbol{\rho}, f_1(\boldsymbol{\rho}), \ldots f_{n-d}(\boldsymbol{\rho})) \in M \subset \mathbb{R}^n$, for $y \in M$ neighboring to x, with smooth functions $f_\ell : T_x M \to \mathbb{R}$ for $\ell = 1, \ldots, n-d$. We also denote the maximum principal curvature at x as K_{max}.*

Theorem 3.1 *Suppose that $\{y_i \in M\}_{i=1}^K$ are the K-nearest neighbor data points of x such that their orthogonal projections, $\boldsymbol{\rho}^{(i)} \sim U[-\sqrt{\epsilon}, \sqrt{\epsilon}]^d \subset T_x M$ at x are i.i.d. Let $\mathbf{W} \in \mathbb{R}^{(n-d)\times(n-d)}$ be a matrix with components given as,*

$$W_{ij} = \mathbb{E}_{\boldsymbol{\rho} \sim U[-\sqrt{\epsilon}, \sqrt{\epsilon}]^d}[f_{q,i}(\boldsymbol{\rho}) f_{q,j}(\boldsymbol{\rho})],$$

where $f_{q,\ell}$ denotes a quadratic form of f_ℓ for $\ell = 1, \dots, n-d$, which is the second-order Taylor expansion about the base point $\mathbf{0}$, involving the curvature at x of M. Then, for any $\tau \in (0,1)$, in high probability,

$$\|\tilde{\mathbf{P}} - \mathbf{P}\|_F \leq \sqrt{2}\tau,$$

if $\epsilon = O(n^{-1}d^{-2}|K_{max}|^{-2})$ and $K = O(\tau^{-2}d^2 \log n)$. Here $\|\cdot\|_F$ denotes the standard Frobenius matrix norm.

Remark 9 *We should point out that the result above implies that the local SVD has a Monte-Carlo error rate, $\tau = O(K^{-1/2})$ and the choice of local neighbor of radius $\sqrt{\epsilon} = O(\rho)$ should be inversely proportional to the maximum principal curvature and the dimension of the manifold and ambient space. Thus, to expect an error $\tau = O(\sqrt{\epsilon}) = O(\rho)$, by balancing $n^{-1}d^{-2}|K_{max}|^{-2} \sim K^{-1}d^2 \log n$, this result suggests that one should choose $K \sim d^4 n \log n |K_{max}|^2$. The numerical result in the torus suggests of rate $\sqrt{\epsilon} = O(N^{-1/2})$ (see Figure 2). For general d-dimensional manifolds, the error rate is expected to be $\sqrt{\epsilon} \sim N^{-1/d}$ due to the fact that $\rho \propto N^{-1/d}$.*

3.2 Second-order Local SVD Method

In this section, we devise an improved scheme to achieve the tangent space approximation with accuracy up to order of $O(\rho^2)$, by accounting for the Hessian components in (22). The algorithm proceeds as follows:

1. Perform the first-order local SVD algorithm and attain the d approximated tangent vectors $\tilde{\mathbf{T}} = [\tilde{\boldsymbol{t}}_1, \dots, \tilde{\boldsymbol{t}}_d] \in \mathbb{R}^{n \times d}$.

2. For each neighbor $\{y_i\}_{i=1,\dots,K}$ of x, compute $\tilde{\boldsymbol{\rho}}^{(i)} = (\tilde{\rho}_1^{(i)}, \dots, \tilde{\rho}_d^{(i)})$, where

$$\tilde{\rho}_j^{(i)} = \mathbf{D}_i^\top \tilde{\boldsymbol{t}}_j, \quad i = 1, \dots, K, j = 1, \dots, d,$$

   where $\mathbf{D}_i := y_i - x$ is the ith column of $\mathbf{D} \in \mathbb{R}^{n \times K}$.

3. Approximate the Hessian $\mathbf{Y}_p = \frac{\partial^2 \iota(\mathbf{0})}{\partial \rho_i \partial \rho_j} \in \mathbb{R}^n$ up to a difference of a vector in $T_x M$ using the following ordinary least squared regression, with $p = 1, \dots, D = d(d+1)/2$ denoting the upper triangular components ($p \mapsto (i,j)$ such that $i \leq j$) of symmetric Hessian matrix. Notice that for each $y_\ell \in \{y_1, \dots, y_K\}$ neighbor of x, the equation (22) can be written as

$$\sum_{i,j=1}^{d} \rho_i^{(\ell)} \rho_j^{(\ell)} \frac{\partial^2 \iota(\mathbf{0})}{\partial \rho_i \partial \rho_j} = 2\left(\iota(\boldsymbol{\rho}^{(\ell)}) - \iota(\mathbf{0})\right) - 2\sum_{i=1}^{d} \rho_i^{(\ell)} \boldsymbol{\tau}_i + O(\rho^3), \qquad \ell = 1, \dots K, \quad (23)$$

   where $\boldsymbol{\rho}^{(\ell)} := (\rho_1^{(\ell)}, \dots, \rho_d^{(\ell)})$ denotes the geodesic coordinate that satisfies $\iota\left(\boldsymbol{\rho}^{(\ell)}\right) = y_\ell$. In compact form, we can rewrite (23) as a linear system,

$$\mathbf{A}\mathbf{Y} = 2\mathbf{D}^\top - 2\mathbf{R} + O(\rho^3), \quad (24)$$

where $\mathbf{R}^\top = (\mathbf{r}_1, \ldots, \mathbf{r}_K) \in \mathbb{R}^{n\times K}$ denotes the order-ρ residual term in the tangential directions,

$$\mathbf{r}_j = \sum_{i=1}^{d} \rho_i^{(j)} \boldsymbol{\tau}_i, \quad j = 1, \ldots, K. \tag{25}$$

Here, $\mathbf{Y} \in \mathbb{R}^{D\times n}$ is a matrix whose pth row is $\mathbf{Y}_p$ and

$$\mathbf{A} = \begin{pmatrix} (\rho_1^{(1)})^2 & (\rho_2^{(1)})^2 & \ldots & (\rho_d^{(1)})^2 & 2(\rho_1^{(1)}\rho_2^{(1)}) & \ldots & 2(\rho_{d-1}^{(1)}\rho_d^{(1)}) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ (\rho_1^{(K)})^2 & (\rho_2^{(K)})^2 & \ldots & (\rho_d^{(K)})^2 & 2(\rho_1^{(K)}\rho_2^{(K)}) & \ldots & 2(\rho_{d-1}^{(K)}\rho_d^{(K)}) \end{pmatrix} \in \mathbb{R}^{K\times D}. \tag{26}$$

With the choice of K in Remark 9, $K > D$, we approximate $\mathbf{Y}$ by solving an over-determined linear problem

$$\tilde{\mathbf{A}}\mathbf{Y} = 2\mathbf{D}^\top, \tag{27}$$

where $\tilde{\mathbf{A}}$ is defined as in (26) except that $\rho_i^{(j)}$ in the matrix entries is replaced by $\tilde{\rho}_i^{(j)}$. The regression solution is given by $\tilde{\mathbf{Y}} = 2(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top\mathbf{D}^\top$. Here $\tilde{\mathbf{Y}}_{ij} = \tilde{Y}_i^{(j)}$, $i = 1, \ldots, D, j = 1, \ldots, n$. Here, each row of $\tilde{\mathbf{Y}}$ is denoted as $\tilde{\mathbf{Y}}_p = (\tilde{Y}_p^{(1)}, \ldots, \tilde{Y}_p^{(n)}) \in \mathbb{R}^{1\times n}$, which is an estimator of $\mathbf{Y}_p$.

4. Apply SVD to

$$2\tilde{\mathbf{R}}^\top := 2\mathbf{D} - (\tilde{\mathbf{A}}\tilde{\mathbf{Y}})^\top \in \mathbb{R}^{n\times K}. \tag{28}$$

Let the leading d left singular vectors be denoted as $\hat{\boldsymbol{\Psi}} = [\hat{\boldsymbol{\psi}}_1, \ldots, \hat{\boldsymbol{\psi}}_d] \in \mathbb{R}^{n\times d}$, which is an estimator of $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_d]$, where $\boldsymbol{\psi}_j$ are the leading d left singular vectors of $\mathbf{R}$ as defined in (24). We define $\hat{\mathbf{P}} = \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^\top$ as the estimator for $\mathbf{P} = \boldsymbol{\Psi}\boldsymbol{\Psi}^\top$, where the last equality is valid due to Proposition 2.1(3).

Figure 2 shows the manifold learning results on a torus with randomly distributed data. One can see that error of the first-order local SVD method is $O(N^{-1/2})$ whereas second-order method is $O(N^{-1})$.

Theoretically, we can deduce the following error bound.

Theorem 3.2 *Let the assumptions in Theorem 3.1 be valid, particularly $\rho \sim \epsilon^{1/2}$. Suppose that the matrix $\mathbf{D} \in \mathbb{R}^{n\times K}$ is defined as in Step 1 of the algorithm with a fixed K is chosen as in Remark 9 in addition to $K > D = \frac{1}{2}d(d+1)$. Assume that $\|\mathbf{A}\|_2/\kappa_2(\mathbf{A}) = K\omega(\epsilon^{3/2})$ as $\epsilon \to 0$, where $\kappa_2(\mathbf{A})$ denotes the condition number of matrix $\mathbf{A}$ based on spectral matrix norm, $\|\cdot\|_2$, and the eigenvalues $\{\lambda_i\}_{i=1,\ldots,n}$ of $\mathbf{R}^\top\mathbf{B}^\top\mathbf{B}\mathbf{R}$, where $\mathbf{R}^\top \in \mathbb{R}^{n\times K}$ as defined in (25), are simple with spectral gap $g_i := \min_{j\neq i}|\lambda_i - \lambda_j| > c\epsilon$ for some $c > 0$ and all $i = 1, \ldots, n$. Here, $\mathbf{B} := \mathbf{I}_K - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top \in \mathbb{R}^{K\times K}$. Let $\hat{\mathbf{P}} = \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^\top$ be the second-order estimator of $\mathbf{P}$, where columns of $\hat{\boldsymbol{\Psi}}$ are the leading d left singular vectors of $\tilde{\mathbf{R}}$ as defined in (28). Then, with high probability,*

$$\|\hat{\mathbf{P}} - \mathbf{P}\|_F = O(\epsilon),$$

as $\epsilon \to 0$.

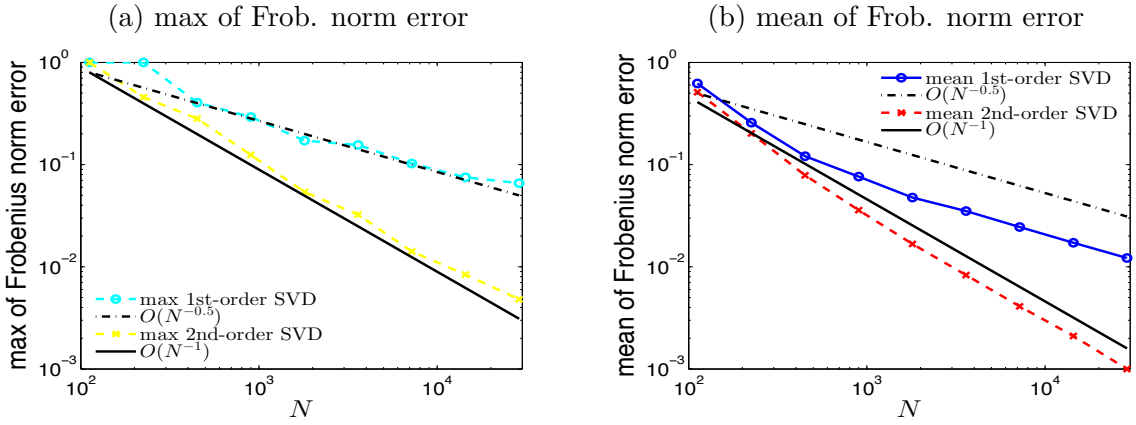

Figure 2: **2D torus in $\mathbb{R}^3$.** Comparison of convergence rates between 1st-order SVD and our 2nd-order SVD for approximating the tangential projection matrix $\mathbf{P}$. Panels (a) and (b) show the maximum and the mean of Frobenius norm errors, respectively. The error of $||\tilde{\mathbf{P}} - \mathbf{P}||_F$ is $O(N^{-1/2})$ for the 1st-order SVD whereas the error of $||\hat{\mathbf{P}} - \mathbf{P}||_F$ is $O(N^{-1})$ for our 2nd-order SVD. The $K = 40$ nearest neighbors are fixed for all the simulations. The data points are uniformly distributed in intrinsic coordinates $[0, 2\pi) \times [0, 2\pi)$ and are then mapped onto the torus in Euclidean space.

The assumption of $\|\mathbf{A}\|_2/\kappa_2(\mathbf{A}) = K\omega(\epsilon^{3/2})$ as $\epsilon \to 0$ is to ensure that the perturbed matrix, $\tilde{\mathbf{A}}$, is still full rank. As we will show, this condition arises from the fact that the minimum relative size of the perturbation for the perturbed matrix to be not full rank is $\frac{1}{\kappa_2(\mathbf{A})}$, i.e.,

$$\min\left\{\frac{\|\tilde{\mathbf{A}} - \mathbf{A}\|_2}{\|\mathbf{A}\|_2} \,:\, \tilde{\mathbf{A}} \text{ is not full rank}\right\} = \frac{1}{\kappa_2(\mathbf{A})}.$$

The simple eigenvalues and spectral gap conditions in the theorem above are two technical assumptions needed for applying the classical perturbation theory of eigenvectors estimation (see e.g. Theorem 5.4 of Demmel (1997)), which allow one to bound the angle between eigenvectors of unperturbed and perturbed matrices by the ratio of the perturbation error and the spectral gap as we shall see in the proof below. One can employ the result in Theorem 3.2 whenever $\mathbf{BR}$ can be approximated by $\tilde{\mathbf{R}}$ sufficiently well (with an error smaller than the spectral gap of the corresponding eigenvalue). One can see from Fig. 2(b) that the asymptotics break down when N decreases to around 500. Moreover, when N is around 500, the max norm errors are already large up to at least 0.3 for both 1st-order and 2nd-order methods as seen from Fig. 2(a).

While we have no access to $\mathbf{R}^\top\mathbf{B}^\top\mathbf{BR}$, since it can be accurately estimated by $\tilde{\mathbf{R}}^\top\tilde{\mathbf{R}}$ in the sense of $\|\tilde{\mathbf{R}}^\top\tilde{\mathbf{R}} - \mathbf{R}^\top\mathbf{B}^\top\mathbf{BR}\|_2 = O(\epsilon^2)$ as we shall see in the following proof, let us investigate the eigenvalues of the approximate matrix $\tilde{\mathbf{R}}^\top\tilde{\mathbf{R}}$ (or equivalently the singular values of $2\tilde{\mathbf{R}}^\top$). Figures 3(a) and (b) show the first two singular values of $2\tilde{\mathbf{R}}^\top$ and their difference for the torus and sphere examples, respectively. One can see the spectral gap $|\tilde{\lambda}_1 -$

(a) Torus, singular values of $2\tilde{\mathbf{R}}^\top$ (b) Sphere, singular values of $2\tilde{\mathbf{R}}^\top$ (c) Pointwise Frobenius norm error

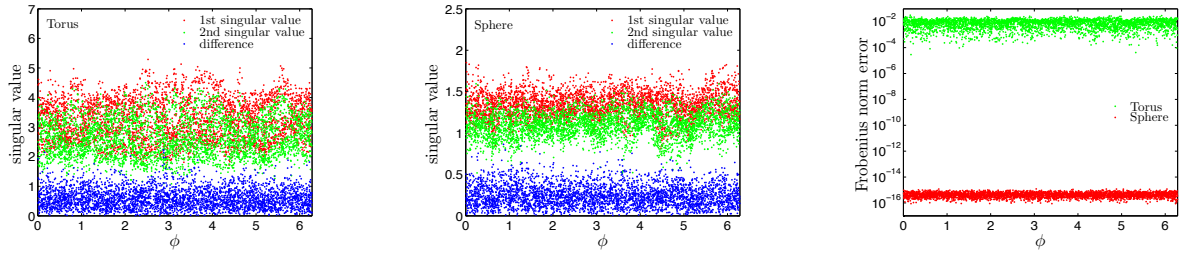

Figure 3: **2D torus and 2D sphere in $\mathbb{R}^3$.** Comparison of the 1st singular value $\tilde{\lambda}_1$ of $2\tilde{\mathbf{R}}^\top$ (red dots), the 2nd singular value $\tilde{\lambda}_2$ of $2\tilde{\mathbf{R}}^\top$ (green dots), and their difference $|\tilde{\lambda}_1 - \tilde{\lambda}_2|$ (blue dots) for (a) 2D torus and (b) 2D unit sphere examples. (c) Pointwise Frobenius norm errors, $\|\hat{\mathbf{P}} - \mathbf{P}\|_F$, for the 2D torus (green dots) and 2D sphere (red dots) examples. The horizontal axis corresponds to one of the intrinsic coordinate $\phi \in [0, 2\pi)$. Randomly distributed $N = 3600$ data points are used on the manifolds.

$\tilde{\lambda}_2|$ on the same scale of $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ almost surely, regardless whether the geometry is highly symmetric (e.g. the sphere) or not (e.g. torus) for a fixed N. This empirical verification suggests that the two assumptions (simple eigenvalues and spectral gap condition) are not unreasonable. In fact, in Fig. 3(c), we found that the corresponding pointwise Frobenius norm errors, $\|\hat{\mathbf{P}} - \mathbf{P}\|_F$, for the torus and sphere examples using the 2nd order SVD method are small, especially for the highly symmetric sphere. While these examples suggest that the eigenvalues are most likely simple for randomly sample data of any fixed N, we believe that one can use other results from perturbation theory that may require different assumptions if the eigenvalues are non-simple. See, for instance, Chapter 2, Section 6.2 of Kato (2013).

We are now ready to prove Theorem 3.2.

Proof Recall that $\mathbf{T} = (\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_d) \in \mathbb{R}^{n\times d}$ is a set of orthonormal tangent vectors such that $\mathbf{P} = \mathbf{T}\mathbf{T}^\top$. Also recall that $\tilde{\mathbf{T}} = (\tilde{\boldsymbol{t}}_1, \ldots, \tilde{\boldsymbol{t}}_d) \in \mathbb{R}^{n\times d}$ is the d approximated tangent vectors such that $\tilde{\mathbf{P}} = \tilde{\mathbf{T}}\tilde{\mathbf{T}}^\top$. Then there exists some orthogonal matrix $\mathbf{O} \in \mathbb{R}^{d\times d}$ such that we have $\tilde{\mathbf{T}} = \mathbf{T}\mathbf{O} + O(\epsilon^{1/2})$ based on the first-order approximation result $\|\tilde{\mathbf{P}} - \mathbf{P}\|_F = O(\epsilon^{1/2})$.

This claim can be verified as follows. First, since columns of $\tilde{\mathbf{T}}$ are eigenvectors of $\tilde{\mathbf{P}}$ corresponding to eigenvalue one, it is clear that $\mathbf{P}\tilde{\boldsymbol{t}}_i = \tilde{\mathbf{P}}\tilde{\boldsymbol{t}}_i + (\mathbf{P} - \tilde{\mathbf{P}})\tilde{\boldsymbol{t}}_i = \tilde{\boldsymbol{t}}_i + O(\epsilon^{1/2})$. This also means that $(\mathbf{P}\tilde{\boldsymbol{t}}_i)^\top(\mathbf{P}\tilde{\boldsymbol{t}}_j) = \delta_{i,j} + O(\epsilon^{1/2})$ for $i \neq j$. Since columns of $\mathbf{P}\tilde{\mathbf{T}}$ are in $\text{span}\{\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_d\}$, then it is clear that there exists an orthogonal matrix $\mathbf{O}$ such that $\mathbf{P}\tilde{\mathbf{T}} = \mathbf{T}\mathbf{O} + O(\epsilon^{1/2})$. Thus, $\tilde{\mathbf{T}} = \tilde{\mathbf{P}}\tilde{\mathbf{T}} = \mathbf{P}\tilde{\mathbf{T}} + O(\epsilon^{1/2}) = \mathbf{T}\mathbf{O} + O(\epsilon^{1/2})$, where the second equality follows from the fact that $\tilde{\mathbf{P}} - \mathbf{P} = O(\epsilon^{1/2})$ for each entry.

Let $\mathbf{T}\mathbf{O} = (\boldsymbol{t}_1, \ldots, \boldsymbol{t}_d) \in \mathbb{R}^{n\times d}$, and we have $\tilde{\boldsymbol{t}}_j - \boldsymbol{t}_j = O(\epsilon^{1/2})$ for $j = 1, \ldots, d$. We can write

$$\rho_j^{(i)} = \mathbf{D}_i^\top \boldsymbol{t}_j + O(\epsilon),$$

using the fact that components of $\mathbf{D}_i$ are $O(\epsilon^{1/2})$. Based on Step 2 of the Algorithm above and Theorem 3.1, we have in high probability, for any $i = 1, \ldots, K$,

$$|\tilde{\rho}_j^{(i)} - \rho_j^{(i)}| = |\mathbf{D}_i^\top(\tilde{\boldsymbol{t}}_j - \boldsymbol{t}_j)| + O(\epsilon) \leq \|\mathbf{D}_i\|\|\tilde{\boldsymbol{t}}_j - \boldsymbol{t}_j\| + O(\epsilon) \leq c\epsilon^{1/2}\epsilon^{1/2} + O(\epsilon) \leq 2c\epsilon.$$

for some $c > 0$, as $\epsilon \to 0$, using the fact that $\|\mathbf{D}_i\| = O(\epsilon^{1/2})$. This implies that,

$$|\tilde{\rho}_i^{(k)}\tilde{\rho}_j^{(k)} - \rho_i^{(k)}\rho_j^{(k)}| \leq |\tilde{\rho}_i^{(k)}||\tilde{\rho}_j^{(k)} - \rho_j^{(k)}| + |\rho_j^{(k)}||\tilde{\rho}_i^{(k)} - \rho_i^{(k)}| = O(\epsilon^{3/2}),$$

where we used again $|\rho_j| = O(\epsilon^{1/2})$. This means, $\|\mathbf{A} - \tilde{\mathbf{A}}\|_2 \leq \sqrt{KD}\|\mathbf{A} - \tilde{\mathbf{A}}\|_{max} = O(K\epsilon^{3/2})$.

We note that the components of each row of matrix $\mathbf{A}$ forms a homogeneous polynomial of degree-2, so they are linearly independent. Since the data points are uniformly sampled from $[-\sqrt{\epsilon}, \sqrt{\epsilon}]^d$, for large enough samples, $K > D$, these sample points will not lie on a subspace of dimension strictly less than d. Thus, $\mathbf{A}$ is not degenerate and $\text{rank}(\mathbf{A}) = D$ almost surely.

Furthermore, there exists a constant $C > 0$ such that,

$$\frac{\|\mathbf{A} - \tilde{\mathbf{A}}\|_2}{\|\mathbf{A}\|_2} \leq \frac{CK\epsilon^{3/2}}{\|\mathbf{A}\|_2} < \frac{1}{\kappa_2(\mathbf{A})},$$

where we used the assumption $\|\mathbf{A}\|_2/\kappa_2(\mathbf{A}) = K\omega(\epsilon^{3/2})$, as $\epsilon \to 0$ for the last inequality, to ensure that the perturbed matrix $\tilde{\mathbf{A}}$ is still full rank.

From (24), one can deduce,

$$\underbrace{\mathbf{Y}}_{O(1)} = \underbrace{(\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top}_{O(\frac{1}{\rho^2})}\underbrace{(2\mathbf{D}^\top - 2\mathbf{R})}_{O(\rho^2)} + O(\rho), \tag{29}$$

where the leading order terms on both sides are $O(1)$ since both terms $\mathbf{A}$ and $2\mathbf{D}^\top - 2\mathbf{R}$ are $O(\rho^2)$. Since $\tilde{\mathbf{A}}$ is full rank, we can solve the regression problem in (27) as,

$$\tilde{\mathbf{Y}} = 2(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top\mathbf{D}^\top. \tag{30}$$

Using (29) and (30), one has

$$\begin{aligned}
2\tilde{\mathbf{R}} &= 2\mathbf{D}^\top - \tilde{\mathbf{A}}\tilde{\mathbf{Y}} = [2\mathbf{D}^\top - \mathbf{A}\mathbf{Y}] + [\mathbf{A}\mathbf{Y} - \tilde{\mathbf{A}}\tilde{\mathbf{Y}}] \\
&= [2\mathbf{R} + O(\rho^3)] + [\mathbf{A}(\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top(2\mathbf{D}^\top - 2\mathbf{R}) + O(\rho^3) - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top 2\mathbf{D}^\top] \\
&= 2\mathbf{R} + (\mathbf{A}(\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top)(2\mathbf{D}^\top - 2\mathbf{R}) - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top 2\mathbf{R} + O(\rho^3) \\
&= \underbrace{(\mathbf{A}(\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top)}_{O(\rho)}\underbrace{(2\mathbf{D}^\top - 2\mathbf{R})}_{O(\rho^2)} \\
&\quad + \underbrace{(\mathbf{I}_K - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top)2\mathbf{R}}_{O(\rho)} + O(\rho^3) \\
&= (\mathbf{I}_K - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^\top\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top)2\mathbf{R} + O(\rho^3).
\end{aligned}$$

The key point here is to notice that the remaining term $2\mathbf{BR}$, where $\mathbf{B} := \mathbf{I}_K - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^\top \tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^\top$, is in the tangent space and all the normal direction terms up to order of $O(\rho^2)$ are cancelled out. Then, one can identify $\text{span}\{\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_d\}$ (or span of the leading d-eigenvectors of $\mathbf{R}^\top\mathbf{B}^\top\mathbf{BR}$ corresponding to nontrivial spectra) by computing the leading d-eigenvectors of $\tilde{\mathbf{R}}^\top\tilde{\mathbf{R}}$. Moreover,

$$\|\tilde{\mathbf{R}}^\top\tilde{\mathbf{R}} - \mathbf{R}^\top\mathbf{B}^\top\mathbf{BR}\|_\infty \le \|\tilde{\mathbf{R}}^\top\|_\infty\|\tilde{\mathbf{R}} - \mathbf{BR}\|_\infty + \|\tilde{\mathbf{R}}^\top - \mathbf{R}^\top\mathbf{B}^\top\|_\infty\|\mathbf{BR}\|_\infty = O(\rho^4).$$

Similarly, $\|\tilde{\mathbf{R}}^\top\tilde{\mathbf{R}} - \mathbf{R}^\top\mathbf{B}^\top\mathbf{BR}\|_1 = O(\rho^4)$. Therefore, $\|\tilde{\mathbf{R}}^\top\tilde{\mathbf{R}} - \mathbf{R}^\top\mathbf{B}^\top\mathbf{BR}\|_2 = O(\rho^4) = O(\epsilon^2)$.

Let $\{\boldsymbol{\psi}_i\}_{i=1,\ldots,n}$ and $\{\tilde{\boldsymbol{\psi}}_i\}_{i=1,\ldots,n}$ be the unit eigenvectors of $\mathbf{R}^\top\mathbf{B}^\top\mathbf{BR}$ and $\tilde{\mathbf{R}}^\top\tilde{\mathbf{R}}$, respectively. By Theorem 5.4 of (Demmel, 1997) and the assumption on the spectral gap of $\mathbf{R}^\top\mathbf{B}^\top\mathbf{BR}$, $g_i > c\epsilon$ for some $c > 0$, the acute angle between $\boldsymbol{\psi}_i$ and $\tilde{\boldsymbol{\psi}}_i$ satisfies,

$$\sin(2\theta) \le 2\frac{\|\mathbf{R}^\top\mathbf{B}^\top\mathbf{BR} - \tilde{\mathbf{R}}^\top\tilde{\mathbf{R}}\|_2}{g_i} = O(\epsilon),$$

where the constant in the big-oh notation above depends on n. Then,

$$\|\boldsymbol{\psi}_i - \tilde{\boldsymbol{\psi}}_i\|^2 \;=\; \|\boldsymbol{\psi}_i\|^2 + \|\tilde{\boldsymbol{\psi}}_i\|^2 - 2\boldsymbol{\psi}_i^\top\tilde{\boldsymbol{\psi}}_i = 2(1-\cos(\theta)) = 4\sin^2\left(\frac{\theta}{2}\right) = O(\epsilon^2). \quad (31)$$

By Proposition 2.1(3), it is clear that $\mathbf{P} := \boldsymbol{\Psi}\boldsymbol{\Psi}^\top$. Based on the step 4 of the Algorithm, $\hat{\mathbf{P}} := \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}^\top$ and from the error bound in (31), the proof is complete.

■

Remark 10 *The result above is consistent with the intuition that the scheme is of order $\rho^2 = O(\epsilon)$. The numerical result in the torus suggests a rate of $\epsilon = O(N^{-1})$. For general d-dimensional manifolds, the error rate for the second-order method is expected to be $\epsilon \sim N^{-2/d}$, due to the fact that $\rho \propto N^{-1/d}$.*

4. Spectral Convergence Results

In this section, we state spectral convergence results for the symmetric estimator $\mathbf{G}^\top\mathbf{G}$ to the Laplace-Beltrami operator Δ_M, as well as analogous results for the symmetric estimator of the Bochner Laplacian $\mathbf{P}^\otimes\mathbf{H}^\top\mathbf{H}\mathbf{P}^\otimes$ to the Bochner Laplacian Δ_B on vector fields. For definitions of the operators and estimators of this section, please see Sections 2.2 and 2.6. To keep the section short, we only present the proof for the spectral convergence of the Laplace-Beltrami operator. We document the proofs of the intermediate bounds needed for this proof in Appendices B, C.1 and C.2. We should point out that the symmetry of discrete estimators $\mathbf{G}^\top\mathbf{G}$ and $\mathbf{P}^\otimes\mathbf{H}^\top\mathbf{H}\mathbf{P}^\otimes$ allows the convergence of the eigenvectors to be proved as well. Since the proof of eigenvector convergence is more technical, we present it in Appendix C.3. The same techniques used to prove convergence results for the Laplace-Beltrami operator are used to show the results for the Bochner Laplacian acting on vector fields, thanks to the similarity in definitions for the Bochner Laplacian and the Laplace-Beltrami operator and the similarity in our discrete estimators. Since these proofs

follow the same arguments as those for the Laplace-Beltrami operator, we document them in Appendix D. While we suspect the proof technique can also be used to show the spectral convergence for the Hodge and Lichnerowicz Laplacians, the exact calculation can be more involved since the weak forms of these two Laplacians have more terms compared to that of the Bochner Laplacian.

4.1 Spectral Convergence for Laplace-Beltrami

Given a data set of points $X = \{x_1, \dots x_N\} \subset M$ and a function $f : M \to \mathbb{R}$, recall that the interpolation $I_{\phi_s} f$ only depends on $f(x_1), \dots, f(x_N)$. Hence, I_{ϕ_s} can be viewed, after restricting to the manifold, as a map either from $\{f : M \to \mathbb{R}\} \to C^{\alpha - \frac{(n-d)}{2}}(M)$, or as a map

$$I_{\phi_s} : \mathbb{R}^N \to C^{\alpha - \frac{(n-d)}{2}}(M).$$

For details regarding the loss of regularity which occurs when restricting to the d-dimensional submanifold $M \subset \mathbb{R}^n$, see the beginning of Section 2.3 in (Fuselier and Wright, 2012). For the Laplace-Beltrami operator, our focus is on continuous estimators, so we presently regard I_{ϕ_s} as a map

$$I_{\phi_s} : \{f : M \to \mathbb{R}\} \to C^{\alpha - \frac{(n-d)}{2}}(M).$$

Based on Theorem 10 in (Fuselier and Wright, 2012), one can deduce the following interpolation error.

Lemma 4.1 *Let ϕ_s be a kernel whose RKHS norm equivalent to Sobolev space of order $\alpha > n/2$. Then there is sufficiently large $N = |X|$ such that with probability higher than $1 - \frac{1}{N}$, for all $f \in H^{\alpha - \frac{(n-d)}{2}}(M)$, we have*

$$\|I_{\phi_s} f - f\|_{L^2(M)} = O\left(N^{\frac{-2\alpha + (n-d)}{2d}}\right).$$

Proof See Appendix B.2. ■

To ensure that derivatives of interpolations of smooth functions are bounded, we must have an interpolator that is sufficiently regular. While many results in this paper hold whenever the RKHS induced by I_{ϕ_s} is norm equivalent to a Sobolev space of order $\alpha > n/2$, we require slightly higher regularity to prove spectral convergence. In particular, we assume the following.

Assumption 4.1 (Sufficiently Regular Interpolator) *Assume that I_{ϕ_s} has an RKHS norm equivalent to $H^\alpha(\mathbb{R}^n)$ with $\alpha \geq n/2 + 3$.*

This assumption allows us to conclude that whenever $f \in C^\infty(M)$, we have the following equation: $\|I_{\phi_s} f\|_{W^{2,\infty}(M)} = O(1)$, where the constants depend on $M, \|f\|_{W^{2,\infty}(M)}$, and $\|f\|_{H^{\alpha-(n-d)/2}(M)}$. This assumption will be needed for the uniform boundedness of random variables in the concentration inequality in the following lemmas. This statement is made reported concisely as Lemma B.3 in Appendix B. Before we prove the spectral convergence result, let us state the following two concentration bounds that will be needed in the proof.

Lemma 4.2 *Let $f, h \in C^\infty(M)$. Then with probability higher than $1 - \frac{2}{N}$,*

$$\left|\langle \Delta_M f, h\rangle_{L^2(M)} - \langle \mathrm{grad}_g I_{\phi_s} f, \mathrm{grad}_g I_{\phi_s} h\rangle_{L^2(\mathfrak{X}(M))}\right| \leq C N^{\frac{-2\alpha+(n-d)}{2d}},$$

as $N \to \infty$, where the constant C depends on $\|\Delta_M f\|_{L^2(M)} + \|\Delta_M I_{\phi_s} h\|_{L^2(M)}$.

Proof See Appendix C.1. In the course of the proof, we note that the prove inherits the interpolation error rate in Lemma 4.1. ■

For the discretization, we need to define an appropriate inner product such that it is consistent with the inner product of $L^2(M)$ as the number of data points N approaches infinity. In particular, we have the following definition.

Definition 11 *Given two vectors $\mathbf{f}, \mathbf{h} \in \mathbb{R}^N$, we define*

$$\langle \mathbf{f}, \mathbf{h}\rangle_{L^2(\mu_N)} := \frac{1}{N}\mathbf{f}^\top \mathbf{h}.$$

Similarly, we denote by $\|\cdot\|_{L^2(\mu_N)}$ the norm induced by the above inner product.

We remark that when $\mathbf{f}$ and $\mathbf{h}$ are restrictions of functions f and h, respectively, then the above can be evaluated as $\langle \mathbf{f}, \mathbf{h}\rangle_{L^2(\mu_N)} = \frac{1}{N}\sum_{i=1}^N f(x_i)h(x_i)$.

By the sufficiently regular interpolator in Assumption 4.1, it is clear that the concentration bound in the lemma above converges as $N \to \infty$. The next concentration bound is as follows:

Lemma 4.3 *Let $f, h \in C^\infty(M)$. Let the sufficiently regular interpolator Assumption 4.1 be valid. Then with probability higher than $1 - \frac{2}{N}$,*

$$\left|\langle \mathbf{G}^\top \mathbf{G}\mathbf{f}, \mathbf{h}\rangle_{L^2(\mu_N)} - \langle \mathrm{grad}_g I_{\phi_s} f, \mathrm{grad}_g I_{\phi_s} h\rangle_{L^2(\mathfrak{X}(M))}\right| \leq C\frac{\sqrt{\log N}}{\sqrt{N}}$$

for some constant $C > 0$, as $N \to \infty$.

Proof See Appendix C.2. ■

The main results for the Laplace-Beltrami operator are as follows. First, we have the following eigenvalue convergence result.

Theorem 4.1 *(convergence of eigenvalues: symmetric formulation) Let λ_i denote the i-th eigenvalue of Δ_M, enumerated $\lambda_1 \leq \lambda_2 \leq \ldots$, and fix some $i \in \mathbb{N}$. Assume that $\mathbf{G}$ is as defined in (6) with interpolation operator that satisfies the hypothesis in Assumption 4.1. Then there exists an eigenvalue $\hat{\lambda}_i$ of $\mathbf{G}^\top\mathbf{G}$ such that*

$$\left|\lambda_i - \hat{\lambda}_i\right| \leq O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right), \tag{32}$$

with probability greater than $1 - \frac{12}{N}$.

Proof Enumerate the eigenvalues of $\mathbf{G}^\top\mathbf{G}$ and label them $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_N$. Let $\mathcal{S}_i' \subseteq C^\infty(M)$ denote an i-dimensional subspace of smooth functions on which the quantity $\max_{f\in\mathcal{S}_i} \frac{\langle \mathbf{G}^\top\mathbf{G}R_N f, R_N f\rangle_{L^2(\mu_N)}}{\|R_N f\|_{L^2(\mu_N)}}$ achieves its minimum. Let $\tilde{f} \in \mathcal{S}_i'$ be the function on which the maximum $\max_{f\in\mathcal{S}_i'}\langle \Delta_M f, f\rangle_{L^2(M)}$ occurs. WLOG, assume that $\|\tilde{f}\|^2_{L^2(M)} = 1$. Assume that N is sufficiently large so that by Hoeffding's inequality $\left|\|R_N\tilde{f}\|^2_{L^2(\mu_N)} - 1\right| \leq \frac{\text{Const}}{\sqrt{N}} \leq 1/2$, with probability $1 - \frac{2}{N}$, so that $\|R_N\tilde{f}\|^2_{L^2(\mu_N)}$ is bounded away from zero. Hence, we can Taylor expand $\frac{\langle \mathbf{G}^\top\mathbf{G}R_N\tilde{f}, R_N\tilde{f}\rangle_{L^2(\mu_N)}}{\|R_N\tilde{f}\|^2_{L^2(\mu_N)}}$ to obtain

$$\frac{\langle \mathbf{G}^\top\mathbf{G}R_N\tilde{f}, R_N\tilde{f}\rangle_{L^2(\mu_N)}}{\|R_N\tilde{f}\|^2_{L^2(\mu_N)}} = \langle \mathbf{G}^\top\mathbf{G}R_N\tilde{f}, R_N\tilde{f}\rangle_{L^2(\mu_N)} - \frac{\text{Const}\langle \mathbf{G}^\top\mathbf{G}R_N\tilde{f}, R_N\tilde{f}\rangle_{L^2(\mu_N)}}{\sqrt{N}}.$$

By Lemmas 4.2 and 4.3, with probability higher than $1 - \frac{4}{N}$, we have that

$$\left|\langle \mathbf{G}^\top\mathbf{G}R_N\tilde{f}, R_N\tilde{f}\rangle_{L^2(\mu_N)} - \langle \Delta_M\tilde{f}, \tilde{f}\rangle_{L^2(M)}\right| = O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

Combining the two bounds above, we obtain that with probability higher than $1 - \frac{6}{N}$,

$$\langle \Delta_M\tilde{f}, \tilde{f}\rangle_{L^2(M)} \leq \frac{\langle \mathbf{G}^\top\mathbf{G}R_N\tilde{f}, R_N\tilde{f}\rangle_{L^2(\mu_N)}}{\|R_N\tilde{f}\|^2_{L^2(\mu_N)}} + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

Since $\tilde{f}$ is the function on which $\langle \Delta_M f, f\rangle_{L^2(M)}$ achieves its maximum over all $f \in \mathcal{S}_i'$, and since certainly

$$\frac{\langle \mathbf{G}^\top\mathbf{G}R_N\tilde{f}, R_N\tilde{f}\rangle_{L^2(\mu_N)}}{\|R_N\tilde{f}\|^2_{L^2(\mu_N)}} \leq \max_{f\in\mathcal{S}_i'}\frac{\langle \mathbf{G}^\top\mathbf{G}R_N f, R_N f\rangle_{L^2(\mu_N)}}{\|R_N f\|^2_{L^2(\mu_N)}},$$

we have the following:

$$\max_{f\in\mathcal{S}_i'}\langle \Delta_M f, f\rangle_{L^2(M)} \leq \max_{f\in\mathcal{S}_i'}\frac{\langle \mathbf{G}^\top\mathbf{G}R_N f, R_N f\rangle_{L^2(\mu_N)}}{\|R_N f\|^2_{L^2(\mu_N)}} + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

But we assumed that $\mathcal{S}_i'$ is the exact subspace on which $\max_{f\in\mathcal{S}_i}\frac{\langle \mathbf{G}^\top\mathbf{G}R_N f, R_N f\rangle_{L^2(\mu_N)}}{\|R_N f\|^2_{L^2(\mu_N)}}$ achieves its minimum. Hence,

$$\max_{f\in\mathcal{S}_i'}\langle \Delta_M f, f\rangle_{L^2(M)} \leq \hat{\lambda}_i + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

But the left-hand-side certainly bounds from above by the minimum of $\max_{f\in\mathcal{S}_i}\langle \Delta_M f, f\rangle_{L^2(M)}$ over all i-dimensional smooth subspaces $\mathcal{S}_i$. Hence,

$$\lambda_i \leq \hat{\lambda}_i + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

The same argument yields that $\hat{\lambda}_i \leq \lambda_i + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right)$, with probability higher than $1 - \frac{6}{N}$. This completes the proof. ■

The first error term in (32) can be seen as coming from discretizing a continuous operator, while the second error term in (32) comes from the fact that continuous estimators in our setting differ from the true Laplace-Beltrami by pre-composing with interpolation. The convergence holds true for eigenvectors, though in this case, the constants involved depend heavily on the multiplicity of the eigenvalues.

The convergence of eigenvector result is stated as follows.

Theorem 4.2 *Let $\epsilon_{\lambda_\ell} := |\lambda_\ell - \hat{\lambda}_\ell|$ denote the error in approximating the ℓ-th distinct eigenvalue, λ_ℓ, as defined in Theorem 4.1. Let Assumption 4.1 be valid. For any ℓ, there is a constant c_ℓ such that whenever $\epsilon_{\lambda_{\ell-1}}, \epsilon_{\lambda_{\ell+1}} < c_\ell$, then with probability higher than $1 - \left(\frac{2m^2+5m+24}{N}\right)$, where m is the geometric multiplicity of eigenvalue λ_ℓ, we have the following situation: for any normalized eigenvector $\mathbf{u}$ of $\mathbf{G}^\top\mathbf{G}$ with eigenvalue $\hat{\lambda}_\ell$, there is a normalized eigenfunction f of Δ_M with eigenvalue λ_ℓ such that*

$$\|R_N f - \mathbf{u}\|_{L^2(\mu_N)} = O\left(N^{-\frac{1}{4}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{4d}}\right).$$

As the proof is more technical, we present it in Appendix C. It is important to note that the results of this section use analytic $\mathbf{P}$. This allows us to conclude that, depending on the smoothness of the kernel, any error slower than the Monte-Carlo convergence rate observed numerically is introduced through the approximation of $\mathbf{P}$, as discussed in the previous section.

4.2 Spectral Convergence for the Bochner Laplacian

The Bochner Laplacian on vector fields is defined in such a way that makes the theoretical discussion in this setting almost identical to that of the Laplace-Beltrami operator. Hence, we relegate the proofs of the results below to Appendix D. We emphasize that the results below will also rely on the Assumption 4.1 which allows us to have a stable interpolator of smooth vector fields. In particular, as a corollary to Lemma B.5, the Assumption 4.1 gives $\|I_{\phi_s} u\|^2_{W^{2,\infty}(\mathfrak{X}(M))} = O(1)$ for any vector field $u \in \mathfrak{X}(M)$ whose components are $C^\infty(M)$ functions. Details regarding the previous statement are found in Appendix B.

The main results for the Bochner Laplacian are as follows. First, we have the following eigenvalue convergence result.

Theorem 4.3 *(convergence of eigenvalues: symmetric formulation) Let λ_i denote the i-th eigenvalue of Δ_B, enumerated $\lambda_1 \leq \lambda_2 \leq \ldots$. Let Assumption 4.1 be valid. For some fixed $i \in \mathbb{N}$, there exists an eigenvalue $\hat{\lambda}_i$ of $\mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes}$ such that*

$$\left|\lambda_i - \hat{\lambda}_i\right| \leq O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right),$$

with probability greater than $1 - \left(\frac{4n+4}{N}\right)$.

The same rate holds true for convergence of eigenvectors, just as in the Laplace-Beltrami case. In fact, the proof of convergence of eigenvectors follows the same argument in Section C.3. Before stating the main result, we have the following definition.

Definition 12 *Given two vectors $\mathbf{U}, \mathbf{V} \in \mathbb{R}^{nN}$ representing restrictions of vector fields, define discrete inner product $L^2(\mu_{N,n})$ in the following way:*

$$\langle \mathbf{U}, \mathbf{V} \rangle_{L^2(\mu_{N,n})} = \frac{1}{N} \sum_{j=1}^{n} \mathbf{U}^j \cdot \mathbf{V}^j,$$

where $\mathbf{U}^j \in \mathbb{R}^N$, $\mathbf{U} = ((\mathbf{U}^1)^\top, \ldots, (\mathbf{U}^n)^\top)^\top$, and similarly for $\mathbf{V}^j$ and $\mathbf{V}$.

With this definition, we can formally state the convergence of eigenvectors result for the Bochner Laplacian.

Theorem 4.4 *Let $\epsilon_{\lambda_\ell} := |\lambda_\ell - \hat{\lambda}_\ell|$ denote the error in approximating the ℓ-th distinct eigenvalue, following the notation in Theorem 4.3. Let the Assumption 4.1 be valid. For any ℓ, assume that there is a constant c_ℓ such that if $\epsilon_{\lambda_{\ell-1}}, \epsilon_{\lambda_{\ell+1}} < c_\ell$, then with probability higher than $1 - \left(\frac{2m^2+2m+3nm+8n+8}{N}\right)$, we have the following situation: for any normalized eigenvector $\mathbf{U}$ of $\mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes}$ with eigenvalue $\hat{\lambda}_\ell$, there is a normalized eigenvector field v of Δ_B with eigenvalue λ_ℓ such that*

$$\|R_N v - \mathbf{U}\|_{L^2(\mu_{N,n})} = O\left(N^{-\frac{1}{4}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{4d}}\right),$$

where m is the geometric multiplicity of eigenvalue λ_ℓ.

Proofs of the results for the Bochner Laplacian can be found in Appendix D.

5. Numerical Study of Eigenvalue Problems

In this section, we first discuss two examples of eigenvalue problems of functions defined on simple manifolds: one being a 2D generalized torus embedded in $\mathbb{R}^{21}$ and the other 4D flat torus embedded in $\mathbb{R}^{16}$. In these two examples, we will compare the results between the Non-symmetric and Symmetric RBFs, which we refer to as NRBF and SRBF respectively, using analytic $\mathbf{P}$ and the approximated $\hat{\mathbf{P}}$. In the first example, we further compare with diffusion maps (DM) algorithm, which is an important manifold learning algorithm that estimates eigen-solutions of the Laplace-Beltrami operator. When the manifold is unknown, as is often the case in practical applications, one does not have access to analytic $\mathbf{P}$. Hence, it is most reasonable to compare DM and RBF with $\hat{\mathbf{P}}$. While DM algorithm can be implemented with a sparse approximation via the K-Nearest Neighbors (KNN) algorithm, we would like to verify how SRBF $\hat{\mathbf{P}}$, which is a dense approximation, performs compared to DM for various degree of sparseness (including not using KNN as reported in Appendix E). Next, we discuss an example of eigenvalue problems of vector fields defined on a sphere. Numerically, we compare the results between the RBF method and the analytic truth. Since the size of our vector Laplacian approximation is $Nn \times Nn$, the current RBF methods are only numerically feasible for data sets with small ambient dimension n.

5.1 Numerical Setups

In the following, we introduce our numerical setups for NRBF, SRBF, and DM methods of finding approximate solutions to eigenvalue problems associated to Laplace-Beltrami or vector Laplacians on manifolds.

Parameter specification for RBF: For the implementation of RBF methods, there are two groups of kernels to be used. One group includes infinitely smooth RBFs, such as Gaussian (GA), multi-quadric (MQ), inverse multiquadric (IMQ), inverse quadratic (IQ), and Bessel (BE) (Fornberg et al., 2011; Flyer and Bengt, 2015; Fornberg and Lehto, 2011). The other group includes piecewise smooth RBFs, such as Polyharmonic spline (PHS), Wendland (WE), and Matérn (MA) (Flyer and Bengt, 2015; Fuselier and Wright, 2009; Wendland, 1995, 2005). In this work, we only apply GA and IQ kernels and test their numerical performances. To compute the interpolant matrix $\mathbf{\Phi}$ in (3), all points are connected and we did not use KNN truncations. The shape parameter s is manually tuned but fixed for different N when we examine the convergence of eigenmodes.

Despite not needing a structured mesh, many RBF techniques impose strict requirements for uniformity of the underlying data points. For uniformly distributed grid points, it often occurs that the operator approximation error decreases rapidly with the number of data N until the calculation breaks down due to the increasing ill-conditioning of the interpolant matrix $\mathbf{\Phi}$ defined in (3) (Tarwater, 1985; Schaback, 1995; Flyer and Bengt, 2015). In this numerical section, we consider data points randomly distributed on manifolds, which means that two neighboring points can be very close to each other. In this case, the interpolant matrix $\mathbf{\Phi}$ involved in most of the global RBF techniques tends to be ill-conditioned or even singular for sufficiently large N. In fact, one can show that with inverse quadratic kernel, the condition number of the matrix $\mathbf{\Phi}$ grows exponentially as a function of N. To resolve such an ill-conditioning issue, we apply the pseudo inversion instead of the direct inversion in approximating the interpolant matrix $\mathbf{\Phi}$. In our implementation, we will take the tolerance parameter of pseudo-inverse around $10^{-9} \sim 10^{-4}$.

If the parametrization or the level set representation of the manifold is known, we can apply the analytic tangential projection matrix $\mathbf{P}$ for constructing the RBF Laplacian matrix. If the parametrization is unknown, that is, only the point cloud is given, we can first learn $\hat{\mathbf{P}}$ using the 2nd-order SVD method and then construct the RBF Laplacian. Notice that we can also construct the Laplacian matrix using $\tilde{\mathbf{P}}$ estimated from the 1st-order SVD (not shown in this work). We found that the results of eigenvalues and eigenvectors using $\tilde{\mathbf{P}}$ are not as good as those using $\hat{\mathbf{P}}$ from our 2nd-order SVD. We also notice that the estimation of $\mathbf{P}$ and the construction of the Laplacian matrix can be performed separately using two sets of points. For example, one can use $10,000$ points to approximate $\mathbf{P}$ but use only 2500 points to construct the Laplacian matrix. This allows one to leverage large amounts of data in the estimation of $\mathbf{P}$, while too much data may not be computationally feasible with graph Laplacian-based approximators such as the diffusion maps algorithm.

For SRBF, the estimated sampling density is needed if the distribution of the data set is unknown. Note that for NRBF, the sampling density is not needed for constructing Laplacian. In our numerical experiments, we apply the MATLAB built-in kernel density estimations (KDEs) function `mvksdensity.m` for approximating the sampling density. We

also apply Silverman's rule of thumb (Silverman, 2018) for tuning the bandwidth parameter in the KDEs.

Eigenvalue problem solver for RBF: For NRBF, we apply the non-symmetric estimator in (9) for solving the eigenvalue problem. The NRBF eigenvectors might be complex-valued and are only linearly independent (i.e., they are not necessarily orthonormal). For SRBF, we apply the symmetric estimator in (10) for solving the generalized eigenvalue problem. When the sampling density q is unknown, we employ the symmetric formulation with the estimated sampling density $\tilde{q}(x)$ obtained from the KDE.

Since we used pseudo inversion to resolve the ill-conditioning issue of $\mathbf{\Phi}$, the resulting RBF Laplacian matrix will be of low rank, L, and will have many zero eigenvalues, depending on the choice of tolerance in the pseudo-inverse algorithm. Two issues naturally arise in this situation. First, it becomes difficult to compute the eigenspace corresponding to the zero eigenvalue(s), especially for the eigenvector-field problem. At this moment, we have not developed appropriate schemes to detect the existence of the nullspace and estimate the harmonic functions in this nullspace if it exists. Second, finding even the leading nonzero eigenvalues (that are close to zero) can be numerically expensive. Based on the rank of the RBF $\mathbf{\Phi}$, for symmetric approximation, one can use the ordered real-valued eigenvalues to attain the nontrivial L eigenvalues in descending (ascending) order when L is small (large). For the non-symmetric approximation, one can also employ a similar idea by sorting the magnitude of the eigenvalues (since the eigenvalues may be complex-valued). This naive method, however, can be very expensive when the number of data points N is large and when the rank of the matrix L is neither $O(10)$ nor close to N.

Comparison of eigenvectors for repeating eigenvalues: When an eigenvalue is non-simple, one needs to be careful in quantifying the errors of eigenvectors for these repeated eigenvalues since the set of true orthonormal eigenvectors is only unique up to a rotation matrix. To quantify the errors of eigenvectors, we apply the following Ordinary Least Square (OLS) method. Let $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_m)$ be the true eigenfunctions located at X corresponding to one repeated eigenvalue λ_i with multiplicity m, and let $\tilde{\mathbf{U}} = (\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_m)$ be their DM or RBF approximations. Assume that the linear regression model is written as $\mathbf{F} = \tilde{\mathbf{U}}\beta + \varepsilon$, where ε is an $N \times m$ matrix representing the errors and β is a $m \times m$ matrix representing the regression coefficients. The coefficients matrix β can be approximated using OLS by $\hat{\beta} = (\tilde{\mathbf{U}}^\top \tilde{\mathbf{U}})^{-1}(\tilde{\mathbf{U}}^\top \mathbf{F})$. The rotated DM or RBF eigenvectors can be written as a linear combination, $\hat{\mathbf{V}} = \tilde{\mathbf{U}}\hat{\beta}$, where these new $\hat{\mathbf{V}} = (\hat{\mathbf{v}}_1, \dots, \hat{\mathbf{v}}_m)$ are in $\mathrm{Span}\{\tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_m\}$. Finally, we can quantify the pointwise errors of eigenvectors between $\mathbf{f}_j$ and $\hat{\mathbf{v}}_j$ for each $j = 1, \dots, m$. For eigenvector fields, we can follow a similar outline to quantify the errors of vector fields. Incidentally, we mention that there are many ways to measure eigenvector errors since the approximation of the rotational coefficient matrix β is not unique. Here, we only provide a practical metric that we will use in our numerical examples below.

5.2 2D General Torus

In this section, we investigate the eigenmodes of the Laplace-Beltrami operator on a general torus. The parameterization of the general torus is given by

$$x = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^{n-2} \\ x^{n-1} \\ x^n \end{bmatrix} = \begin{bmatrix} (a+\cos\theta)\cos\phi \\ (a+\cos\theta)\sin\phi \\ \vdots \\ \frac{2}{n-1}(a+\cos\theta)\cos\frac{n-1}{2}\phi \\ \frac{2}{n-1}(a+\cos\theta)\sin\frac{n-1}{2}\phi \\ \sqrt{\sum_{i=1}^{(n-1)/2}\frac{1}{i^2}}\sin\theta \end{bmatrix}, \tag{33}$$

where the two intrinsic coordinates $0 \le \theta, \phi \le 2\pi$ and the radius $a = 2 > 1$. The Riemannian metric is

$$g = \begin{bmatrix} b_{\frac{n-1}{2}} & 0 \\ 0 & \frac{n-1}{2}(a+\cos\theta)^2 \end{bmatrix}, \tag{34}$$

where $b_{\frac{n-1}{2}} := \sum_{i=1}^{(n-1)/2}\frac{1}{i^2}$. We solve the following eigenvalue problem for Laplace-Beltrami operator:

$$\Delta_g\psi_k = \frac{-1}{(a+\cos\theta)}\left[\frac{\partial}{\partial\theta}\left((a+\cos\theta)\,\frac{1}{b_{\frac{n-1}{2}}}\frac{\partial\psi_k}{\partial\theta}\right) + \frac{\partial}{\partial\phi}\left(\frac{2}{n-1}\frac{1}{(a+\cos\theta)}\frac{\partial\psi_k}{\partial\phi}\right)\right] = \lambda_k\psi_k, \tag{35}$$

where λ_k and ψ_k are the eigenvalues and eigenfunctions, respectively. After separation of variables (that is, we set $\psi_k = \Phi_k(\phi)\,\Theta_k(\theta)$ and substitute ψ_k back into (35) to deduce the equations for Φ_k and Θ_k), we obtain:

$$\begin{aligned} \frac{\mathrm{d}^2\Phi_k}{\mathrm{d}\phi^2} + m_k^2\Phi_k &= 0, \\ \frac{\mathrm{d}}{\mathrm{d}\theta}\left((a+\cos\theta)\frac{\mathrm{d}\Theta_k}{\mathrm{d}\theta}\right) - \frac{b_{\frac{n-1}{2}}m_k^2}{\frac{n-1}{2}(a+\cos\theta)}\Theta_k &= -b_{\frac{n-1}{2}}(a+\cos\theta)\,\lambda_k\Theta_k. \end{aligned}$$

The eigenvalues to the first equation are $m_k^2 = k^2$ with $k = 0, 1, 2, \ldots$ and the associated eigenvectors are $\Phi_k = \{1, \sin k\phi, \cos k\phi\}$. The second eigenvalue problem can be written in the Sturm–Liouville form and then numerically solved on a fine uniform grid with N_θ points $\{\theta_j = \frac{2\pi j}{N_\theta}\}_{j=0}^{N_\theta-1}$ (Pryce, 1993). The eigenvalues λ_k associated with the eigenfunctions ψ_k obtained above are referred to as the true semi-analytic solutions to the eigenvalue problem (35).

In our numerical implementation, data points are randomly sampled from the general torus with uniform distribution in intrinsic coordinates. For this example, we also show results based on the diffusion maps algorithm with a Gaussian kernel, $K_\epsilon(x,y) = \exp(-\frac{\|x-y\|^2}{4\epsilon})$, where ϵ denotes the bandwidth parameter to be specified. The sampling density is estimated using the KDE estimator proposed by Loftsgaarden and Quesenberry (1965). For an efficient implementation, we use the K-nearest neighbors algorithm to avoid computing the graph affinity between pair of points that are sufficiently far away. In Appendix E, additional numerical results for other choices of K (including $K = N$ or no KNN

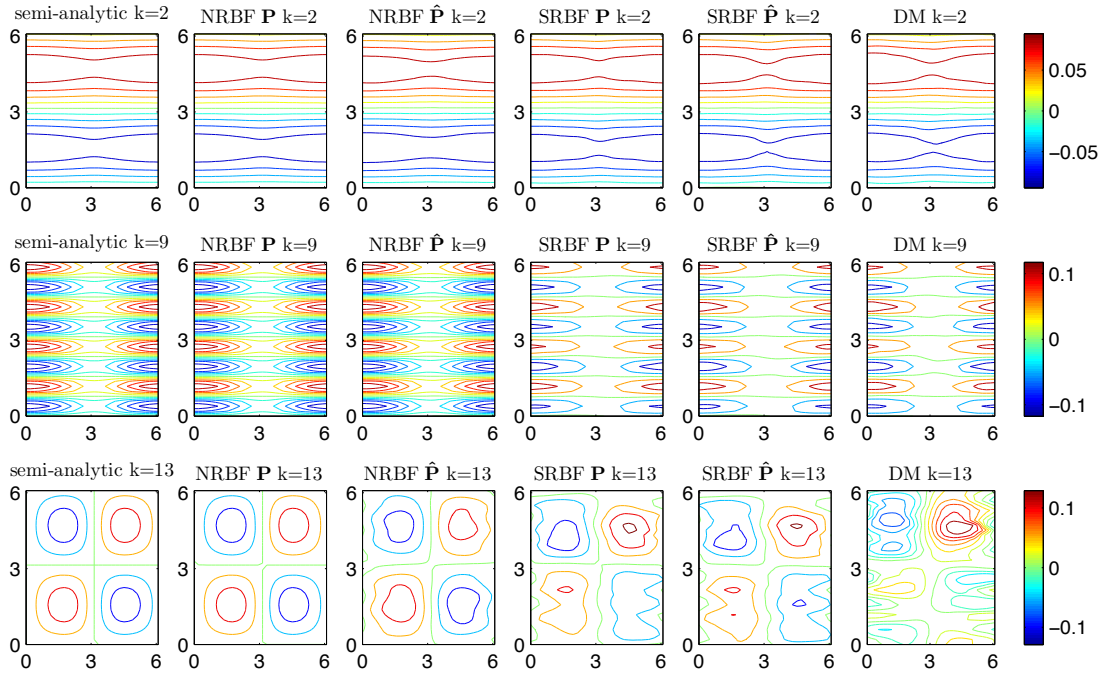

Figure 4: **2D general torus in $\mathbb{R}^{21}$.** Comparison of eigenfunctions of Laplace -Beltrami for $k = 2, 9, 13$ among NRBF using $\mathbf{P}$ and $\hat{\mathbf{P}}$, SRBF using $\mathbf{P}$ and $\hat{\mathbf{P}}$, and DM using $K = 100$. For NRBF, IQ kernel with $s = 0.5$ is used, and for SRBF, IQ kernel with $s = 0.1$ is used. For DM with $K = 100$, the auto-tuned algorithm for the bandwidth ϵ is used. The horizontal and vertical axes correspond to θ and ϕ, respectively. $N = 2500$ randomly distributed data points on the manifold are used for computing the eigenvalue problem. The eigenvectors are then generalized onto the 32×32 well-sampled grid points $\{\theta_i, \phi_j\} = \{\frac{2\pi i}{32}, \frac{2\pi j}{32}\}_{i,j=0}^{31}$ for plotting.

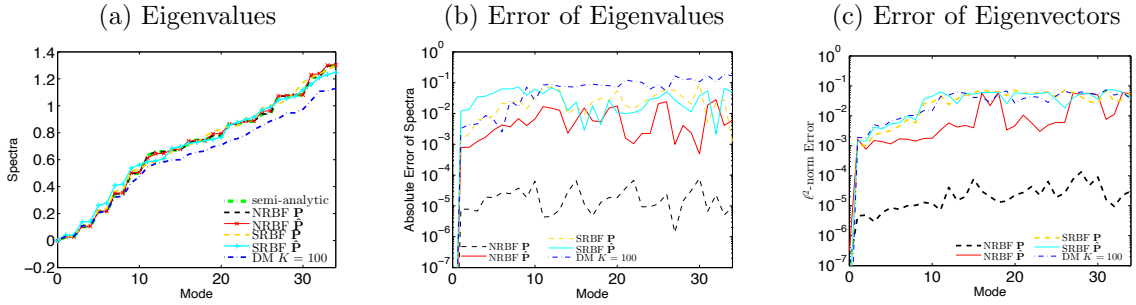

Figure 5: **2D general torus in $\mathbb{R}^{21}$.** Comparison of (a) eigenvalues, (b) error of eigenvalues, and (c) error of eigenvectors, among NRBF, SRBF, and DM. For NRBF, IQ kernel with $s = 0.5$ is used, and for SRBF, IQ kernel with $s = 0.1$ is used. For SRBF, the KDE estimated $\tilde{q}$ is used. Randomly distributed $N = 2500$ data points on the manifold are used for solving the eigenvalue problem.

being used) are reported. It is worth noting that the additional results in Appendix E suggest that the accuracy of the estimation of eigenvectors does not improve for any values of K that we tested. To further check the numerical optimality of the choice of bandwidth parameter ϵ, we also empirically check whether an improved estimate can be attained by varying ϵ around the auto-tuned value. We found that the auto-tuned method is more effective when K is relatively small, which motivates the use of small K in verifying the numerical convergence. Figure 6 shows the sensitivity of the estimates as ϵ is varied for fixed $K = 100$ and $N = 2500$. For the convergence result, we choose $K \sim \sqrt{N}$ following the theoretical guideline in Calder and Trillos (2022) that guarantees a convergence rate of $\epsilon = O\left((\frac{K}{N})^{2/d}\right)$. Once K is fixed, the parameter ϵ is selected based on the auto-tuned algorithm introduced in Coifman et al. (2008). Specifically, we set $K = 60, 100, 144, 200$ for $N = 1024, 2500, 5000, 10000$, respectively.

To apply NRBF, we use IQ kernel with $s = 0.5$. To apply SRBF, we use IQ kernel with $s = 0.1$. Figure 4 shows the comparison of eigenfunctions for modes $k = 2, 9, 13$ among the semi-analytic truth, NRBF with $\mathbf{P}$ and $\hat{\mathbf{P}}$, SRBF with $\mathbf{P}$ and $\hat{\mathbf{P}}$, and DM. One can see from the first row of Fig. 4 that when $k = 2$ is very small, all the methods can provide excellent approximations of eigenfunctions. For larger k, such as 9 and 13, NRBF methods with $\mathbf{P}$ or $\hat{\mathbf{P}}$ provide more accurate approximations compared to SRBF with $\mathbf{P}$ and $\hat{\mathbf{P}}$ and DM. In fact, the eigenvectors obtained from NRBF with $\mathbf{P}$ are accurate and very smooth as seen from the second column of Fig. 4. On the other hand, SRBF with $\hat{\mathbf{P}}$ does not produce eigenvectors that are qualitatively much better than those of DM.

Figure 5 further quantifies the errors of eigenvalues and eigenfunctions for all the methods. One can see that NRBF with $\mathbf{P}$ performs much better than all other methods on this 2D manifold example. When the manifold is unknown, one can diagnose the manifold learning capabilities of the symmetric and non-symmetric RBF using $\hat{\mathbf{P}}$ compared to DM. One can see that NRBF with $\hat{\mathbf{P}}$ (red curve) performs better than the other two methods. One can also see that DM (blue curve) performs slightly better than SRBF with $\hat{\mathbf{P}}$ (cyan curve) in estimating the leading eigenvalues but somewhat worst in estimating eigenvalues

corresponding to the higher modes (Fig. 5(b)). In terms of the estimation of eigenvectors, they are comparable (blue and cyan curves Fig. 5(c)). Additionally, SRBF with $\mathbf{P}$ (yellow curve) and SRBF with $\hat{\mathbf{P}}$ (cyan curve) produce comparable accuracies in terms of the eigenvector estimation. This result, where no advantage is observed using the analytic $\mathbf{P}$ over the approximated $\hat{\mathbf{P}}$, is consistent with the theory which suggests that the error bound is dominated by the Monte-Carlo error rate provided smooth enough kernels are used.

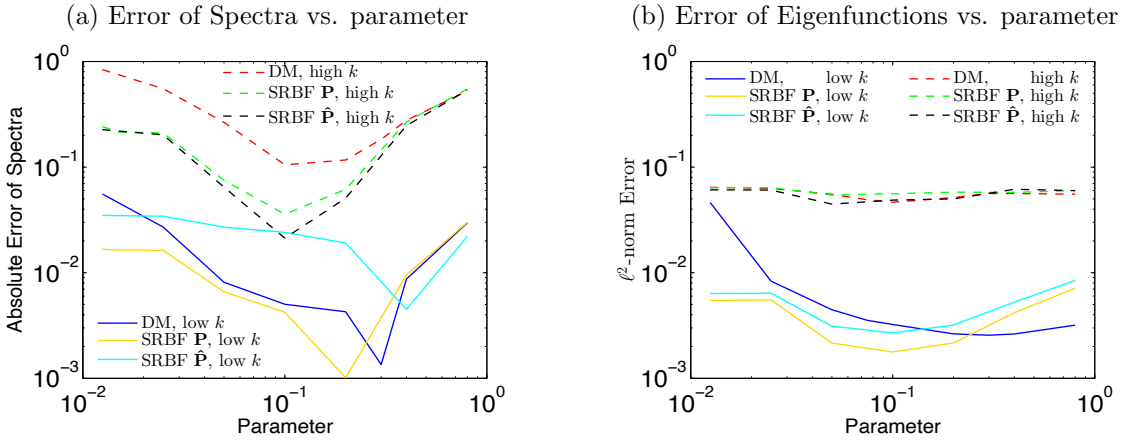

Figure 6: **2D general torus in $\mathbb{R}^{21}$.** Sensitivity of (a) eigenvalues and (b) eigenfunctions with respect to the parameter (bandwidth for DM and shape parameter for SRBF using true $\mathbf{P}$ and SRBF using estimated $\hat{\mathbf{P}}$). The blue, yellow and cyan curves denote the average error of eigenvalues or eigenvectors over modes 2-5 (low k), while the red, green, and black curves denote the average error of eigenvalues or eigenvectors over modes 21-30 (high k). For SRBF, IQ kernel is used and shape parameter s ranges from 10^{-2} to 10^{0}. For DM, $K = 100$ nearest neighbors are used and bandwidth ϵ ranges from 10^{-2} to 10^{0}. In this experiment, we fixed the $N = 2500$ data points which are randomly distributed on the general torus with uniform distribution in the intrinsic coordinates.

In the previous two figures, we showed the SRBF estimates corresponding to a specific choice of $s = 0.1$. Now let us check the robustness of the method with respect to other choices of the shape parameter. In Fig. 6, we show the errors in the estimation of eigenvalues and eigenvectors. Specifically, we report the average errors of low modes (between modes 2-5) for DM (blue), SRBF with $\mathbf{P}$ (yellow), and SRBF with $\hat{\mathbf{P}}$ (cyan). For these low modes, notice that with the optimal shape parameter, $s = 0.4$, the eigenvalue estimates from SRBF with $\hat{\mathbf{P}}$ are slightly less accurate than the diffusion maps. For this shape parameter value, the SRBF with $\hat{\mathbf{P}}$ produces an even more accurate estimation of the leading eigenvalues. However, the accuracy of the SRBF eigenvectors decreases slightly under this parameter value. We also report the average errors of high modes (between modes 21-30) for DM (red), SRBF with $\mathbf{P}$ (green), and SRBF with $\hat{\mathbf{P}}$ (black). For these high modes, both SRBFs are uniformly more accurate than DM in the estimation of eigenvalues, but the accuracies of the estimation of eigenvectors are comparable. More comparisons can be found in Appendix

E. Overall, SRBF and DM show comparable results for the eigenvalue problem. For both SRBF and DM, the eigenvalue estimates are sensitive to the choice of the parameter while the eigenvector estimates are not. Based on numerical experiments with a wide range of parameters, we empirically found that DM has a slight advantage in the estimation of leading spectrum while the SRBF has a slight advantage in the estimation of non-leading spectrum.

In Fig. 7, we examine the convergence of eigenvalues and eigenvectors for NRBF with $\hat{\mathbf{P}}$ in the case of unknown manifold. For NRBF with $\hat{\mathbf{P}}$, since the eigenvalues are complex value, Fig. 7(a) displays all the eigenvalues on a complex plane. Here are four observations from our numerical results:

1. When N increases, more eigenvalues with large magnitudes flow to the "tail" as a cluster packet.

2. The magnitude of imaginary parts decays as N increases.

3. For the leading modes with small magnitudes, NRBF eigenvalues converge fast to the real axis and converge to the true spectra at the same time.

4. It appears that all of the eigenvalues lie in the right half plane with positive real parts for this 2D manifold as long as N is large enough ($N > 1000$). Notice that this result is consistent with the previous result reported in (Fuselier and Wright, 2013). In that paper, the authors considered the negative definite Laplace-Beltrami operator and numerically observed that all eigenvalues are in the left half plane with negative real parts for many complicated 2D manifolds for large enough data.

In Fig. 7(b)-(d), we would like to verify that the convergence rate of the NRBF is dominated by the error rate in the estimation of $\mathbf{P}$. In all numerical experiments in these panels, we solve eigenvalue problems of discrete approximation with a fixed 2500 data points as in previous examples. Here, we verify the error rate in terms of the number of points used to construct $\hat{\mathbf{P}}$, which we denote as N_p. For the 2D manifolds, we found that the convergence rate (panel (b)) for the leading 12 modes decay with the rate N_p^{-1}, which is consistent with the theoretical rate deduced in Theorem 3.2 and the discussion in Remark 10. This rate is faster than the Monte-Carlo rate even for randomly distributed data. In panels (c)-(d), we report the detailed errors in the eigenvalue and eigenvector estimation for each mode. This result suggests that if N is large enough (as we point out in bullet point 4 above), one can attain accurate estimation by improving the accuracy of the estimation of $\hat{\mathbf{P}}$ by increasing the sample size, N_p.

In Fig. 8, we examine the convergence of eigenvalues and eigenvectors for DM and SRBF with $\hat{\mathbf{P}}$ in the case of the unknown manifold. Previously in Figs. 4-6, we showed result with $N = 2500$ fixed, now we examine the convergence rate as N increases. For DM and SRBF with $\hat{\mathbf{P}}$, the Laplacian matrix is always symmetric positive definite, so that their eigenvalues and eigenvectors must be real-valued and their eigenvalues must be positive. Figures 8(c)-(e) display the errors of eigenvalues and eigenvectors for DM and SRBF for $N = 1024, 2500, 5000, 10000$. For robustness, we report estimates from 16 experiments, where each estimate corresponds to independent randomly drawn data (see thin lines). The thicker lines in each panel correspond to the average of these experiments. One can observe

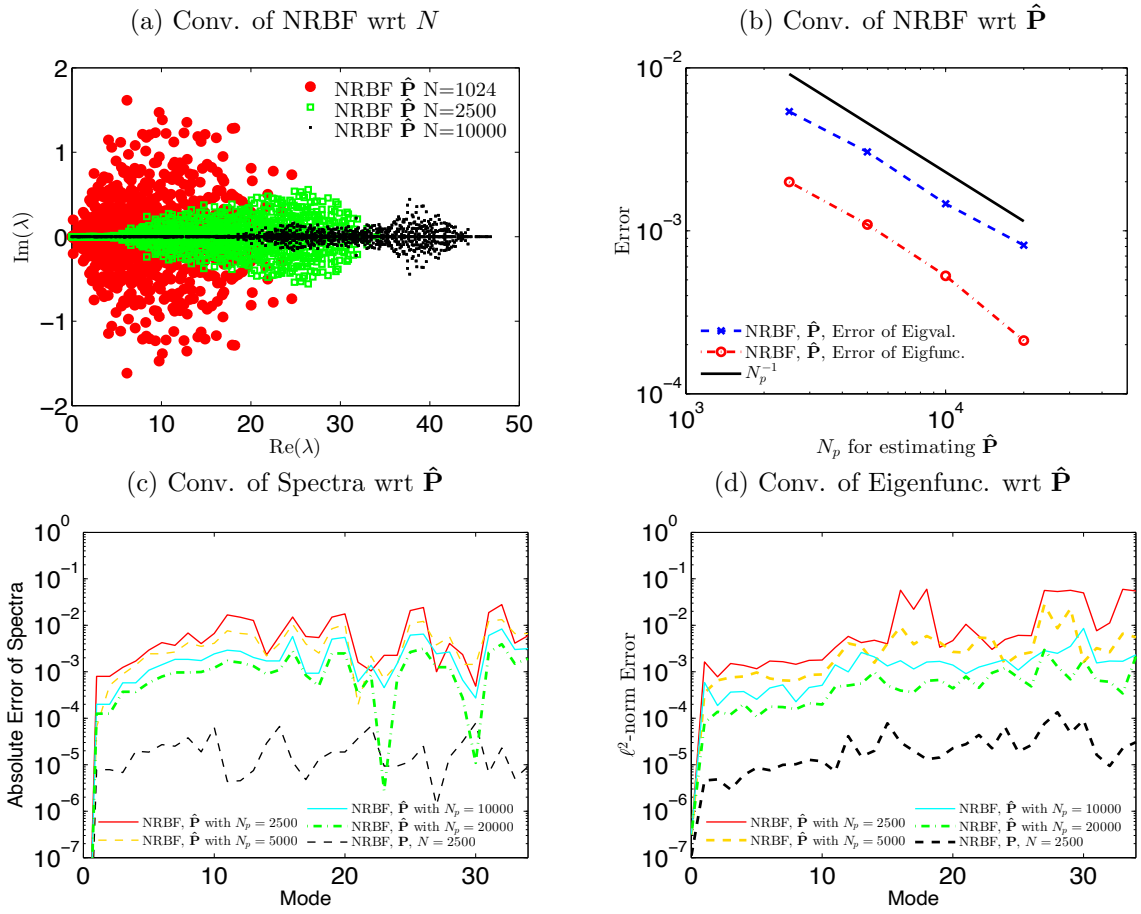

Figure 7: **2D general torus in $\mathbb{R}^{21}$.** Convergence of eigenvalues and eigenvectors for the NRBF method. Here N denotes the number of points for solving the eigenvalue problem, while N_p denotes the number of points for estimating $\hat{\mathbf{P}}$. (a) Convergence of eigenvalues with respect to N for NRBF using $\hat{\mathbf{P}}$. In panel (a), N data points are used for both approximating $\hat{\mathbf{P}}$ and evaluating NRBF matrices. In panels (c) and (d), shown is the convergence of NRBF eigenvalues and NRBF eigenfunctions with respect to $\hat{\mathbf{P}}$ for varying values of N_p, but with fixed $N = 2500$ data points used for solving the NRBF eigenvalue problem. (b) For each N_p, plotted are the averages of errors of eigenvalues or eigenfunctions for the leading 12 modes (2nd-13rd modes). The convergence rate of eigenvalues and eigenfunctions are both N_p^{-1}. IQ kernel with $s = 0.5$ was fixed for all cases. The data points are randomly distributed on the general torus according to a uniform distribution in the intrinsic coordinates.

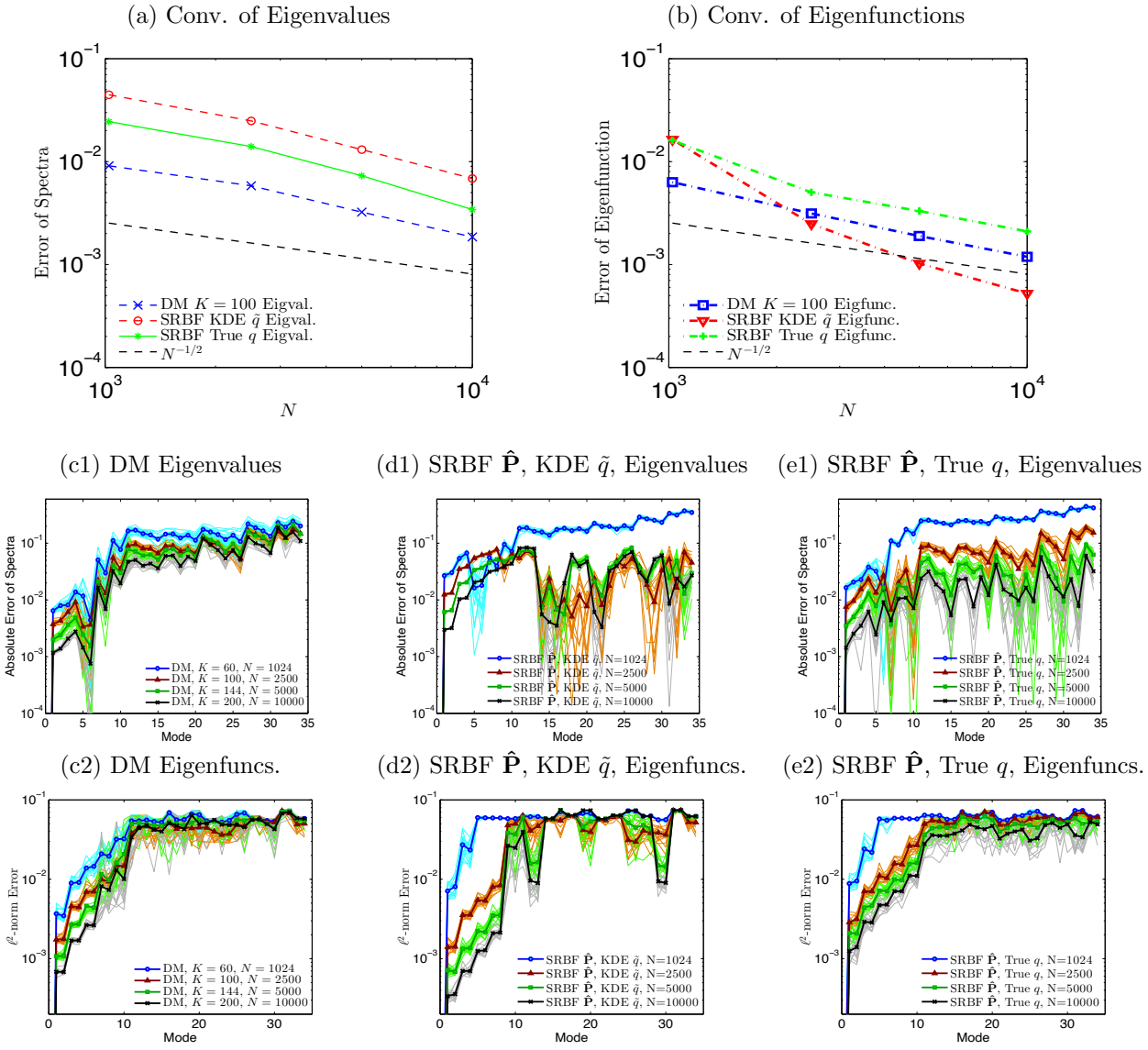

Figure 8: **2D general torus in $\mathbb{R}^{21}$.** Convergence of (a) eigenvalues and (b) eigenfunctions for DM and SRBF using $\hat{\mathbf{P}}$. For each N, the average error of eigenvalues or eigenvectors over the leading four modes (2nd-5th modes) are plotted. For SRBF, IQ kernel with $s = 0.1$ was fixed for each N. Comparison of errors of eigenvalues for (c1) DM, (d1) SRBF using estimated sampling density $\tilde{q}$, and (e1) SRBF using the true sampling density q are shown. Plotted in (c2)-(e2) are the corresponding comparison of errors of eigenvectors. For each N, 16 independent trials are run and depicted by light color. For each N, the average of all 16 trials are depicted by dark color. For each trial, randomly distributed data points on the manifold are used for computation.

from Figs. 8(a) and (b) that the errors of the leading four modes for both methods are comparable and decay on the order of $N^{-1/2}$. This rate is consistent with Theorems 4.1 and 4.2 with smooth kernels for the SRBF. On the other hand, this rate is faster than the theoretical convergence rate predicted in (Calder and Trillos, 2022), $N^{-\frac{1}{d+4}}$. One can also see that SRBF using KDE $\tilde{q}$ (red dash-dotted curve) provides a slightly faster convergence rate for the error of eigenfunction compare to those of SRBF with analytic q, which is counterintuitive. In the next example, we will show the opposite result, which is more intuitive. Lastly, more detailed errors of the eigenvalues and eigenvectors estimates for each leading mode are reported in Figs. 8(c)-(e), from which one can see the convergence for each leading mode for both SRBF and DM. One can also see the slight advantage of SRBF over DM in the estimation of non-leading eigenvalues while the slight advantage of DM over SRBF in the estimation of leading eigenvalues (consistent to the result with a fixed $N = 2500$ shown in Figs. 5 and 6).

5.3 4D Flat Torus

We consider a $d-$dimensional manifold embedded in $\mathbb{R}^{2md}$ with the following parameterization,

$$x = \frac{1}{\sqrt{1+\cdots+m^2}} \begin{pmatrix} \cos(t_1), & \sin(t_1), & \cdots & \cos(mt_1), & \sin(mt_1), \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ \cos(t_d), & \sin(t_d), & \cdots & \cos(mt_d), & \sin(mt_d) \end{pmatrix},$$

with $0 \le t_1 \le 2\pi, \cdots, 0 \le t_d \le 2\pi$. The Riemannian metric is given by a $d \times d$ identity matrix $\mathbf{I}_d$. The Laplace-Beltrami operator can be computed as $\Delta_g u = -\sum_{i=1}^{d} \frac{\partial^2 u}{\partial t_i^2}$. For each dimension t_i, the eigenvalues of operator $-\frac{\partial^2}{\partial t_i^2}$ are $\{0, 1, 1, 4, 4, \ldots, k^2, k^2, \ldots\}$. The exact spectrum and multiplicities of the Laplace-Beltrami operator on the general flat torus depends on the intrinsic dimension d of the manifold. In this section, we study the eigenmodes of Laplace-Beltrami operator for a flat torus with dimension $d = 4$ and ambient space $\mathbb{R}^{2md} = \mathbb{R}^{16}$. In this case, the spectra of the flat torus can be calculated as

$$\begin{array}{lccccc} \text{spectra} & 0 & 1 & 2 & 3 & 4 \\ \text{mode } k & 1 & 2 \sim 9 & 10 \sim 33 & 34 \sim 65 & 66 \sim 89 \end{array}, \tag{36}$$

where the eigenvalues $0, 1, 2, 3, 4$ have multiplicities of $1, 8, 24, 32, 24$, respectively. Our motivation here is to investigate if RBF methods suffer from curse of dimensionality when solving eigenvalue problems.

Numerically, data points are randomly distributed on the flat torus with uniform distribution. To apply NRBF, we use GA kernel with $s = 0.5$. To apply SRBF, we use IQ kernel with $s = 0.3$. Figure 9 shows the results of eigenvalues and eigenfunctions for NRBF with $\mathbf{P}$ and $\hat{\mathbf{P}}$, and SRBF with $\hat{\mathbf{P}}$. One can see from Figs. 9(b) and (c) that when $N = 30000$ is large enough, NRBF with $\mathbf{P}$ (black dashed curve) performs much better than the other methods. One can also see that NRBF with $\hat{\mathbf{P}}$ (red curve) and SRBF with $\hat{\mathbf{P}}$ (cyan curve) are comparable when the manifold is assumed to be unknown.

For the NRBF method using $\hat{\mathbf{P}}$, all the four observations in Example 5.2 persist. However, we should point out that for the last observation (for all numerical eigenvalues in the

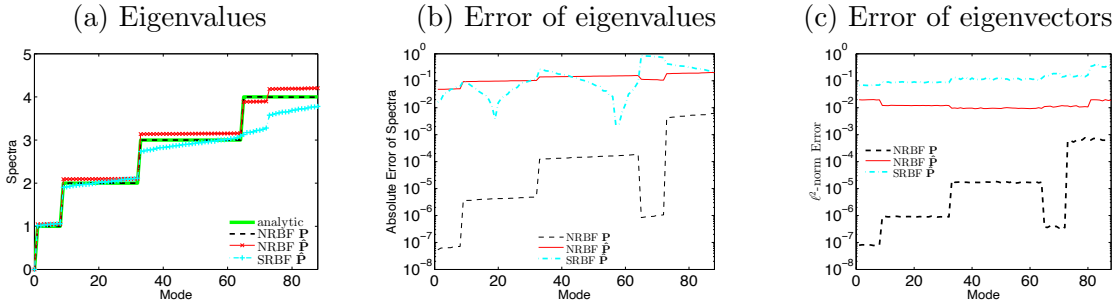

Figure 9: **4D flat torus in $\mathbb{R}^{16}$.** Comparison of (a) eigenvalues, (b) error of eigenvalues, and (c) error of eigenvectors, among NRBF and SRBF. For NRBF, GA kernel with $s = 0.5$ is used, and for SRBF, IQ kernel with $s = 0.3$ is used. The $N = 30,000$ data points are randomly distributed on the flat torus.

right half plane with positive real parts) to hold, the number of points N has to be around 20000 as can be seen from Fig. 10(a). When $N = 10000$ is not large enough, there are many irrelevant eigenvalues with negative real parts (blue crosses), which is a sign of spectral pollution. Moreover, the leading eigenvalues (yellow dots) are not close to the truth (red circles) [the inset of Fig. 10(a)]. When N increases to 20000 or 30000, the irrelevant eigenvalues do not completely disappear (they appear near eigenvalues with larger real components) even though NRBF eigenvalues (magenta and green dots) lie in the right half plane. Here, the leading NRBF eigenvalues (dark green dots) approximate the truth (red circles) more accurately [the inset of Fig. 10(a)].

For NRBF, we also repeat the experiment with $N = 10000$ data points using the analytic $\mathbf{P}$. We found that the profile of blue crosses in Fig. 10(a) still persists even if the analytic $\mathbf{P}$ was used to replace the approximated $\hat{\mathbf{P}}$ [not shown]. The only slight difference was that the errors of the leading modes became smaller if $\hat{\mathbf{P}}$ was replaced with $\mathbf{P}$ [red and black curves in Fig. 9(b)]. This numerical result suggests that these irrelevant eigenvalues (blue crosses in Fig. 10(a)) are due to not enough data points rather than the inaccuracy of $\hat{\mathbf{P}}$.

In contrast, when ($N = 30000$) data points are used, we already observed in Fig. 9(b) and (c) that NRBF with $\mathbf{P}$ (black dashed curve) performs much better than NRBF with $\hat{\mathbf{P}}$ (red solid curve). One can expect the errors of the NRBF with $\hat{\mathbf{P}}$ (red curve in Fig. 9(b)) to decay to the errors of the NRBF with $\mathbf{P}$ (black curve in Fig. 9(b)) as the data points used to learn the tangential projection matrices $\hat{\mathbf{P}}$ is increased (beyond $30,000$ points) while the same fixed $N = 30000$ data points are used to construct the Laplacian matrix for the eigenvalue problem. This scenario is practically useful since the approximation of $\hat{\mathbf{P}}$ is computationally cheap even with a large number of data, whereas solving the eigenvalue problem of a dense, non-symmetric RBF matrix is very expensive when N is large. In Figs. 10(b) and (c) for NRBF, we demonstrate the improvement with this estimation scenario, especially with respect to the number of points N_p used to construct $\hat{\mathbf{P}}$. We found that the rate is $N_p^{-1/2}$ which is consistent with our expectation as noted in Remark 10 for 4-dimensional examples.

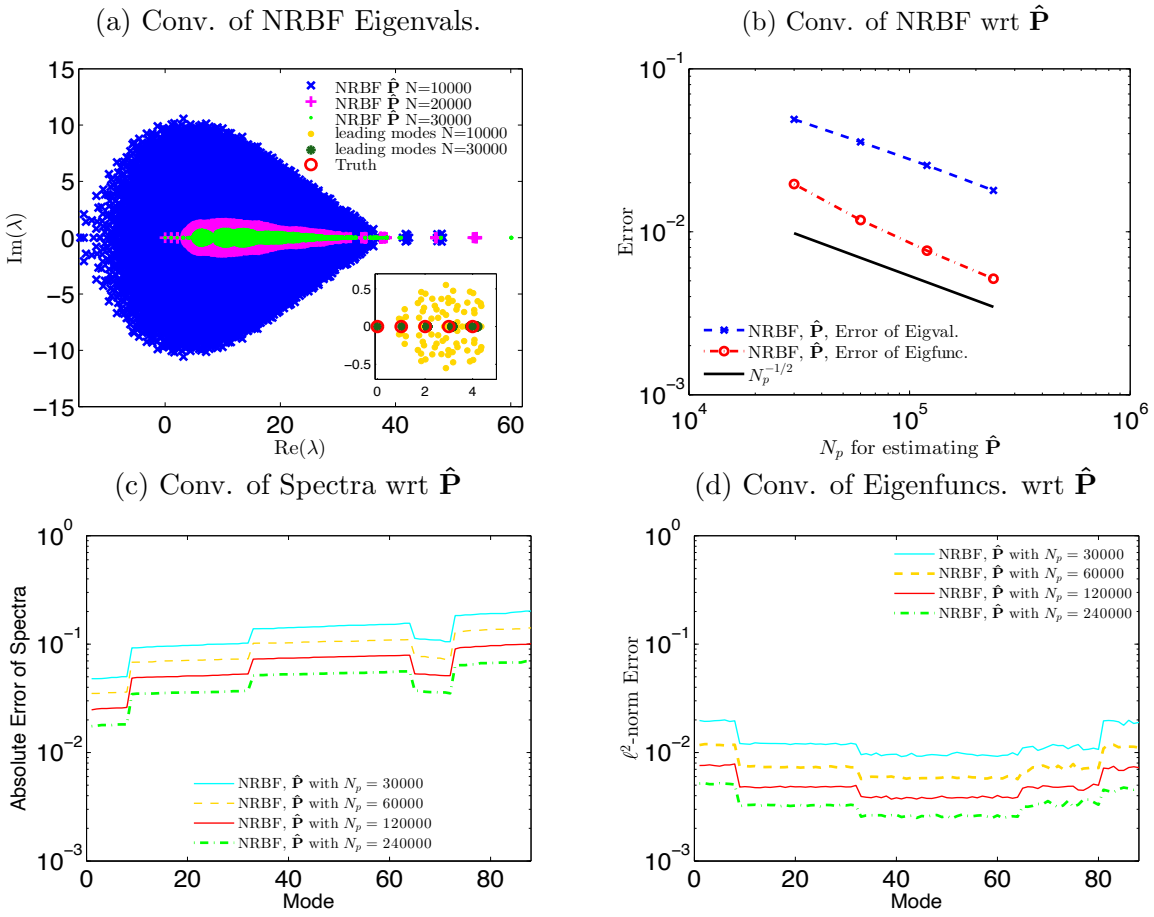

Figure 10: **4D flat torus in $\mathbb{R}^{16}$.** (a) Convergence of eigenvalues with respect to N for NRBF using $\hat{\mathbf{P}}$. For panel (a), N data points are used for both approximating $\hat{\mathbf{P}}$ and evaluating NRBF matrices. The leading modes in legend are referred to as the 89 numerical NRBF eigenvalues closest to the truth listed in equation (36). Convergence of (c) NRBF eigenvalues and (d) NRBF eigenfunctions with respect to $\hat{\mathbf{P}}$ estimated from different N_p number of points. For panels (c) and (d), the same $N = 30,000$ data points are used for evaluating NRBF matrices but different N_p data points are used for approximating $\hat{\mathbf{P}}$ at those $30,000$ data points. In panel (b), plotted is the convergence rate of $O(N_p^{-1/2})$ for the leading 8 modes (2nd-9th modes). GA kernel with $s = 0.5$ was fixed for all N. The data points are randomly distributed on the flat torus.

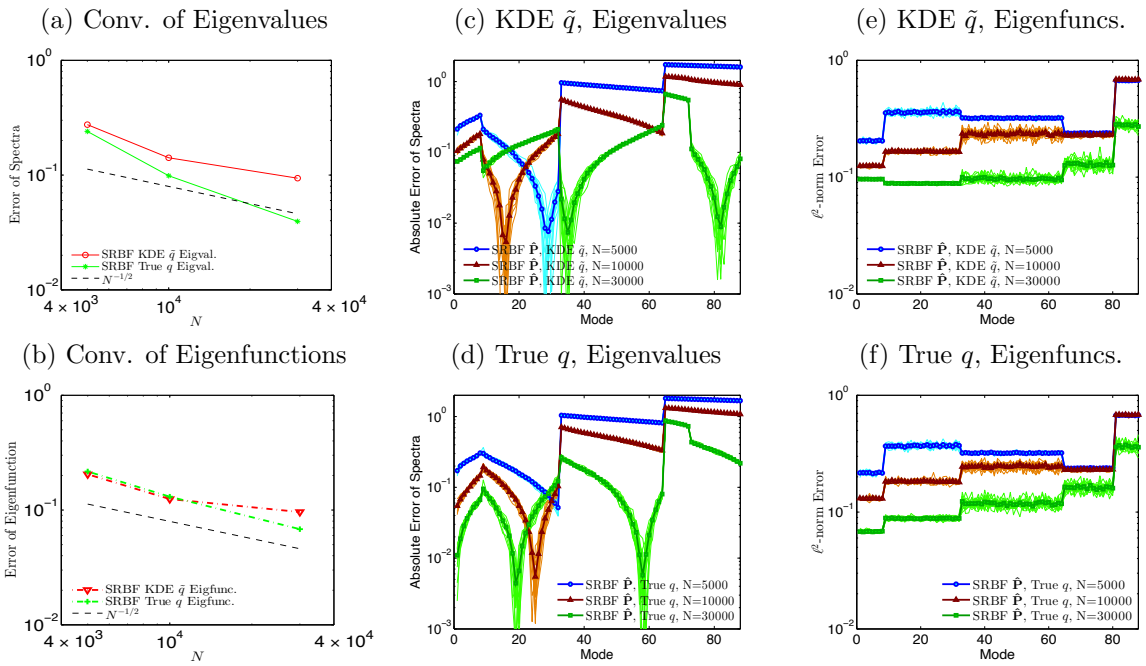

Figure 11: **4D flat torus in $\mathbb{R}^{16}$.** Convergence of (a) eigenvalues and (b) eigenvectors for SRBF using $\hat{\mathbf{P}}$ averaged over the first eight modes (2nd-9th modes). In (c), and (d) we plot the error of eigenvalues for each mode, while in (e), and (f) we plot the error of eigenvectors. For each panel, IQ kernel with $s = 0.3$ was used. Panels (c) and (d) show the errors of eigenvalues using the estimated sampling density $\tilde{q}$ and true sampling density q, respectively. Panels (e) and (f) show the corresponding errors of eigenvectors. In each of these panels, 16 independent trials are run and depicted by light color. The average of all 16 trials are depicted by dark color. The data points used for computation are randomly distributed on the manifold.

We also inspect the convergence of NRBF when the dimension of the manifold varies. In particular, we compare the numerical result with a 3D flat torus example in $\mathbb{R}^{12}$ as well as a 5D flat torus example in $\mathbb{R}^{20}$. For the 3D flat torus, we found that irrelevant eigenvalues appear in the left half plane when $N = 3,000$, and those eigenvalues disappear when N increases to $10,000$ [not shown]. For the 5D flat torus, we found that irrelevant eigenvalues always exist in the left half plane even when N increases to $30,000$ [not shown]. Based on these experiments, we believe that more data points are needed for accurate estimations of the leading order spectra of NRBF in higher dimensional cases. This implies that NRBF methods suffer from curse of dimensionality.

We now analyze the results of SRBF. Figure 11 displays the errors of eigenmodes for SRBF for $N = 5000$, 10000, and 30000. One can see from Figs. 11(a) and (b) that the errors of eigenvalues and eigenvectors of the leading eight modes decrease on order of $N^{-1/2}$ for SRBF with $\hat{\mathbf{P}}$ and true q. The convergence can also be examined for each leading mode for SRBF using $\hat{\mathbf{P}}$ and true q in Figs. 11(d) and (f). However, the convergence is not clear for SRBF using $\hat{\mathbf{P}}$ and KDE $\tilde{q}$ as shown in Figs. 11(a)-(c)(e). This implies that $\hat{\mathbf{P}}$ is accurate enough, whereas the sampling density $\tilde{q}$ is not. We suspect that the use of KDE for the density estimation in 4D may not be optimal. This leaves room for future investigations with more accurate density estimation methods.

5.4 2D Sphere

In this section, we study the eigenvector field problem for the Hodge, Bochner and Lichnerowicz Laplacians, on a 2D unit sphere. The parameterization of the unit sphere is given by

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}, \quad \text{for} \quad \begin{matrix} \theta \in [0, \pi] \\ \phi \in [0, 2\pi) \end{matrix}, \tag{37}$$

with Riemannian metric given as,

$$g = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix}. \tag{38}$$

The analytical solution of the eigenvalue problem for the Laplace-Beltrami operator on the unit sphere consists of the set of Laplace's spherical harmonics with corresponding eigenvalue $\lambda = l(l+1)$ for $l \in \mathbb{N}^+$. In particular, the leading spherical harmonics can be written in Cartesian coordinate as

$$\begin{aligned} &\psi = x_i, && \text{for } \lambda = 2, \\ &\psi = 3x_i x_k - \delta_{ik} r^2, && \text{for } \lambda = 6, \\ &\psi = 15 x_i x_k x_n - 3\delta_{ik} r^2 x_n - 3\delta_{kn} r^2 x_i - 3\delta_{ni} r^2 x_k, && \text{for } \lambda = 12, \end{aligned} \tag{39}$$

for indices $i, k, n = 1, 2, 3$, where $(x_1, x_2, x_3) = (x, y, z)$ and $r^2 = x^2 + y^2 + z^2$.

We start with the analysis of spectra and corresponding eigenvector fields of Hodge Laplacian on a unit sphere. The Hodge Laplacian defined on a k-form is

$$\Delta_H = \mathrm{d}_{k-1}\mathrm{d}_{k-1}^* + \mathrm{d}_k^*\mathrm{d}_k,$$

which is self-adjoint and positive definite. Here, d is the exterior derivative and d^* is the adjoint of d given by

$$\mathrm{d}_{k-1}^* = (-1)^{d(k-1)+1} * \mathrm{d}_{d-k} * : \Omega^k(M) \to \Omega^{k-1}(M),$$

where d denotes the dimension of manifold M and $* : \Omega^k(M) \to \Omega^{d-k}(M)$ is the standard Hodge star operator. Obviously, one has $\mathrm{dd} = 0$ and $\mathrm{d}^*\mathrm{d}^* = 0$ and the following diagram:

$$\Omega^0(M) \underset{\mathrm{d}_0^*}{\overset{\mathrm{d}_0}{\rightleftarrows}} \Omega^1(M) \underset{\mathrm{d}_1^*}{\overset{\mathrm{d}_1}{\rightleftarrows}} \Omega^2(M) \underset{\mathrm{d}_2^*}{\overset{\mathrm{d}_2}{\rightleftarrows}} \cdots,$$

with $\Omega^0(M) = C^\infty(M)$. We note that Hodge Laplacian Δ_H reduces to Laplace-Beltrami Δ_M when acting on functions. To obtain the solution of the eigenvalue problem for Hodge of 1-form, we can apply the results from (Folland, 1989) or we can provide one derivation based on the following proposition:

Proposition 5.1 *Let M be a $d-$dimensional manifold and Δ_H be the Hodge Laplacian on k-form. Then,*
(1) $\Omega^k(M) = \operatorname{im} \mathrm{d}_{k-1} \oplus \operatorname{im} \mathrm{d}_k^* \oplus \ker \Delta_H$.
(2) $\lambda(\Delta_H)\setminus\{0\} = \lambda\left(\mathrm{d}_{k-1}\mathrm{d}_{k-1}^*\right)\setminus\{0\} \cup \lambda\left(\mathrm{d}_k^*\mathrm{d}_k\right)\setminus\{0\}$.
(3) $\lambda(\mathrm{d}_k\mathrm{d}_k^*)\setminus\{0\} = \lambda(\mathrm{d}_k^*\mathrm{d}_k)\setminus\{0\}$ *for* $k = 0, 1, \ldots, d$.
(4) $\lambda(\mathrm{d}_k\mathrm{d}_k^*) = \lambda\left(\mathrm{d}_{d-k-1}^*\mathrm{d}_{d-k-1}\right)$ *for* $k = 0, 1, \ldots, d-1$.

Proof (1) is the Hodge Decomposition Theorem. The proof of (2) and (3) can be referred to as *pp* 138 of (Jost and Jost, 2008). For (4), we first show that $\lambda(\mathrm{d}_k\mathrm{d}_k^*) \subseteq \lambda\left(\mathrm{d}_{d-k-1}^*\mathrm{d}_{d-k-1}\right)$. Assume that $\mathrm{d}_k\mathrm{d}_k^*\alpha = \lambda\alpha$ for $\alpha \in \Omega^{k+1}(M)$. Taking Hodge star $*$ at both sides, one arrives at the result, $(-1)^{dk+1} * \mathrm{d}_k * \mathrm{d}_{d-k-1} * \alpha = \mathrm{d}_{d-k-1}^*\mathrm{d}_{d-k-1}(*\alpha) = \lambda * \alpha$. Vice versa one can prove $\lambda(\mathrm{d}_k\mathrm{d}_k^*) \supseteq \lambda\left(\mathrm{d}_{d-k-1}^*\mathrm{d}_{d-k-1}\right)$. ■

Corollary 5.1 *Let M be a 2D surface embedded in $\mathbb{R}^3$ and $\Delta_H = \sharp(\mathrm{dd}^* + \mathrm{d}^*\mathrm{d})\flat$ be the Hodge Laplacian on vector fields $\mathfrak{X}(M)$. Then, the non-trivial eigenvalues of Hodge Laplacian, $\lambda(\Delta_H)$, are identical with the non-trivial eigenvalues of Laplace-Beltrami operator, $\lambda(\Delta_M)\setminus\{0\}$, where the number of eigenvalues of the Hodge Laplacian doubles those of the Laplace-Beltrami operator. Specifically, the corresponding eigenvector fields to nonzero eigenvalues are $\mathbf{P}\nabla_{\mathbb{R}^3}\psi$ and $\boldsymbol{n} \times \nabla_{\mathbb{R}^3}\psi$, where ψ is the eigenfunction of Laplace-Beltrami Δ_M and $\boldsymbol{n}$ is the normal vector to surface M.*

Proof First, we have $\lambda(\mathrm{d}_0\mathrm{d}_0^*)\setminus\{0\} = \lambda(\mathrm{d}_0^*\mathrm{d}_0)\setminus\{0\} = \lambda(\Delta_M)\setminus\{0\}$ using Proposition 5.1 (3). Second, we have $\lambda(\mathrm{d}_1^*\mathrm{d}_1)\setminus\{0\} = \lambda(\mathrm{d}_0\mathrm{d}_0^*)\setminus\{0\} = \lambda(\Delta_M)\setminus\{0\}$ using Proposition 5.1 (3) and (4). Using Proposition 5.1 (2), we obtain the first part in Corollary 5.1. That is, the eigenvalues of Hodge Laplacian for 1-forms or vector fields are exactly of eigenvalues of Laplace-Beltrami, and the number of each eigenvalue of Hodge Laplacian doubles the multiplicity of the corresponding eigenspace of the Laplace-Beltrami operator.

To simplify the notation, we use d to denote d_k for arbitrary k, which is implicitly identified by whichever $k-$form it acts on. We now examine the eigenforms. Assume that

ψ is an eigenfunction for the Laplace-Beltrami operator Δ_M associated with the eigenvalue λ, that is, $\Delta_M\psi = \mathrm{d}_0^*\mathrm{d}_0\psi = \lambda\psi$. Then, one can show that $\mathrm{d}\psi \in \Omega^1(M)$ is an eigenform of Δ_H,

$$\Delta_H \mathrm{d}\psi = (\mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d})\,\mathrm{d}\psi = \mathrm{d}\mathrm{d}^*\mathrm{d}\psi = \mathrm{d}\lambda\psi = \lambda\mathrm{d}\psi. \tag{40}$$

One can also show that $*\mathrm{d}\psi \in \Omega^1(M)$ is an eigenform,

$$\begin{aligned}\Delta_H * \mathrm{d}\psi &= (\mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}) * \mathrm{d}\psi = \mathrm{d}^*\mathrm{d} * \mathrm{d}\psi \\ &= -*\,\mathrm{d} * \mathrm{d} * \mathrm{d}\psi = *\mathrm{d}\mathrm{d}^*\mathrm{d}\psi = *\mathrm{d}\lambda\psi = \lambda * \mathrm{d}\psi,\end{aligned} \tag{41}$$

where we have used $\mathrm{d}_1^* = (-1)^{2(1)+1} * \mathrm{d}_0 *$ and $\mathrm{d}_0^* = (-1)^1 * \mathrm{d}_1 *$. Thus, based on Hodge Decomposition Theorem 5.1 (1), harmonic forms and above eigenforms $\mathrm{d}\psi$ and $*\mathrm{d}\psi$ form a complete space of $\Omega^1(M)$.

Last, we compute the corresponding eigenvector fields. The corresponding eigenvector field for $\mathrm{d}\psi$ is

$$\sharp\mathrm{d}\psi = g^{ij}\psi_i\frac{\partial}{\partial\theta^j} = \mathbf{P}\nabla_{\mathbb{R}^3}\psi. \tag{42}$$

Assume that $\left\{\theta^1, \theta^2\right\}$ is the local normal coordinate so that the Riemannian metric is locally identity $g_{ij} = \delta_{ij}$ at point $\boldsymbol{x}$. The corresponding eigenvector field for $*\mathrm{d}\psi$ is

$$\begin{aligned}\sharp * \mathrm{d}\psi &= \sharp * \psi_i\mathrm{d}\theta^i = \sharp\delta_{ij}^{12}\psi^i\mathrm{d}\theta^j = \sharp\left(\psi^1\mathrm{d}\theta^2 - \psi^2\mathrm{d}\theta^1\right) \\ &= -\frac{\partial\psi}{\partial\theta^2}\frac{\partial}{\partial\theta^1} + \frac{\partial\psi}{\partial\theta^1}\frac{\partial}{\partial\theta^2} = -(\nabla_{\mathbb{R}^3}\psi\cdot\frac{\partial\boldsymbol{x}}{\partial\theta^2})\frac{\partial}{\partial\theta^1} + (\nabla_{\mathbb{R}^3}\psi\cdot\frac{\partial\boldsymbol{x}}{\partial\theta^1})\frac{\partial}{\partial\theta^2} \\ &= \nabla_{\mathbb{R}^3}\psi\times(\frac{\partial\boldsymbol{x}}{\partial\theta^2}\times\frac{\partial\boldsymbol{x}}{\partial\theta^1}) = \boldsymbol{n}\times\nabla_{\mathbb{R}^3}\psi,\end{aligned} \tag{43}$$

where the equality in last line holds true for 2D surface in $\mathbb{R}^3$. ■

Based on (42) and (43), one can immediately obtain the leading eigenvalues and corresponding eigenvector fields for Hodge Laplacian. When $\lambda^H = 2$, the 6 corresponding Hodge eigen vector fields are

$$U_{1,2,3} = \boldsymbol{n}\times\nabla_{\mathbb{R}^3}\psi = \left[\begin{array}{c|c|c} y & -z & 0 \\ -x & 0 & z \\ 0 & x & -y \end{array}\right], U_{4,5,6} = \mathbf{P}\nabla_{\mathbb{R}^3}\psi = \left[\begin{array}{c|c|c} xz & xy & -y^2-z^2 \\ yz & -x^2-z^2 & xy \\ -x^2-y^2 & zy & xz \end{array}\right],$$

where $U_{1,2,3}$ are computed from the curl formula (43) and $U_{4,5,6}$ are computed from the projection formula (42) by taking $\psi = x_i$ $(i = 1, 2, 3)$ in (39). When $\lambda^H = 6$, the 10 Hodge eigenvector fields are

$$\begin{aligned}U_{7\sim11} &= \boldsymbol{n}\times\nabla_{\mathbb{R}^3}\psi = \left[\begin{array}{c|c|c|c|c} -xz & y^2-z^2 & xy & 0 & -yz \\ yz & -xy & z^2-x^2 & xz & 0 \\ x^2-y^2 & xz & -yz & -xy & xy \end{array}\right], \\ U_{12\sim16} &= \mathbf{P}\nabla_{\mathbb{R}^3}\psi = \left[\begin{array}{c|c|c|c|c} y-2x^2y & z-2x^2z & -2xyz & x-x^3 & -xy^2 \\ x-2xy^2 & -2xyz & z-2y^2z & -x^2y & y-y^3 \\ -2xyz & x-2xz^2 & y-2yz^2 & -x^2z & -y^2z \end{array}\right],\end{aligned}$$

where $U_{7\sim11}$ are computed from the curl (43) and $U_{12\sim16}$ are computed from the projection (42) by taking $\psi = 3x_ix_k - \delta_{ik}r^2$ in (39).

For the Bochner Laplacian, we notice that it is different from the Hodge Laplacian by a Ricci tensor and Ricci curvature is constant on the sphere. Therefore, the Bochner and Hodge Laplacians share the same eigenvector fields but have different eigenvalues. We examined that $U_{1\sim 6}$ are eigenvector fields for $\lambda^B = 1$ and $U_{7\sim 16}$ are eigenvector fields for $\lambda^B = 5$. In general, the spectrum are $\lambda^B = l(l+1) - 1$ for the Bochner Laplacian and $\lambda^H = l(l+1)$ for the Hodge Laplacian for $l \in \mathbb{N}^+$.

For the Lichnerowicz Laplacian, we can verify that $U_{1,2,3}$ are in the null space $\ker \Delta_L$, which is often referred to as the Killing field. We can further verify that $U_{4,5,6}$ correspond to $\lambda^L = 2$, $U_{7\sim 11}$ correspond to $\lambda^L = 4$, and $U_{12\sim 16}$ correspond to $\lambda^L = 10$. Moreover, we verify that $U_{17\sim 23}$ corresponds to $\lambda^L = 10$, where $U_{17\sim 23}$ are computed from the curl equation (43) by taking $\psi = 15x_ix_kx_n - 3\delta_{ik}r^2x_n - 3\delta_{kn}r^2x_i - 3\delta_{ni}r^2x_k$. We refer to the eigenvalues λ associated with the eigenvector fields U obtained above as the analytic solutions to the eigenvalue problem of vector Laplacians.

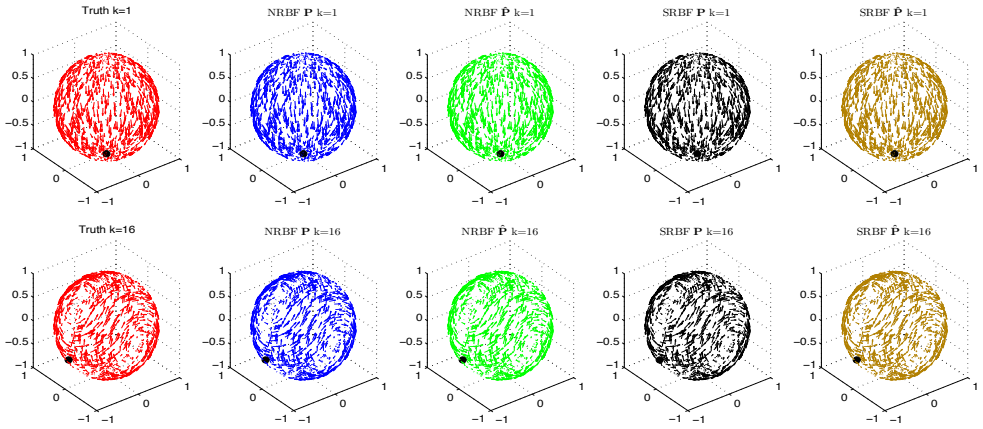

Figure 12: **2D Sphere in $\mathbb{R}^3$.** Comparison of eigen-vector fields of Bochner Laplacian for $k = 1, 16$ among NRBF using $\mathbf{P}$ and $\hat{\mathbf{P}}$, and SRBF using $\mathbf{P}$ and $\hat{\mathbf{P}}$. Black dots correspond the poles where the vector fields vanish. For NRBF, GA kernel with $s = 1.0$ is used, and for SRBF, IQ kernel with $s = 0.5$ is used. The $N = 1024$ data points are randomly distributed on the manifold.

In the following, we compute eigenvalues and associated eigenvector fields of the Hodge, Bochner, and Lichnerowicz Laplacians on a unit sphere. We use uniformly random sample data points on the sphere for one trial comparison in most of the following figures, except for Fig. 15 in which we use well-sampled data points for verifying the convergence of the SRBF method. Figure 12 displays the comparison of eigenvector fields of Bochner Laplacian for $k = 1$ and 16. When $\mathbf{P}$ is used, we have examined that these vector fields at each point are orthogonal to the normal directions $\boldsymbol{n} = \boldsymbol{x} = (x, y, z)$ of the sphere. When $\hat{\mathbf{P}}$ is used (i.e. in an unknown manifold scenario), the vector fields lie in the approximated tangent space which is orthogonal to the normal $\boldsymbol{x}$ within numerical accuracy. Based on Hairy ball theorem, a vector field vanishes at one or more points on the sphere. Here, we plot the poles where the vector field vanishes with black dots.

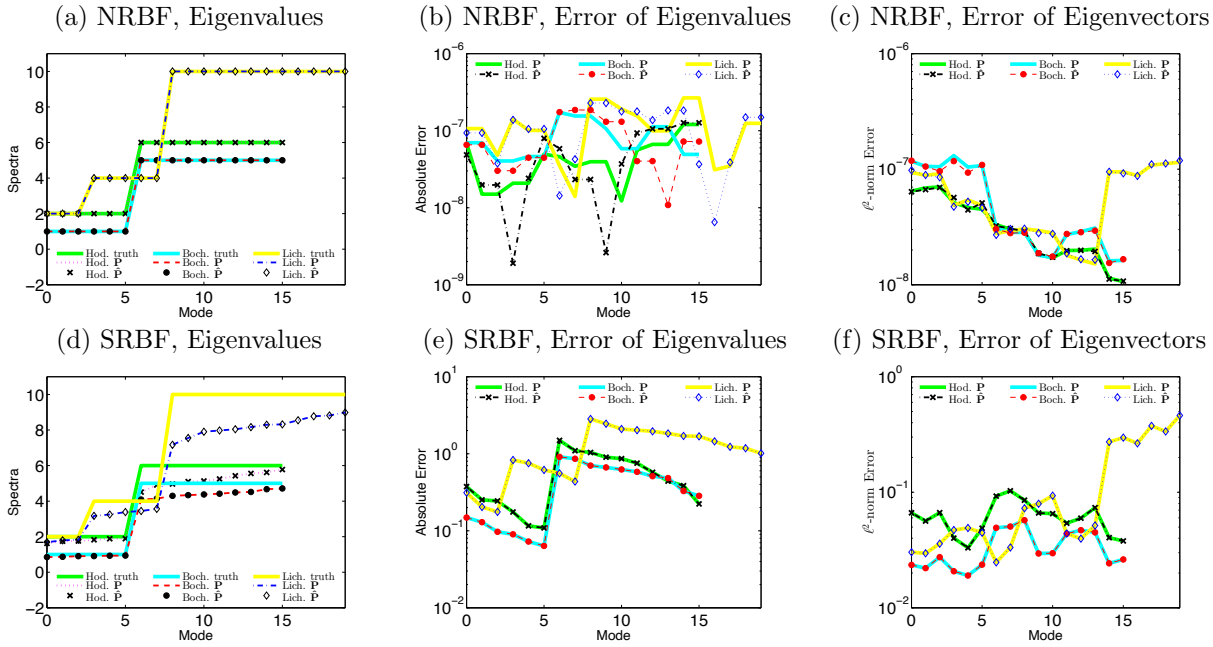

Figure 13: (Color online) 2D Sphere in $\mathbb{R}^3$. Comparison of (a)(d) eigenvalues, (b)(e) error of eigenvalues, and (c)(f) error of eigenvector-fields for Hodge, Bochner, Lichnerowicz Laplacians. (a)(b)(c) correspond to NRBF and (d)(e)(f) correspond to SRBF results. For NRBF, GA kernel with $s = 1.0$ is used, and for SRBF, IQ kernel with $s = 0.5$ is used. The data points are randomly distributed with $N = 1024$.

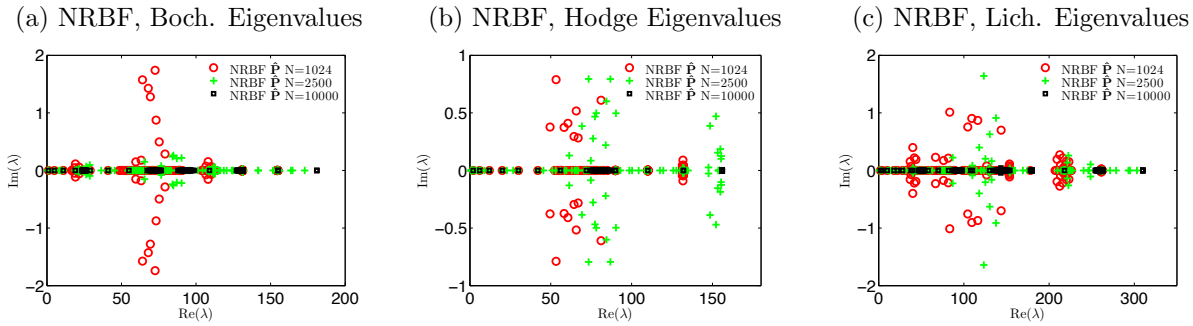

Figure 14: **2D Sphere in $\mathbb{R}^3$.** Convergence of eigenvalues for NRBF using $\hat{\mathbf{P}}$ for (a) Bochner, (b) Hodge, and (c) Lichnerowicz Laplacians. GA kernel with $s = 1.0$ was fixed for different N. The data points are randomly distributed.

Figure 13 displays the comparison of the leading eigenvalues and eigenvector fields for the Hodge, Bochner, and Lichnerowicz Laplacians. For NRBF, one can see from Fig. 13(a)-(c) that both eigenvalues and eigenvector fields can be approximated very accurately using either the analytic $\mathbf{P}$ or estimated $\hat{\mathbf{P}}$. This small error result can be expected using analytic $\mathbf{P}$ for the known manifold. It is a little unexpected that the estimation produces such a small error result using estimated $\hat{\mathbf{P}}$ when the manifold is unknown. After further inspection, we found that our 2nd order SVD provides a super-convergence for $\hat{\mathbf{P}}$ on this particular 2D sphere example. For SRBF, one can see from Fig. 13(d)-(f) that eigenvalues and eigenvector fields can be approximated within certain errors for both $\mathbf{P}$ and $\hat{\mathbf{P}}$. This means that the Monte-Carlo error dominates the error for $\hat{\mathbf{P}}$ in this example.

To inspect more of these spectra, we compare the leading eigenvalues of the truth, NRBF, and SRBF with $\hat{\mathbf{P}}$ up to $k = 80$ in Table 3. For NRBF, one can see from Table 3 that the leading NRBF eigenvalues are often in excellent agreement with the truth. Then, irrelevant eigenvalues may appear but are followed by the true eigenvalues again. For instance, the first 30 Bochner eigenvalues are accurately approximated, the $31^{\text{st}} \sim 40^{\text{th}}$ eigenvalues are irrelevant (spectral pollution), and $41^{\text{st}} \sim 58^{\text{th}}$ eigenvalues become very accurate again. To further understand the irrelevant spectra phenomenon, we also solve the problem with different kernels and/or tune different shape parameters. Interestingly, we found that the leading $\sqrt{N} \sim 30$ true eigenvalues can always be accurately approximated even under different kernels or shape parameters. Unfortunately, the irrelevant eigenvalues also appear with different values starting from at least $k \geq \sqrt{N}$ for different kernels or shape parameters [not shown here]. In these experiments, the data points are always fixed. In contrast, SRBFs do not exhibit such an issue despite the fact that the eigenvalue estimates are not as accurate as NRBFs. Several advantages of SRBF are that their eigenvalues are real-valued, well-ordered, and do not produce irrelevant estimates.

Figure 14 shows the convergence of eigenvalues of NRBF methods for vector Laplacians. The data in this figure are randomly distributed. One can see that the four observations in Example 5.2 are still valid for the behavior of NRBF eigenvalues. Figure 15 shows the convergence of eigenvalues and eigenvectors of the SRBF method for vector Laplacians. For

Table 3: 2D sphere in $\mathbb{R}^3$. Comparison of eigenvalues of Bochner, Hodge, and Lichnerowicz Laplacians from NRBF and SRBF using approximated $\hat{\mathbf{P}}$. The $N = 1024$ data points are randomly distributed on the sphere. The eigenvalues of NRBF are shown with their absolute values when they are complex. For NRBF, GA kernel with $s = 1.0$ is used, and for SRBF, IQ kernel with $s = 0.5$ is used.

k	Boch	Boch	Boch	Hodg	Hodg	Hodg	Lich	Lich	Lich	k	Boch	Boch	Boch	Hodg	Hodg	Hodg	Lich	Lich	Lich
	True	NRBF	SRBF	True	NRBF	SRBF	True	NRBF	SRBF		True	NRBF	SRBF	True	NRBF	SRBF	True	NRBF	SRBF
1	1	1.0	0.85	2	2.0	1.63	0	–	–	41	19	19.0	15.7	20	20.0	15.3	28	28.0	19.0
2	1	1.0	0.87	2	2.0	1.75	0	–	–	42	19	19.0	15.8	20	20.0	15.5	28	28.0	19.8
3	1	1.0	0.90	2	2.0	1.76	0	–	–	43	19	19.0	16.1	20	20.0	15.8	28	28.0	20.4
4	1	1.0	0.91	2	2.0	1.82	2	2.0	1.69	44	19	19.0	16.1	20	20.0	15.8	28	28.0	20.7
5	1	1.0	0.93	2	2.0	1.88	2	2.0	1.80	45	19	19.0	16.4	20	20.0	16.2	28	28.0	21.4
6	1	1.0	0.94	2	2.0	1.89	2	2.0	1.82	46	19	19.0	16.6	20	20.0	16.4	28	28.0	21.6
7	5	5.0	4.09	6	6.0	4.51	4	4.0	3.17	47	19	19.0	17.0	20	20.0	16.5	28	28.0	22.1
8	5	5.0	4.14	6	6.0	4.91	4	4.0	3.25	48	19	19.0	17.1	20	20.0	16.7	28	28.0	22.7
9	5	5.0	4.30	6	6.0	4.96	4	4.0	3.38	49	29	19.0	18.9	30	30.0	16.8	28	28.0	23.3
10	5	5.0	4.34	6	6.0	5.10	4	4.0	3.44	50	29	19.0	19.2	30	30.0	17.3	28	28.0	23.9
11	5	5.0	4.38	6	6.0	5.14	4	4.0	3.56	51	29	19.0	19.7	30	30.0	17.4	38	33.1	24.2
12	5	5.0	4.42	6	6.0	5.24	10	10.0	7.17	52	29	19.0	20.1	30	30.0	17.6	38	33.4	24.9
13	5	5.0	4.49	6	6.0	5.43	10	10.0	7.55	53	29	19.0	20.9	30	30.0	17.7	38	33.9	25.6
14	5	5.0	4.52	6	6.0	5.56	10	10.0	7.90	54	29	19.0	21.1	30	30.0	17.8	38	34.5	26.2
15	5	5.0	4.67	6	6.0	5.62	10	10.0	7.97	55	29	19.0	21.4	30	30.0	17.9	38	35.0	26.4
16	5	5.0	4.72	6	6.0	5.78	10	10.0	8.05	56	29	19.0	21.5	30	30.0	18.2	38	35.2	26.8
17	11	11.0	8.27	12	12.0	7.92	10	10.0	8.17	57	29	19.0	22.1	30	30.0	18.4	38	36.5	27.0
18	11	11.0	8.46	12	12.0	8.38	10	10.0	8.30	58	29	19.0	22.3	30	30.0	18.6	38	36.9	27.3
19	11	11.0	8.65	12	12.0	8.64	10	10.0	8.32	59	29	19.1	22.6	30	30.0	18.7	38	36.9	27.8
20	11	11.0	8.80	12	12.0	9.26	10	10.0	8.55	60	29	19.1	22.9	30	30.0	18.8	40	37.7	28.0
21	11	11.0	8.90	12	12.0	9.44	10	10.0	8.77	61	29	19.5	23.6	30	30.0	19.0	40	37.9	28.3
22	11	11.0	9.03	12	12.0	9.61	10	10.0	8.82	62	29	20.1	23.9	30	30.0	19.1	40	37.9	28.7
23	11	11.0	9.13	12	12.0	9.74	10	10.0	8.99	63	29	20.1	24.2	30	30.0	19.2	40	38.0	29.3
24	11	11.0	9.27	12	12.0	9.99	18	18.0	12.2	64	29	20.6	24.4	30	30.0	19.5	40	38.0	29.4
25	11	11.0	9.47	12	12.0	10.2	18	18.0	12.7	65	29	20.9	24.8	30	30.0	19.6	40	38.0	30.2
26	11	11.0	9.56	12	12.0	10.5	18	18.0	13.2	66	29	21.1	25.2	30	30.0	19.8	40	38.0	30.7
27	11	11.0	9.87	12	12.0	10.6	18	18.0	13.4	67	29	21.6	25.7	30	30.0	20.0	40	38.0	31.3
28	11	11.0	9.98	12	12.0	10.9	18	18.0	13.6	68	29	21.6	25.7	30	30.0	20.1	40	38.0	31.5
29	11	11.0	10.3	12	12.0	11.4	18	18.0	13.8	69	29	21.8	26.1	30	30.0	20.5	40	38.0	32.4
30	11	11.0	10.5	12	12.0	11.6	18	18.0	14.1	70	29	22.0	26.2	30	30.0	20.7	40	38.0	32.6
31	19	16.7	13.3	20	20.0	11.7	18	18.0	14.7	71	41	22.4	26.9	42	42.0	20.7	40	38.0	33.7
32	19	16.8	13.5	20	20.0	12.6	18	18.0	15.6	72	41	22.8	27.0	42	42.0	21.0	40	38.5	33.7
33	19	17.1	13.7	20	20.0	12.8	22	22.0	15.9	73	41	23.3	27.4	42	42.0	21.1	54	40.0	34.8
34	19	17.4	13.8	20	20.0	13.1	22	22.0	16.3	74	41	23.7	27.6	42	42.0	21.2	54	40.0	34.9
35	19	17.7	14.1	20	20.0	13.8	22	22.0	16.6	75	41	24.3	28.0	42	42.0	21.6	54	40.0	35.7
36	19	18.0	14.4	20	20.0	13.9	22	22.0	17.2	76	41	25.4	28.2	42	42.0	21.7	54	40.0	35.9
37	19	18.5	14.6	20	20.0	14.2	22	22.0	17.3	77	41	25.8	28.6	42	42.0	21.9	54	40.0	36.4
38	19	18.6	14.8	20	20.0	14.6	22	22.0	17.7	78	41	29.0	28.8	42	42.0	22.0	54	40.0	36.8
39	19	18.6	15.2	20	20.0	14.7	22	22.0	18.3	79	41	29.0	29.2	42	42.0	22.3	54	40.0	37.2
40	19	18.9	15.3	20	20.0	15.1	28	28.0	18.7	80	41	29.0	29.5	42	42.0	22.5	54	40.0	38.7

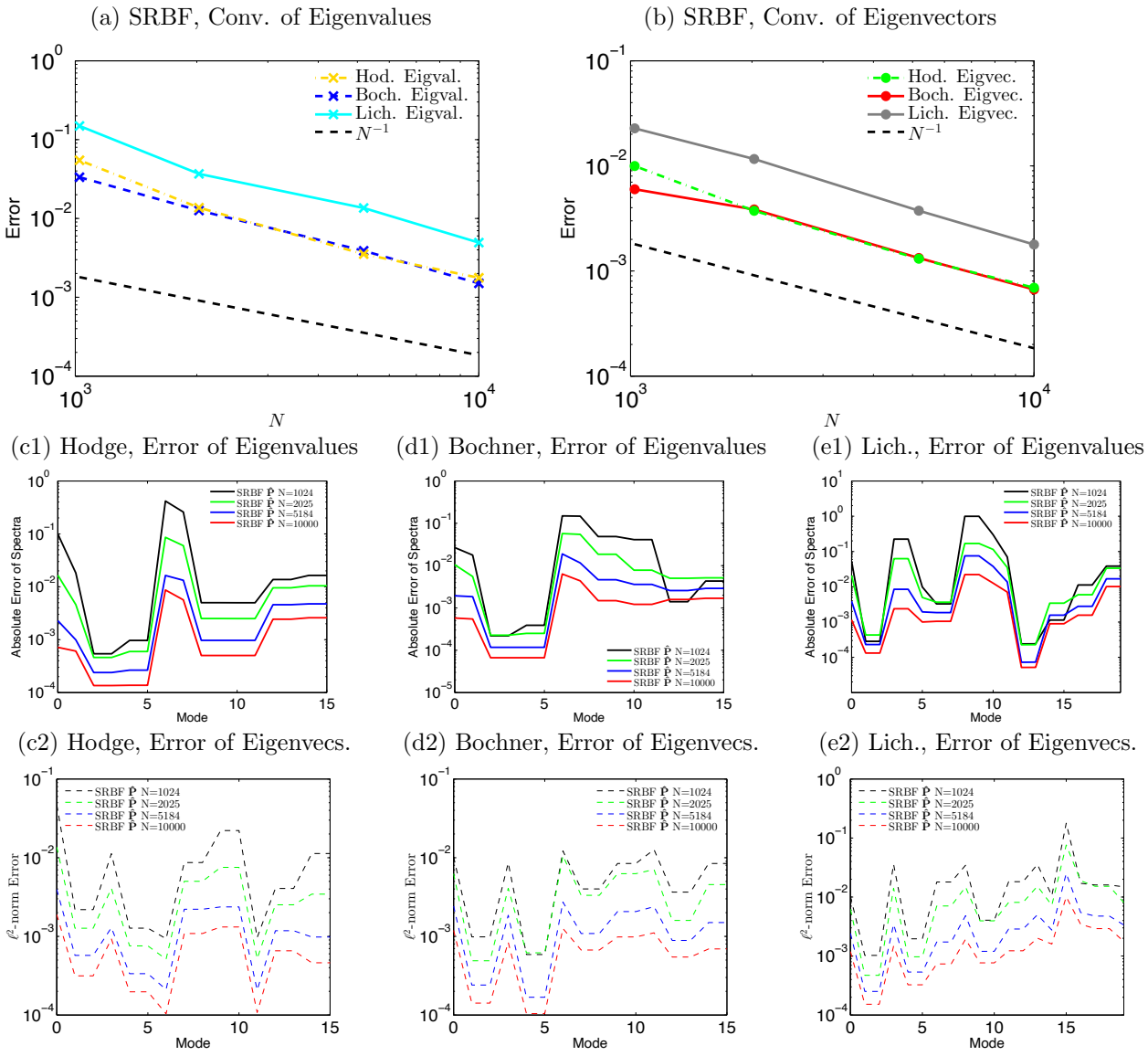

Figure 15: **2D Sphere in $\mathbb{R}^3$.** The data points are uniformly well-sampled on the manifold and the true sampling density q was used. Convergence of (a) eigenvalues and (b) eigenvectors for the average of the leading 16 modes for Bochner and Hodge Laplacians, and the average of the leading 20 modes for the Lichnerowicz Laplacian for SRBF using $\hat{\mathbf{P}}$ and true q are shown. Plotted in the second row are the errors of eigenvalues for (c1) Hodge, (d1) Bochner, and (e1) Lichnerowicz Laplacians for each leading mode for SRBF. Plotted in (c2)-(e2) are the corresponding errors of eigenvectors. IQ kernel with $s = 0.5$ was fixed for SRBF with different N.

an efficient computational labor, we show results with well-sampled data points $\{\theta_i, \phi_j\}$, defined as equally spaced points in both direction of the intrinsic coordinates $[0, \pi] \times [0, 2\pi)$ such that the total number of points are $N = 32^2, 45^2, 71^2, 100^2$. We use the approximated $\hat{\mathbf{P}}$ and the true sampling density q to construct the Laplacians. One can see from Figs. 15(a) and (b) that the errors of eigenvalues and eigenvectors for all vector Laplacians decay on order of N^{-1}. More details can be found in Figs. 15(c)-(e) for the leading eigenmodes. Notice that for the well-sampled data, the errors here are relatively small compared to those in the case of random data shown in Figs. 13(d)-(f). Also, notice that the error rates here decay on order of N^{-1} for well-sampled data, which is faster than Monte Carlo error rates of $N^{-1/2}$ for random data.

6. Summary and Discussion

In this paper, we studied the Radial Basis Function (RBF) approximation to differential operators on smooth, closed Riemannian submanifolds of Euclidean space, identified by randomly sampled point cloud data. We summarize the findings in this paper and conclude with some open issues that arise from this study:

- The classical pointwise RBF formulation is an effective and accurate method for approximating general differential operators on closed manifolds. For unknown manifolds, i.e., identified only by point clouds, the approximation error is dominated by the accuracy of the estimation of local tangent space. With sufficient regularity on the functions and manifolds, we improve the local SVD technique to achieve a second-order accuracy, by appropriately accounting for the curvatures of the manifolds.

- The pointwise non-symmetric RBF formulation can produce very accurate estimations of the leading eigensolutions of Laplacians when the size of data used to construct the RBF matrix N is large enough and the projection operator $\mathbf{P}$ is accurately estimated. We empirically found that the accuracy is significantly better than those produced by the graph-based method, which is supported by the results in Figure 5. In Figure 7, for sufficiently large N, one can further improve the estimates by increasing $N_p \gg N$, the number of points used to estimate $\mathbf{P}$. While the second-order local SVD method is numerically efficient even for large N_p, the ultimate computational cost depends on the size of N as we are solving the eigenvalue problem of dense, non-symmetric matrices of size $N \times N$ for Laplace-Beltrami or $nN \times nN$ for vector Laplacians. When both N and N_p are not large enough, this approximation produces eigensolutions that are not consistent with the spectral properties of the Laplacians in the sense that the eigenvalue estimates can be complex-valued and do not correspond to any of the underlying real-valued spectra. Since the spectral accuracy of NRBF relies on a very accurate estimation of the projection operator $\mathbf{P}$, we believe that this method may not be robust when the data is corrupted by noise, and thus, may not be reliable for manifold learning.

- The proposed symmetric RBF approximation for Laplacians overcomes the issues in the pointwise formulation. The only caveat with the symmetric formulations is that this approximation may not be as accurate as the pointwise approximation whenever

the latter works. Based on our analysis, the error of SRBF is dominated by the error of discretizing integrals in the weak formulation of appropriate Hilbert spaces, provided sufficiently smooth kernels are used. See Theorems 4.1 and 4.2 for the approximation of eigenvalues and eigenfunctions of the Laplace-Beltrami operator, respectively. See Theorems 4.3 and 4.4 for the approximation of eigenvalues and eigenvector fields of Bochner Laplacian, respectively. Empirically, we found that SRBF consistently produces more accurate estimations of the non-leading eigenvalues and less accurate estimations of the leading eigenvalues compared to those obtained from diffusion maps algorithm for Laplacians acting on functions. These encouraging results suggest that the symmetric RBF discrete Laplacian approximation is potentially useful to complement graph-based methods when $d \ll n \ll N$, which arises in many PDE models on unknown sub-manifolds of moderately high dimensions, e.g., $n = O(10)$. Unlike graph-based approaches, which are restricted to approximating Laplacians, the general formulation in this paper supports PDE applications involving other differential operators. In a companion paper (Yan et al., 2023), we have introduced a spectral Galerkin framework to solve elliptic PDEs based on SRBF.

- The main limitation of this formulation is that it is not scalable for manifold learning problems where $n = O(N)$. One of the key computational issues with this formulation is that it requires n number of dense $N \times N$ matrices $\mathbf{G}_i$ to be stored, which is memory intensive when $n = O(N)$. Further memory issues are encountered in the approximation of operators on vector fields and higher-order tensor fields. For instance, in our formulation, a Laplacian on vector fields is numerically resolved with an $nN \times nN$ matrix, which again requires significant memory whenever $n = O(N)$. An alternative approach to avoid this significant memory issue is to resort to an intrinsic approach (Liang and Zhao, 2013), where the authors use approximate tangent vectors to directly construct the metric tensor with a local polynomial regression and subsequently resolve the derivatives on manifolds using local tangent coordinates. While they have demonstrated fast convergence of their method for solving PDEs on relatively low dimensional manifolds with low co-dimensions and their approach has been implemented to approximate differential operators on vector fields on 2-dimensional surfaces embedded in $\mathbb{R}^3$ (Gross et al., 2020), in all of their examples the manifolds are well-sampled in the sense that the data points are visually regularly spaced. It remains unclear how accurate and stable this method is on randomly sampled data as in our examples. In addition, the intrinsic formulation may be difficult to implement since the local representation of a specific differential operator requires a separate hardcode that depends on the inverse and/or derivatives of the metric tensor components that are being approximated.

- Parameter tuning: There are two key parameters to be tuned whenever RBF is used. First is the tolerance parameter in the pseudo-inverse, which determines the rank of the discrete approximation. A smaller tolerance parameter value increases the condition number, while too large of a tolerance parameter value reduces the number of non-trivial eigensolutions that can be estimated. Second, is the shape parameter of the RBF kernel. More accurate approximations seem to occur with smaller values of shape parameters, which yields a denser matrix such that the corresponding eigen-

value problem becomes expensive, especially for large data. This is the opposite of the graph-based approach which tends to be more accurate with smaller bandwidth parameters (inversely proportional to shape parameters) which induce sparser matrices. In a forthcoming paper, we will avoid parameter tuning by replacing the radial basis interpolant with a local polynomial interpolant defined on the estimated tangent space.

- For the discrete approximation corresponding to an unweighted L^2-space, one needs to de-bias the sampling distribution induced by the random data with appropriate Monte-Carlo sampling weights (or density function). When the sampling density of the data is unknown, one needs to approximate it as in the graph-based approximation. Hence, the accuracy of the estimation depends on the accuracy of the density estimation, which we did not analyze in this paper as we only implement the classical kernel density estimation technique. In the case when the operator to be estimated is defined weakly over an L^2-space that is weighted with respect to the sampling distribution, then one does not need to estimate the sampling density function. This is in contrast to the graph Laplacian-based approach. Particularly, graph Laplacian approximation based on the diffusion maps asymptotic expansion imposes a "right normalization" over the square root of the approximate density.

- For the non-symmetric approximation of the Laplacians, the spectral consistency of the approximation in the limit of large data still remains open. While this approximation is subject to spectral pollution, it is worthwhile to consider or develop methods to identify spectra without pollution, following ideas in Colbrook and Townsend (2024).

- While we believe the symmetric formulation should be robust to small amplitude noises in the ambient space as numerically studied in our companion paper (Yan et al., 2023), this claim remains to be theoretically justified. Particularly, further investigation is needed to extend Theorem 3.1 for noisy data.

Acknowledgment

The research of J.H. was partially supported by the NSF grants DMS-1854299, DMS-2207328, and the ONR grant N00014-22-1-2193. S. J. was supported by the NSFC Grant No. 12101408 and the HPC Platform of ShanghaiTech University. S. J. also thanks Chengjian Yao for useful discussion.

Appendix A. More Operators on Manifolds

A.1 Proof for Proposition 2.2

We first provide the proof for equation (13) in Proposition 2.2.

Proof From a direct calculation, we have

$$\sum_r g^{ij}\frac{\partial X^r}{\partial\theta^j}\frac{\partial U^r}{\partial\theta^k}=\sum_r g^{ij}\frac{\partial X^r}{\partial\theta^j}\frac{\partial}{\partial\theta^k}\left(u^p\frac{\partial X^r}{\partial\theta^p}\right)=\sum_r g^{ij}\frac{\partial X^r}{\partial\theta^j}\left(\frac{\partial u^p}{\partial\theta^k}\frac{\partial X^r}{\partial\theta^p}+u^p\frac{\partial^2 X^r}{\partial\theta^k\partial\theta^p}\right),$$

where we have used (12) in the first equality above. Using the fact that $\sum_r \frac{\partial X^r}{\partial\theta^i}\frac{\partial X^r}{\partial\theta^k}=g_{ik}$,

$$\sum_r g^{ij}\frac{\partial X^r}{\partial\theta^j}\frac{\partial U^r}{\partial\theta^k}=g^{ij}g_{jp}\frac{\partial u^p}{\partial\theta^k}+\sum_r g^{ij}u^p\left(\frac{\partial X^r}{\partial\theta^j}\frac{\partial^2 X^r}{\partial\theta^k\partial\theta^p}\right)=\frac{\partial u^i}{\partial\theta^k}+\sum_r g^{ij}u^p\left(\frac{\partial X^r}{\partial\theta^j}\frac{\partial^2 X^r}{\partial\theta^k\partial\theta^p}\right). \quad (44)$$

It remains to observe that

$$\sum_r\left(\frac{\partial X^r}{\partial\theta^j}\frac{\partial^2 X^r}{\partial\theta^k\partial\theta^p}\right)=\left\langle\frac{\partial\mathbf{X}}{\partial\theta^j},\frac{\partial^2\mathbf{X}}{\partial\theta^k\partial\theta^p}\right\rangle=\left\langle\frac{\partial\mathbf{X}}{\partial\theta^j},\sum_r\Gamma^r_{pk}\frac{\partial\mathbf{X}}{\partial\theta^r}+\mathbf{n}\right\rangle=\sum_r\Gamma^r_{pk}\left\langle\frac{\partial\mathbf{X}}{\partial\theta^j},\frac{\partial\mathbf{X}}{\partial\theta^k}\right\rangle=\sum_r g_{jr}\Gamma^r_{pk},$$

where we have used the fact that $\frac{\partial^2\mathbf{X}}{\partial\theta^k\partial\theta^p}$ is a linear combination of partials of $\mathbf{X}$ with Christoffel symbols as coefficients, plus a vector $\mathbf{n}$ orthogonal to the tangent space. Plugging this to (44) yields Equation (13). ■

In the remainder of this appendix, we now show explicitly how other differential operators can be approximated using the tangential projection formulation. In particular, we derive formulas for approximations of the Lichnerowicz and Hodge Laplacians, as well as the covariant derivative. Before deriving such results, we explicitly compute the matrix form of divergence of a $(2,0)$ tensor field, which will be needed in the following subsections. Throughout the following calculations, we will use the following shorthand notation for simplification:

$$v^{jk}_{,s}:=v^{mk}\Gamma^j_{sm}+\frac{\partial v^{jk}}{\partial\theta^s}+v^{jm}\Gamma^k_{sm},$$

where $v=v^{jk}\frac{\partial}{\partial\theta^j}\otimes\frac{\partial}{\partial\theta^k}$ is a $(2,0)$ tensor field. While we do not prove convergence in the probabilistic sense, the authors suspect that techniques similar to those used in Section 4 can be used to obtain the convergence results for the other differential operators. In Section 5, convergence is demonstrated numerically for the approximations derived in this appendix.

A.2 Derivation of Divergence of a (2,0) Tensor Field

Let v be a $(2,0)$ tensor field of $v=v^{jk}\frac{\partial}{\partial\theta^j}\otimes\frac{\partial}{\partial\theta^k}$. The divergence of v is defined as

$$\mathrm{div}^r_1(v)=C^r_1(\nabla v)$$

for $r=1$ or $r=2$, where C^r_1 denotes the contraction. More explicitly, the divergence $\mathrm{div}^r_1(v)$ can be calculated as,

$$\mathrm{div}^r_1(v)=C^r_1\left(v^{jk}_{,s}\frac{\partial}{\partial\theta^j}\otimes\frac{\partial}{\partial\theta^k}\otimes\mathrm{d}\theta^s\right)=\begin{cases}v^{jk}_{,j}\frac{\partial}{\partial\theta^k}, & r=1\\ v^{jk}_{,k}\frac{\partial}{\partial\theta^j} & r=2\end{cases}, \quad (45)$$

where $v^{jk}_{,s} = v^{mk}\Gamma^j_{sm} + \frac{\partial v^{jk}}{\partial \theta^s} + v^{jm}\Gamma^k_{sm}$. For any $p \in M$, we can extend v to V, defined on a neighborhood of p with ambient space representation

$$V = \frac{\partial X^s}{\partial \theta^j} v^{jk} \frac{\partial X^t}{\partial \theta^k} \frac{\partial}{\partial X^s} \otimes \frac{\partial}{\partial X^t} \equiv V^{st} \frac{\partial}{\partial X^s} \otimes \frac{\partial}{\partial X^t},$$

so that $V|_p = v|_p$ agrees at $p \in M$. We show the following formula for the divergence for a $(2,0)$ tensor field:

$$\mathrm{div}^r_1(v) = \mathcal{P}C^r_1\left(\mathcal{P}_3 \bar{\nabla}_{\mathbb{R}^n} V\right), \tag{46}$$

where the right hand side is defined as

$$\begin{aligned}
\mathcal{P}C^r_1\left(\mathcal{P}_3 \bar{\nabla}_{\mathbb{R}^n} V\right) &\equiv \mathcal{P}C^r_1\left(\left[\delta_{cq} \frac{\partial X^c}{\partial \theta^a} g^{ab} \frac{\partial X^d}{\partial \theta^b} \mathrm{d}X^q \otimes \frac{\partial}{\partial X^d}\right]\left[\frac{\partial V^{st}}{\partial X^m} \frac{\partial}{\partial X^s} \otimes \frac{\partial}{\partial X^t} \otimes \mathrm{d}X^m\right]\right) \\
&= \mathcal{P}C^r_1\left(\delta_{cq} \frac{\partial X^c}{\partial \theta^a} g^{ab} \frac{\partial X^m}{\partial \theta^b} \frac{\partial V^{st}}{\partial X^m} \frac{\partial}{\partial X^s} \otimes \frac{\partial}{\partial X^t} \otimes \mathrm{d}X^q\right).
\end{aligned}$$

Here, $\mathcal{P}_3$ is as defined in (14) and $\mathcal{P}_3$ acts on the last 1-form component $\mathrm{d}X^m$. For $r = 1$, RHS of (46) can be calculated as,

$$\begin{aligned}
&\mathcal{P}C^r_1\left(\mathcal{P}_3 \bar{\nabla}_{\mathbb{R}^n} V\right) \\
= \;& \mathcal{P}\left(\delta_{cs} \frac{\partial X^c}{\partial \theta^a} g^{ab} \frac{\partial X^m}{\partial \theta^b} \frac{\partial V^{st}}{\partial X^m} \frac{\partial}{\partial X^t}\right) = \mathcal{P}\left(\sum_{s=1}^n \frac{\partial X^s}{\partial \theta^a} g^{ab} \frac{\partial}{\partial \theta^b}\left(\frac{\partial X^s}{\partial \theta^j} v^{jk} \frac{\partial X^t}{\partial \theta^k}\right) \frac{\partial}{\partial X^t}\right) \\
= \;& \mathcal{P}\left(\sum_{s=1}^n \frac{\partial X^s}{\partial \theta^a} g^{ab}\left(\frac{\partial^2 X^s}{\partial \theta^b \partial \theta^j} v^{jk} \frac{\partial X^t}{\partial \theta^k} + \frac{\partial X^s}{\partial \theta^j} \frac{\partial v^{jk}}{\partial \theta^b} \frac{\partial X^t}{\partial \theta^k} + \frac{\partial X^s}{\partial \theta^j} v^{jk} \frac{\partial^2 X^t}{\partial \theta^b \partial \theta^k}\right) \frac{\partial}{\partial X^t}\right),
\end{aligned}$$

where the second line follows from Leibniz rule. Unraveling definitions, one obtains

$$\begin{aligned}
\mathcal{P}C^r_1\left(\mathcal{P}_3 \bar{\nabla}_{\mathbb{R}^n} V\right) &= \mathcal{P}\left(\left(g^{ab} v^{jk} \frac{\partial X^t}{\partial \theta^k} g_{am} \Gamma^m_{bj} + g_{aj} g^{ab} \frac{\partial v^{jk}}{\partial \theta^b} \frac{\partial X^t}{\partial \theta^k} + g_{aj} g^{ab} v^{jk} \frac{\partial^2 X^t}{\partial \theta^b \partial \theta^k}\right) \frac{\partial}{\partial X^t}\right) \\
&= \left(v^{jk} \frac{\partial X^t}{\partial \theta^k} \Gamma^m_{mj} + \frac{\partial v^{jk}}{\partial \theta^j} \frac{\partial X^t}{\partial \theta^k}\right) \frac{\partial}{\partial X^t} + \mathcal{P}\left(v^{jk} \frac{\partial^2 X^t}{\partial \theta^j \partial \theta^k} \frac{\partial}{\partial X^t}\right),
\end{aligned}$$

where we have used that $\mathcal{P}v = v$ for $v \in \mathfrak{X}(M)$, which holds for the first and second terms above. Using the definition of projection tensor and the ambient space formulation of the connection for the third term above, one obtains

$$\begin{aligned}
&\mathcal{P}C^r_1\left(\mathcal{P}_3 \bar{\nabla}_{\mathbb{R}^n} V\right) \\
= \;& \left(v^{jk} \frac{\partial X^t}{\partial \theta^k} \Gamma^m_{mj} + \frac{\partial v^{jk}}{\partial \theta^j} \frac{\partial X^t}{\partial \theta^k}\right) \frac{\partial}{\partial X^t} + \left(\delta_{il} \frac{\partial X^i}{\partial \theta^r} g^{rp} \frac{\partial X^h}{\partial \theta^p} \mathrm{d}X^l \otimes \frac{\partial}{\partial X^h}\right)\left(v^{jk} \frac{\partial^2 X^t}{\partial \theta^j \partial \theta^k} \frac{\partial}{\partial X^t}\right) \\
= \;& \left(v^{jk} \frac{\partial X^t}{\partial \theta^k} \Gamma^m_{mj} + \frac{\partial v^{jk}}{\partial \theta^j} \frac{\partial X^t}{\partial \theta^k}\right) \frac{\partial}{\partial X^t} + \sum_{t=1}^n g^{rp} \frac{\partial X^h}{\partial \theta^p} v^{jk} \frac{\partial^2 X^t}{\partial \theta^j \partial \theta^k} \frac{\partial X^t}{\partial \theta^r} \frac{\partial}{\partial X^h} \\
= \;& \left(v^{jk} \frac{\partial X^t}{\partial \theta^k} \Gamma^m_{mj} + \frac{\partial v^{jk}}{\partial \theta^j} \frac{\partial X^t}{\partial \theta^k}\right) \frac{\partial}{\partial X^t} + g^{rp} v^{jk} g_{rm} \Gamma^m_{jk} \frac{\partial X^h}{\partial \theta^p} \frac{\partial}{\partial X^h} \\
= \;& \left(v^{jk} \Gamma^m_{mj} + \frac{\partial v^{jk}}{\partial \theta^j} + v^{jm} \Gamma^k_{jm}\right) \frac{\partial X^t}{\partial \theta^k} \frac{\partial}{\partial X^t} = v^{jk}_{,j} \frac{\partial X^t}{\partial \theta^k} \frac{\partial}{\partial X^t} = v^{jk}_{,j} \frac{\partial}{\partial \theta^k} = \mathrm{div}^1_1(v).
\end{aligned}$$

Following the same steps, we can obtain $\mathrm{div}_1^r(v) = \mathcal{P}C_1^r\left(\mathcal{P}_3\bar{\nabla}_{\mathbb{R}^n}V\right)$ for $r=2$. In matrix form, the divergence of a $(2,0)$ tensor field can be written as

$$\mathrm{div}_1^r(v) = \mathbf{P}\mathrm{tr}_1^r\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}(V)\right).$$

The above formula is needed when deriving approximations for the Lichnerowicz and Hodge Laplacian.

A.3 RBF Approximation for Lichnerowicz Laplacian

The Lichnerowicz Laplacian can be computed as

$$\begin{aligned} -\Delta_L u &\equiv 2\mathrm{div}_1^1(Su) = 2C_1^1(\nabla Su) = C_1^1\left[\left(g^{ij}u^k_{,is} + g^{ki}u^j_{,is}\right)\frac{\partial}{\partial\theta^j}\otimes\frac{\partial}{\partial\theta^k}\otimes \mathrm{d}\theta^s\right] \\ &= \left(g^{ij}u^k_{,ij} + g^{ki}u^j_{,ij}\right)\frac{\partial}{\partial\theta^k} = \mathrm{div}_1^1(\mathrm{grad}_g u) + \mathrm{div}_1^2(\mathrm{grad}_g u), \end{aligned} \tag{47}$$

where S denotes the symmetric tensor, defined as

$$Su = \frac{1}{2}\left(g^{ki}u^j_{,i} + g^{ij}u^k_{,i}\right)\frac{\partial}{\partial\theta^j}\otimes\frac{\partial}{\partial\theta^k}.$$

In matrix form, (47) can be written as,

$$-\Delta_L u = \mathrm{div}_1^1(\mathrm{grad}_g u) + \mathrm{div}_1^2(\mathrm{grad}_g u) = \mathbf{P}\mathrm{tr}_1^1\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}\left(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P}\right)\right) + \mathbf{P}\mathrm{tr}_1^2\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}\left(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P}\right)\right), \tag{48}$$

at each $x\in M$. One also has the following formula for the Lichnerowicz Laplacian:

$$\Delta_L u = -\mathrm{div}_1^1\left(\mathrm{grad}_g u + (\mathrm{grad}_g u)^\top\right)$$

for a vector field $u\in\mathfrak{X}(M)$. Hence, after extension to Euclidean space, the Lichnerowicz Laplacian can be written in a matrix form in the following way:

$$\begin{aligned} \bar{\Delta}_L U &= -\mathbf{P}\mathrm{tr}_1^1\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P} + (\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P})^\top)\right) \\ &= -\mathbf{P}\mathrm{tr}_1^1\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}(\mathrm{Sym}(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P}))\right), \end{aligned} \tag{49}$$

where $\mathrm{Sym}(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P}) = \mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P} + (\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P})^\top$ is a $(2,0)$ tensor field.

We now provide the non-symmetric approximation to the Lichnerowicz Laplacian acting on a vector field U. Following the notation and methodology used for the Bochner Laplacian, one can approximate Lichnerowicz Laplacian as,

$$\mathbf{U}\mapsto -\begin{bmatrix}\mathbf{H}_1\\ \vdots\\ \mathbf{H}_n\end{bmatrix}\cdot\mathrm{Sym}\begin{bmatrix}\mathbf{H}_1\mathbf{U}\\ \vdots\\ \mathbf{H}_n\mathbf{U}\end{bmatrix} := -\begin{bmatrix}\mathbf{H}_1\\ \vdots\\ \mathbf{H}_n\end{bmatrix}\cdot\mathrm{Sym}\begin{bmatrix}\bar{\mathbf{V}}_1\\ \vdots\\ \bar{\mathbf{V}}_n\end{bmatrix},$$

where $\bar{\mathbf{V}}_i := \mathbf{H}_i\mathbf{U} = [\mathbf{V}_{i1},\ldots,\mathbf{V}_{in}]^\top$ is the ith row of the tensor $\mathbf{V} = [\mathbf{V}_{ij}]_{i,j=1}^n$. The symmetric operator can be written in detail as

$$\mathrm{Sym}\begin{bmatrix}\bar{\mathbf{V}}_1\\ \vdots\\ \bar{\mathbf{V}}_n\end{bmatrix} = \mathrm{Sym}\begin{bmatrix}\mathbf{V}_{11} & \mathbf{V}_{12} & \cdots & \mathbf{V}_{1n}\\ \mathbf{V}_{21} & \mathbf{V}_{22} & \cdots & \vdots\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{V}_{n1} & \cdots & \cdots & \mathbf{V}_{nn}\end{bmatrix} = \begin{bmatrix}2\mathbf{V}_{11} & \mathbf{V}_{12}+\mathbf{V}_{21} & \cdots & \mathbf{V}_{1n}+\mathbf{V}_{n1}\\ \mathbf{V}_{21}+\mathbf{V}_{12} & 2\mathbf{V}_{22} & \cdots & \vdots\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{V}_{n1}+\mathbf{V}_{1n} & \cdots & \cdots & 2\mathbf{V}_{nn}\end{bmatrix}.$$

Then Lichnerowicz Laplacian can be written as

$$\mathbf{U} \mapsto -[\mathbf{H}_1 \left(\bar{\mathbf{V}}_1 + \mathbf{V}|_1\right) + \cdots + \mathbf{H}_n \left(\bar{\mathbf{V}}_n + \mathbf{V}|_n\right)], \tag{50}$$

where $\mathbf{V}|_i = [\mathbf{V}_{1i}, \ldots, \mathbf{V}_{ni}]^\top$ is the ith column of the tensor $\mathbf{V} = [\mathbf{V}_{ij}]_{i,j=1}^n$.

We now write the approximate Lichnerowicz Laplacian in (50) in terms of $\mathbf{U}$. Since $\bar{\mathbf{V}}_i := \mathbf{H}_i\mathbf{U}$, we only need to focus on $\mathbf{V}|_i$ for $i = 1, \ldots, n$. Since the $(2,0)$ tensor $V_{ij} = [\mathrm{grad}_g u]_{ij}$, we see that the ith column of $[V_{ij}]_{i,j=1}^n$ can be calculated as,

$$\begin{aligned}
V|_i &= \begin{bmatrix} V_{1i} \\ V_{2i} \\ \vdots \\ V_{ni} \end{bmatrix}_{n\times 1} = \begin{bmatrix} \mathcal{G}_1U^1 & \mathcal{G}_1U^2 & \cdots & \mathcal{G}_1U^n \\ \mathcal{G}_2U^1 & \mathcal{G}_2U^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathcal{G}_nU^1 & \cdots & \cdots & \mathcal{G}_nU^n \end{bmatrix}_{n\times n} \begin{bmatrix} P_{1i} \\ P_{2i} \\ \vdots \\ P_{ni} \end{bmatrix}_{n\times 1} \\
&= \begin{bmatrix} P_{1i}\mathcal{G}_1U^1 + \cdots + P_{ni}\mathcal{G}_1U^n \\ P_{1i}\mathcal{G}_2U^1 + \cdots + P_{ni}\mathcal{G}_2U^n \\ \vdots \\ P_{1i}\mathcal{G}_nU^1 + \cdots + P_{ni}\mathcal{G}_nU^n \end{bmatrix}_{n\times 1} \\
&= \begin{bmatrix} P_{1i}\mathcal{G}_1 & P_{2i}\mathcal{G}_1 & \cdots & P_{ni}\mathcal{G}_1 \\ P_{1i}\mathcal{G}_2 & P_{2i}\mathcal{G}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ P_{1i}\mathcal{G}_n & \cdots & \cdots & P_{ni}\mathcal{G}_n \end{bmatrix}_{n\times n} \begin{bmatrix} U^1 \\ U^2 \\ \vdots \\ U^n \end{bmatrix}_{n\times 1} := \mathcal{S}_iU.
\end{aligned} \tag{51}$$

After discretization using RBF approximation, above (51) can be approximated by

$$\mathbf{V}|_i = \begin{bmatrix} \mathrm{diag}\left(\mathbf{p}_{1i}\right)\mathbf{G}_1 & \mathrm{diag}\left(\mathbf{p}_{2i}\right)\mathbf{G}_1 & \cdots & \mathrm{diag}\left(\mathbf{p}_{ni}\right)\mathbf{G}_1 \\ \mathrm{diag}\left(\mathbf{p}_{1i}\right)\mathbf{G}_2 & \mathrm{diag}\left(\mathbf{p}_{2i}\right)\mathbf{G}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathrm{diag}\left(\mathbf{p}_{1i}\right)\mathbf{G}_n & \cdots & \cdots & \mathrm{diag}\left(\mathbf{p}_{ni}\right)\mathbf{G}_n \end{bmatrix}_{Nn\times Nn} \begin{bmatrix} \mathbf{U}^1 \\ \mathbf{U}^2 \\ \vdots \\ \mathbf{U}^n \end{bmatrix}_{Nn\times 1} := \mathbf{S}_i\mathbf{U},$$

where $\mathbf{p}_{ij} = (P_{ij}(\boldsymbol{x}_1), \ldots, P_{ij}(\boldsymbol{x}_N)) \in \mathbb{R}^{N\times 1}$. Then, Lichnerowicz Laplacian in (50) can be derived as,

$$\begin{aligned}
\mathbf{U} &\mapsto -[\mathbf{H}_1 \left(\bar{\mathbf{V}}_1 + \mathbf{V}|_1\right) + \cdots + \mathbf{H}_n \left(\bar{\mathbf{V}}_n + \mathbf{V}|_n\right)] \\
&= -\left[\mathbf{H}_1 \left(\mathbf{H}_1 + \mathbf{S}_1\right) + \cdots + \mathbf{H}_n \left(\mathbf{H}_n + \mathbf{S}_n\right)\right]\mathbf{U}.
\end{aligned}$$

Thus, the symmetric operator can be expressed as

$$\mathrm{Sym} \begin{bmatrix} \mathbf{H}_1\mathbf{U} \\ \vdots \\ \mathbf{H}_n\mathbf{U} \end{bmatrix} = \begin{bmatrix} \left(\mathbf{H}_1 + \mathbf{S}_1\right)\mathbf{U} \\ \vdots \\ \left(\mathbf{H}_n + \mathbf{S}_n\right)\mathbf{U} \end{bmatrix}, \tag{52}$$

and Lichnerowicz Laplacian matrix given by $-\mathbf{H}_1 \left(\mathbf{H}_1 + \mathbf{S}_1\right) - \cdots - \mathbf{H}_n \left(\mathbf{H}_n + \mathbf{S}_n\right)$ can be used to study the spectral properties of the Lichnerowicz Laplacian numerically.

A.4 RBF Approximation for Hodge Laplacian

The Hodge Laplacian is initially defined on k forms as

$$\Delta_H = (\mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}),$$

where the Hodge Laplacian is positive definite. Here, d is the exterior derivative and d^* is the adjoint of d given by

$$\mathrm{d}^* = (-1)^{dr+d+1} * \circ \mathrm{d} \circ * : \Omega^r(M) \to \Omega^{r-1}(M),$$

where $*$ is Hodge star operator. See (Chow et al., 2007) for details. The Hodge Laplacian acting on a vector field $u = u^k \frac{\partial}{\partial \theta^k}$ is defined as

$$\Delta_H u = \sharp\,(\mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d})\,\flat u,$$

where $\sharp$ and $\flat$ are the standard musical isomorphisms. First, we can compute

$$\sharp \mathrm{d}\mathrm{d}^* \flat u = \sharp \mathrm{d}(-g^{jk}u_{j,k}) = -g^{jk}u^i_{,ij}\frac{\partial}{\partial\theta^k} = -\mathrm{grad}_g\left(\mathrm{div}_g\left(u\right)\right), \tag{53}$$

where $\flat u = u_i \mathrm{d}\theta^i = g_{ij}u^j \mathrm{d}\theta^i$. Next, we can compute

$$\begin{aligned} \sharp \mathrm{d}^* \mathrm{d} \flat u &= \sharp \mathrm{d}^* \mathrm{d}\left(u_i \mathrm{d}\theta^i\right) = \sharp \mathrm{d}^* (\frac{1}{2!}\left(u_{j,i} - u_{i,j}\right)\mathrm{d}\theta^i \wedge \mathrm{d}\theta^j) = \sharp\left(-g^{ij}\left(u_{k,ij} - u_{i,kj}\right)\mathrm{d}\theta^k\right) \\ &= -g^{ij}u^k_{,ij}\frac{\partial}{\partial\theta^k} + g^{kj}u^i_{,ji}\frac{\partial}{\partial\theta^k} = -\mathrm{div}_1^1(\mathrm{grad}_g u) + \mathrm{div}_1^2(\mathrm{grad}_g u), \end{aligned} \tag{54}$$

where in last equality we have used definitions of gradient and divergence. Lastly, from (53) and (54), we can compute the Hodge Laplacian as

$$\begin{aligned} \Delta_H u &= \sharp\,(\mathrm{d}^*\mathrm{d} + \mathrm{d}\mathrm{d}^*)\,\flat u = -\left(g^{ij}u^k_{,ij} - g^{kj}u^i_{,ji} + g^{jk}u^i_{,ij}\right)\frac{\partial}{\partial\theta^k} \\ &= \Delta_B u + g^{jk}\left(u^i_{,ji} - u^i_{,ij}\right)\frac{\partial}{\partial\theta^k} = \Delta_B u + g^{jk}u^m R_{mj}\frac{\partial}{\partial\theta^k} \equiv \Delta_B u + \mathrm{Ri}(u), \end{aligned} \tag{55}$$

where the Ricci identity was used in the second line. One can see from (55) that the Hodge Laplacian Δ_H is different from the Bochner Laplacian Δ_B by a Ricci tensor $\mathrm{Ri}(u) \equiv g^{jk}u^m R_{mj}\frac{\partial}{\partial\theta^k}$ (see *pp* 478 of (Chow et al., 2007) for k-form for example).

Now we present the Hodge Laplacian written in matrix form. First, using previous formulas for the divergence and gradient, we have

$$\sharp \mathrm{d}\mathrm{d}^* \flat u = -\mathrm{grad}_g\left(\mathrm{div}_g\left(u\right)\right) = -\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}\left(\mathrm{tr}_1^1\left[\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}\mathbf{U}\right]\right).$$

Secondly, we have

$$\begin{aligned} \sharp \mathrm{d}^* \mathrm{d} \flat u &= -g^{ij}u^k_{,ij}\frac{\partial}{\partial\theta^k} + g^{kj}u^i_{,ji}\frac{\partial}{\partial\theta^k} = -\mathrm{div}_1^1(\mathrm{grad}_g u) + \mathrm{div}_1^2(\mathrm{grad}_g u) \\ &= -\mathbf{P}\mathrm{tr}_1^1\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P})\right) + \mathbf{P}\mathrm{tr}_1^2\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P})\right) \end{aligned}$$

Then, we can calculate Hodge Laplacian as

$$\begin{aligned}
\Delta_H u &= \sharp\left(\mathrm{d}^*\mathrm{d}+\mathrm{d}\mathrm{d}^*\right)\flat u = -\left(g^{ij}u^k_{,ij} - \left(g^{kj}u^i_{,ji} - g^{jk}u^i_{,ij}\right)\right)\frac{\partial}{\partial\theta^k} \\
&= -\mathrm{div}_1^1(\mathrm{grad}_g u) + \left[\mathrm{div}_1^2\left(\mathrm{grad}_g u\right) - \mathrm{grad}_g\left(\mathrm{div}_g\left(u\right)\right)\right] \\
&= -\mathbf{P}\mathrm{tr}_1^1\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P})\right) \\
&+ \left[\mathbf{P}\mathrm{tr}_1^2\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}(\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}U\mathbf{P})\right) - \mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}\left(\mathrm{tr}_1^1\left[\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}U\right]\right)\right], \qquad (56)
\end{aligned}$$

where the last line holds at each $x \in M$. The first term of (56) indeed is the Bochner Laplacian $\Delta_B u$. Hence, it only remains to compute the second and third terms. At each $x \in M$, we can write the second term as

$$\begin{aligned}
\mathbf{P}\mathrm{tr}_1^2\left(\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}\mathbf{V}\right) &= \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix}_{n\times n} \begin{bmatrix} \sum_{k=1}^n \mathcal{G}_k V_{1k} \\ \vdots \\ \sum_{k=1}^n \mathcal{G}_k V_{nk} \end{bmatrix}_{n\times 1} \\
&= \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix}_{n\times n} \begin{bmatrix} \mathcal{G}_1V_{11} + \mathcal{G}_2V_{12} + \cdots + \mathcal{G}_nV_{1n} \\ \vdots \\ \mathcal{G}_1V_{n1} + \mathcal{G}_2V_{n2} + \cdots + \mathcal{G}_nV_{nn} \end{bmatrix}_{n\times 1}, \qquad (57)
\end{aligned}$$

and write the third term as

$$\mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}\left(\mathrm{tr}_1^1\left[\mathbf{P}\bar{\nabla}_{\mathbb{R}^n}U\right]\right) = \mathbf{P}\overline{\mathrm{grad}}_{\mathbb{R}^n}\left(\sum_{k=1}^n \mathcal{G}_k U_k\right) = \begin{bmatrix} \mathcal{G}_1 \\ \vdots \\ \mathcal{G}_n \end{bmatrix}\left(\sum_{k=1}^n \mathcal{G}_k U_k\right). \qquad (58)$$

Substituting (57),(58), as well as the formula for the Bochner Laplacian, into (56), we obtain the formula for Hodge Laplacian.

Another way to write the Hodge Laplacian is given by

$$\begin{aligned}
-\Delta_H u &= \mathrm{div}_1^1(\mathrm{grad}_g u) - \mathrm{div}_1^2(\mathrm{grad}_g u) + \mathrm{grad}_g\left(\mathrm{div}_g\left(u\right)\right) \\
&= \mathrm{div}_1^1(\mathrm{grad}_g u - (\mathrm{grad}_g u)^\top) + \mathrm{grad}_g\left(\mathrm{div}_g\left(u\right)\right). \qquad (59)
\end{aligned}$$

The first term, involving the anti-symmetric part of $\mathrm{grad}_g u$, after pre-composing with the interpolating operator, can be approximated by

$$\mathrm{Ant}\begin{bmatrix} \mathbf{H}_1\mathbf{U} \\ \vdots \\ \mathbf{H}_n\mathbf{U} \end{bmatrix} = \begin{bmatrix} \left(\mathbf{H}_1 - \mathbf{S}_1\right)\mathbf{U} \\ \vdots \\ \left(\mathbf{H}_n - \mathbf{S}_n\right)\mathbf{U} \end{bmatrix}.$$

Thus, we only need to consider the RBF approximation for the last term $\mathrm{grad}_g\left(\mathrm{div}_g\left(u\right)\right)$ in (59). Using the formula for gradient of a function and divergence of a vector field in (58)

directly, we can obtain that

$$\begin{aligned}
&\operatorname{grad}_g\left(\operatorname{div}_g(u)\right)\\
=&\begin{bmatrix}\mathcal{G}_1\\ \vdots\\ \mathcal{G}_n\end{bmatrix}\sum_{i=1}^n \mathcal{G}_i U^i = \begin{bmatrix}\mathcal{G}_1\sum_{i=1}^n \mathcal{G}_i U^i\\ \vdots\\ \mathcal{G}_n\sum_{i=1}^n \mathcal{G}_i U^i\end{bmatrix} = \begin{bmatrix}\mathcal{G}_1\\ \vdots\\ \mathcal{G}_n\end{bmatrix}\begin{bmatrix}\mathcal{G}_1\cdots\mathcal{G}_n\end{bmatrix}\begin{bmatrix}U^1\\ \vdots\\ U^n\end{bmatrix}\\
=&\begin{bmatrix}\mathcal{G}_1\mathcal{G}_1 & \mathcal{G}_1\mathcal{G}_2 & \cdots & \mathcal{G}_1\mathcal{G}_n\\ \mathcal{G}_2\mathcal{G}_1 & \mathcal{G}_2\mathcal{G}_2 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \vdots\\ \mathcal{G}_n\mathcal{G}_1 & \cdots & \cdots & \mathcal{G}_n\mathcal{G}_n\end{bmatrix}_{n\times n}\begin{bmatrix}U^1\\ U^2\\ \vdots\\ U^n\end{bmatrix}_{n\times 1}.
\end{aligned}$$

After discretization, we obtain that this term can be approximated by

$$\mathbf{U}\mapsto\begin{bmatrix}\mathbf{G}_1\mathbf{G}_1 & \mathbf{G}_1\mathbf{G}_2 & \cdots & \mathbf{G}_1\mathbf{G}_n\\ \mathbf{G}_2\mathbf{G}_1 & \mathbf{G}_2\mathbf{G}_2 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \vdots\\ \mathbf{G}_n\mathbf{G}_1 & \cdots & \cdots & \mathbf{G}_n\mathbf{G}_n\end{bmatrix}_{Nn\times Nn}\begin{bmatrix}\mathbf{U}^1\\ \mathbf{U}^2\\ \vdots\\ \mathbf{U}^n\end{bmatrix}_{Nn\times 1} := \mathbf{T}\mathbf{U}.$$

It follows that the Hodge Laplacian matrix can be approximated by

$$\bar{\Delta}_H = -\left(\mathbf{H}_1\left(\mathbf{H}_1-\mathbf{S}_1\right)+\cdots+\mathbf{H}_n\left(\mathbf{H}_n-\mathbf{S}_n\right)\right)-\mathbf{T}. \tag{60}$$

A.5 RBF Approximation for Covariant Derivative

In the following, we examine $\nabla_u y = \mathcal{P}\bar{\nabla}_U Y$ via direct calculation, where the RHS will be defined below. Let $u = u^i\frac{\partial}{\partial\theta^i}\in\mathfrak{X}(M)$ and $y = y^i\frac{\partial}{\partial\theta^i}\in\mathfrak{X}(M)$ be two vector fields, where $u^i, y^i\in C^\infty(M)$ are smooth functions. Then, the covariant derivative is defined as

$$\nabla_u y = u^k y^i_{,k}\frac{\partial}{\partial\theta^i} = u^k\left(\frac{\partial y^i}{\partial\theta^k}+y^j\Gamma^i_{jk}\right)\frac{\partial}{\partial\theta^i},$$

where the covariant derivative operator $\nabla:\mathfrak{X}(M)\times\mathfrak{X}(M)\to\mathfrak{X}(M)$ takes vector fields u and y in $\mathfrak{X}(M)$ to a vector field $\nabla_u y$ in $\mathfrak{X}(M)$. We can rewrite covariant derivative as

$$\begin{aligned}
\nabla_u y &= u^k\left(\delta_{rs}g^{ij}\frac{\partial X^r}{\partial\theta^j}\frac{\partial Y^s}{\partial\theta^k}\right)\frac{\partial}{\partial\theta^i} = u^k\delta_{rs}g^{ij}\frac{\partial X^r}{\partial\theta^j}\left(\frac{\partial Y^s}{\partial X^m}\frac{\partial X^m}{\partial\theta^k}\right)\left(\frac{\partial X^p}{\partial\theta^i}\frac{\partial}{\partial X^p}\right)\\
&= \delta_{rs}\frac{\partial X^p}{\partial\theta^i}g^{ij}\frac{\partial X^r}{\partial\theta^j}\frac{\partial Y^s}{\partial X^m}U^m\frac{\partial}{\partial X^p}\\
&= \left(\delta_{rs}\frac{\partial X^p}{\partial\theta^i}g^{ij}\frac{\partial X^r}{\partial\theta^j}\mathrm{d}X^s\otimes\frac{\partial}{\partial X^p}\right)\left(U^m\frac{\partial Y^k}{\partial X^m}\frac{\partial}{\partial X^k}\right) := \mathcal{P}\bar{\nabla}_U Y,
\end{aligned}\tag{61}$$

where the first line follows from the chain rule, the second line follows from $U^m = u^k\frac{\partial X^m}{\partial\theta^k}$, and the last line defines $\bar{\nabla}_U Y\equiv U^m\frac{\partial Y^k}{\partial X^m}\frac{\partial}{\partial X^k}$ for the covariant derivative in Euclidean space. In matrix-vector form, (61) is straightforward to write as

$$\nabla_u y = \mathbf{P}\bar{\nabla}_U Y = \begin{bmatrix}P_{11} & \cdots & P_{1n}\\ \vdots & \ddots & \vdots\\ P_{n1} & \cdots & P_{nn}\end{bmatrix}\begin{bmatrix}\frac{\partial Y^1}{\partial X^1} & \cdots & \frac{\partial Y^1}{\partial X^n}\\ \vdots & \ddots & \vdots\\ \frac{\partial Y^n}{\partial X^1} & \cdots & \frac{\partial Y^n}{\partial X^n}\end{bmatrix}\begin{bmatrix}U^1\\ \vdots\\ U^n\end{bmatrix}. \tag{62}$$

The RBF approximation for covariant derivative is slightly different from above linear Laplacian operators since it is nonlinear involving two vector fields U and Y as in (62). We first compute the covariant derivative $\bar{\nabla}_U Y$ in Euclidean space. For each Y^r, its RBF interpolant we can use the RBF interpolant to evaluate the rth row of $\bar{\nabla}_U Y$ in (62) at all node locations. This yields

$$\bar{\nabla}_{\mathbf{U}} \mathbf{Y}^r \;:=\; \sum_{k=1}^{n} \frac{\partial I_{\phi_s} Y^r}{\partial X^k}\, U^k\Big|_X \in \mathbb{R}^N.$$

Concatenating all the $\bar{\nabla}_{\mathbf{U}} \mathbf{Y}^r$ for $r = 1, \ldots, n$ to form an augmented vector $\bar{\nabla}_{\mathbf{U}} \mathbf{Y} = (\bar{\nabla}_{\mathbf{U}} \mathbf{Y}^1, \ldots, \bar{\nabla}_{\mathbf{U}} \mathbf{Y}^n) \in \mathbb{R}^{Nn \times 1}$, we can obtain the RBF formula for covariant derivative as

$$\mathbf{P} \bar{\nabla}_U I_{\phi_s} Y = \mathbf{P}^{\otimes} \bar{\nabla}_{\mathbf{U}} \mathbf{Y}. \tag{63}$$

Appendix B. Interpolation Error

In this appendix, we develop results regarding interpolation error in the probabilistic setting. Namely, we formally state and prove interpolation results for functions, vector fields, and $(1,1)$ tensor fields. The setting for this section is self-contained, and can be summarized as follows. Let $X = \{x_1, \ldots, x_N\}$ be finitely many uniformly sampled data points of a closed, smooth Riemannian manifold M of dimension d, embedded in a higher dimensional Euclidean space $M \subseteq \mathbb{R}^n$. We assume additionally that the injectivity radius $\iota(M)$ is bounded away from zero from below by a constant $r > 0$.

B.1 Probabilistic Mesh Size Result

In this setting, we have the following result from (Croke, 1980):

Lemma B.1 *(Proposition 14 in (Croke, 1980)) Let $B_\delta(x)$ denote a geodesic ball of radius δ around a point $x \in M$. For $\delta < \iota(M)/2$, we have*

$$\mathit{Vol}(B_\delta(x)) \geq C(d)\delta^d,$$

where $C(d)$ is a constant depending only on the dimension d of the manifold.

Consider the quantity

$$h_{X,M} = \sup_{x \in M} \min_{x_i \in X} \|x - x_i\|_g,$$

often referred to as mesh-size. We will show that this quantity converges to 0, in high probability, after $N \to \infty$. Here, $\|\cdot\|_g$ denotes the geodesic distance.

Lemma B.2 *We have the following result regarding the mesh size $h_{X,M}$:*

$$\mathbb{P}_{X \sim \mathcal{U}} \left(h_{X,M} > \delta \right) \leq \exp(-CN\delta^d).$$

Here, $C = C(d)/\mathit{Vol}(M)$, and $\mathcal{U}$ denotes the uniform distribution on M.

Proof Suppose $h_{X,M} > \delta$, so there is $x \in M$ such that $\min_{x_i \in X} \|x - x_i\|_g \geq \delta$. So $B_\delta(x) \cap X = \emptyset$. In other words, each $x_i \in X$ is in $M \setminus B_\delta(x)$. This has measure $1 - \frac{\text{Vol}(B_\delta(x)}{\text{Vol}(M)} \leq 1 - C\delta^d$. Hence, $\min_{x_i \in X} \|x - x_i\|_g \geq \delta$ occurs with probability less than $(1 - C\delta^d)^N \leq \exp(-CN\delta^d)$. This completes the proof. ■

B.2 Interpolation of Functions, Vector Fields, and (2,0) Tensors

Given a radial basis function $\phi_s : M \times M \to \mathbb{R}$, consider the interpolation map

$$I_{\phi_s} f = \sum_{i=1}^{N} c_i \phi_s(\cdot, x_i), \tag{64}$$

where c_i's are chosen so that $I_{\phi_s} f|_X = f|_X$. In the following discussion, we assume that the kernel ϕ_s is a Mercer kernel that is at least C^2. Our probabilistic interpolation result relies heavily on Theorem 10 from (Fuselier and Wright, 2012). We adapt this theorem to our notation and state it here, for convenience.

Theorem B.1 *(adapted from Corollary 13 in Fuselier and Wright (2012)) Let M be a d-dimensional submanifold of $\mathbb{R}^n$, and let ϕ_s be a kernel with RKHS norm equivalent to a Sobolev space of order $\alpha > n/2$. Let also $2 \leq q \leq \infty$, and $0 \leq \mu \leq \alpha - n/2 + d/q - 1$. Then there exists a constant h_M such that whenever a finite node set X satisfies $h_{X,M} < h_M$, then for all $f \in H^{\alpha - \frac{(n-d)}{2}}(M)$, we have*

$$\|f - I_{\phi_s} f\|_{W^{\mu,q}(M)} \leq C h_{X,M}^{\alpha - \frac{(n-d)}{2} - \mu - n(1/2 - 1/q)} \|f\|_{H^{\alpha - \frac{(n-d)}{2}}(M)}.$$

The above result is stated to proof Lemma 4.1.

Proof of Lemma 4.1: By Lemma B.2, choosing $\delta = \left(\frac{\log N}{CN}\right)^{1/d}$, we have that $h_{X,M} = O\left(N^{-1/d}\right)$ with probability $1 - \frac{1}{N}$. For N sufficiently large, it follows that $h_{X,M} \leq h_M$, where h_M is as in Theorem B.1. Hence, the hypotheses of Theorem B.1 are satisfied (with $q = 2$ and $\mu = 0$) and we have that

$$\|I_{\phi_s} f - f\|_{L^2(M)} = O(h_{X,M}^{\alpha - (n-d)/2}).$$

Using the scaling of the mesh size with N, we obtain the desired result. ■

When the interpolator is sufficiently regular, the same result from Fuselier and Wright (2012) can be used to obtain convergence in certain Sobolev spaces. While this is not necessary for our results, we do need Sobolev norms of interpolated functions $I_{\phi_s} f$ to be bounded independent of N. We state and proof this result in the following lemma.

Lemma B.3 *Let ϕ_s be a kernel with RKHS norm equivalent to a Sobolev space of order $\alpha \geq n/2 + 3$. Then for any $f \in H^{\alpha - \frac{(n-d)}{2}}(M)$, we have*

$$\|I_{\phi_s} f\|_{W^{2,\infty}(M)} = O(1),$$

for all sufficiently large N with probability higher than $1 - \frac{1}{N}$.

Proof Let $0 \leq \mu \leq \alpha - n/2 - 1$. It follows from Theorem B.1 that there is a constant h_M such that if $h_{X,M} \leq h_M$, then for all $f \in H^{\alpha-(n-d)/2}(M)$, we have

$$\|f - I_{\phi_s} f\|_{W^{\mu,\infty}(M)} \leq h_{X,M}^{\alpha-n/2-\mu} \|f\|_{H^{\alpha-(n-d)/2}(M)}.$$

By the same argument as Lemma 4.1, this occurs with probability $1 - \frac{1}{N}$ for sufficiently large N. Since $\alpha \geq n/2 + 3$ suggests that $\alpha - n/2 - 1 \geq 2$, so $\mu = 2$ lies in the valid parameter regime, we have that

$$\|f - I_{\phi_s} f\|_{W^{2,\infty}(M)} = O(1),$$

where the constant depends on the manifold and $\|f\|_{H^{\alpha-(n-d)/2}(M)}$. Hence

$$\|I_{\phi_s} f\|_{W^{2,\infty}(M)} \leq \|f - I_{\phi_s} f\|_{W^{2,\infty}(M)} + \|f\|_{W^{2,\infty}(M)} = O(1),$$

where the above constant depends on the manifold, $\|f\|_{H^{\alpha-(n-d)/2}(M)}$, and $\|f\|_{W^{2,\infty}(M)}$. This completes the proof. ∎

For the interpolation of vector fields, recall that given a vector field $u = u^i \frac{\partial}{\partial \theta^i}$ on M, we can extend it to a smooth vector field $U = U^i \frac{\partial}{\partial X^i}$ defined on an open $\mathbb{R}^n$ neighborhood of M. Moreover, if U and V are extensions of u and v respectively, then we have that

$$\langle u, v \rangle_x = \begin{bmatrix} U^1(x) \\ U^2(x) \\ \vdots \\ U^n(x) \end{bmatrix} \cdot \begin{bmatrix} V^1(x) \\ V^2(x) \\ \vdots \\ V^n(x) \end{bmatrix},$$

at each $x \in M$, where $\langle \cdot, \cdot \rangle_x$ denotes the Riemannian inner product, and $\cdot$ denotes the standard Euclidean inner product. Recall that the interpolation of a vector field is defined component-wise in the ambient space coordinates:

$$I_{\phi_s} u = I_{\phi_s} U^i \frac{\partial}{\partial X^i},$$

where again U^i denotes the ambient space components of the extension of u to a neighborhood of M. We are now ready to prove the following lemma.

Lemma B.4 *For any $u = u^i \frac{\partial}{\partial \theta^i} \in \mathfrak{X}(M)$, we have that with probability $1 - \frac{n}{N}$,*

$$\|u - I_{\phi_s} u\|_{L^2(\mathfrak{X}(M))} = O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

Proof Notice that

$$\|u - I_{\phi_s} u\|^2_{L^2(\mathfrak{X}(M))} = \int_M \langle u - I_{\phi_s} u, u - I_{\phi_s} u \rangle_x d\mathrm{Vol}(x) = \sum_{i=1}^n \int_M (U^i(x) - I_{\phi_s} U^i(x))^2 d\mathrm{Vol}(x).$$

Using Lemma 4.1 n times, we see that with probability higher than $1 - \frac{n}{N}$,

$$\left(\int_M (U^i(x) - I_{\phi_s} U^i(x))^2 d\mathrm{Vol}(x)\right)^{1/2} = O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right). \qquad i = 1, 2, \ldots, n.$$

Hence,

$$\|u - I_{\phi_s} u\|_{L^2(\mathfrak{x}(M))} = O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

which complete the proof. ■

Similar to before, consider the norm

$$\|u\|^2_{W^{2,\infty}(\mathfrak{x}(M))} := \sum_{i=1}^n \|U^i\|^2_{W^{2,\infty}(M)},$$

where u is a vector field with ambient space coefficients U^i as before. The exact same reasoning as Lemma B.3 applied to each coefficient yields a result analogous to Lemma B.3 but for the interpolation of vector fields. We now state the result.

Lemma B.5 *Let ϕ_s be a kernel with RKHS norm equivalent to a Sobolev space of order $\alpha \geq n/2 + 3$. For any vector field u with ambient space coefficients U^i satisfying $U^i \in H^{\alpha - \frac{(n-d)}{2}}(M)$, we have*

$$\|I_{\phi_s} u\|^2_{W^{2,\infty}(\mathfrak{x}(M))} = O(1)$$

for sufficiently large N, with probability higher than $1 - \frac{n}{N}$.

Similarly, if $a = a_{ij} \frac{\partial}{\partial \theta^i} \otimes \frac{\partial}{\partial \theta^j}$ is a $(2,0)$ tensor field, we can extend a to $A = A_{ij} \frac{\partial}{\partial X^i} \otimes \frac{\partial}{\partial X^j}$, defined on an $\mathbb{R}^n$ neighborhood of M. Moreover, if a and b are $(2,0)$ tensor fields on M with extensions A and B respectively, we have that

$$\langle a, b \rangle_x = \mathrm{tr}(A(x)^\top B(x)),$$

for each $x \in M$, where $A(x), B(x)$ are thought of $n \times n$ matrices with components $A_{ij}(x), B_{ij}(x)$ respectively. We can now prove the following lemma.

Lemma B.6 *For any $a = \sum_{ij} a_{ij} \frac{\partial}{\partial \theta^i} \otimes \frac{\partial}{\partial \theta^j} \in T^{(2,0)}TM$, we have that with probability $1 - \frac{n^2}{N}$,*

$$\|a - I_{\phi_s} a\|_{L^2(T^{(2,0)}TM)} = O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

Proof Again, notice that

$$\langle a - I_{\phi_s} a, a - I_{\phi_s} a \rangle_x = \sum_{j=1}^n \sum_{i=1}^n (A_{ij}(x) - I_{\phi_s} A_{ij}(x))^2,$$

for every $x \in M$. Integrating the above over M and using Lemma 4.1 n^2 times, we see that

$$\left(\int_M \langle a - I_{\phi_s} a, a - I_{\phi_s} a \rangle_x d\mathrm{Vol}\right)^{1/2} = O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right),$$

with probability higher than $1 - \frac{n^2}{N}$. This completes the proof. ■

Appendix C. Proof of Spectral Convergence: the Laplace-Beltrami Operator

In this appendix, we study the consistency of the symmetric matrix,

$$\mathbf{G}^\top \mathbf{G} = \sum_{i=1}^{n} \mathbf{G}_i^\top \mathbf{G}_i,$$

as an approximation of the Laplace-Beltrami operator. We begin by focusing on the continuous (unrestricted) operator acting on a fixed, smooth function f, $\sum_i (\mathcal{G}_i I_{\phi_s})^* (\mathcal{G}_i I_{\phi_s}) f$ and prove its convergence to $\Delta_M f$ with high probability. Such results depend on the accuracy with which I_{ϕ_s} can approximate f and its derivatives. We then quantify the error obtained when restricting to the data set and constructing a matrix. Finally, we prove convergence of eigenvalues and eigenvectors. Since the estimator in this case is symmetric, convergence of eigenvalues requires only a weak convergence result. To prove convergence of eigenvectors, however, we need convergence of our estimator in L^2 sense.

For the discrete approximation, we consider an inner product over restricted functions which is consistent with the inner product of $L^2(M)$ as the number of data points N approaches infinity. The notation for this discrete inner product is given in Definition 11. In the remainder of this section, we will use the notation based on the following definition to account for all of the errors.

Definition 13 *Denote the L^2 norm error between $I_{\phi_s} f$ and f by ϵ_f, i.e.,*

$$\epsilon_f := \|I_{\phi_s} f - f\|_{L^2(M)}.$$

For concentration bound, we will also define a parameter $0 \leq \delta_f \leq 1$ to probabilistically characterize an upper bound for ϵ_f. For example, Lemma 4.1 states that $\epsilon_f = O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right)$, with probability higher than $1 - \delta_f$, where $\delta_f = 1/N$.

For the next spectral result, we define the formal adjoint of $\mathrm{grad}_g I_{\phi_s}$ to be the unique linear operator $(\mathrm{grad}_g I_{\phi_s})^* : \mathfrak{X}(M) \to C^\infty(M)$ satisfying

$$\langle (\mathrm{grad}_g I_{\phi_s}) f, u \rangle_{L^2(\mathfrak{X}(M))} = \langle f, (\mathrm{grad}_g I_{\phi_s})^* u \rangle_{L^2(M)}$$

for any $f \in C^\infty(M), u \in \mathfrak{X}(M)$. Before proving the spectral convergence results, we prove a number of necessary pointwise and weak convergence results.

C.1 Pointwise and Weak Convergence Results: Interpolation Error

Using Cauchy-Schwarz, paried with the fact that the formal adjoint of grad_g is $-\mathrm{div}_g$, we immediately have the following Lemma.

Lemma C.1 *Let $f \in C^\infty(M)$, and let $u \in \mathfrak{X}(M)$. Then with probability higher than $1 - \delta_f$,*

$$\left| \langle \mathrm{grad}_g f - \mathrm{grad}_g I_{\phi_s} f, u \rangle_{L^2(\mathfrak{X}(M))} \right| \leq \epsilon_f \|\mathrm{div}_g(u)\|_{L^2(M)}.$$

A little more work also yields the following.

Lemma C.2 *Let $f \in C^\infty(M)$. Then with probability higher than $1-\delta_f$,*

$$\left| \|\mathrm{grad}_g f\|^2_{L^2(\mathfrak{X}(M))} - \|\mathrm{grad}_g I_{\phi_s} f\|^2_{L^2(\mathfrak{X}(M))} \right| \le \epsilon_f \left(\|\Delta_M f\|_{L^2(M)} + \|\Delta_M I_{\phi_s} f\|_{L^2(M)} \right).$$

Proof Without ambiguity, we use the notation $\langle \cdot, \cdot \rangle$ for both inner products with respect to $L^2(M)$ and $L^2(\mathfrak{X}(M))$ in the derivations below to simplify the notation. We begin by adding and subtracting the mixed term $\langle \mathrm{grad}_g f, \mathrm{grad}_g I_{\phi_s} f \rangle$:

$$\begin{aligned} & \left| \|\mathrm{grad}_g f\|^2 - \langle \mathrm{grad}_g f, \mathrm{grad}_g I_{\phi_s} f \rangle + \langle \mathrm{grad}_g f, \mathrm{grad}_g I_{\phi_s} f \rangle - \|\mathrm{grad}_g I_{\phi_s} f\|^2 \right| \\ \le & \left| \|\mathrm{grad}_g f\|^2 - \langle \mathrm{grad}_g f, \mathrm{grad}_g I_{\phi_s} f \rangle \right| + \left| \langle \mathrm{grad}_g f, \mathrm{grad}_g I_{\phi_s} f \rangle - \|\mathrm{grad}_g I_{\phi_s} f\|^2 \right|. \end{aligned}$$

By the previous lemma, the first term is $\epsilon_f \|\Delta_M f\|_{L^2(M)}$, while the second term is $\epsilon_f \|\Delta_M I_{\phi_s} f\|_{L^2(M)}$. ■

We also need the following result for studying the convergence of eigenvectors.

Corollary C.1 *Let $f \in C^\infty(M)$. Then with probability higher than $1-\delta_f$,*

$$\|\mathrm{grad}_g f - \mathrm{grad}_g I_{\phi_s} f\|^2_{L^2(\mathfrak{X}(M))} \le \epsilon_f \left(\|\Delta_M f\|_{L^2(M)} + \|\Delta_M I_{\phi_s} f\|_{L^2(M)} \right).$$

Proof Expanding yields

$$\begin{aligned} \|\mathrm{grad}_g f - \mathrm{grad}_g I_{\phi_s} f\|^2_{L^2(\mathfrak{X}(M))} &= \langle \mathrm{grad}_g f, \mathrm{grad}_g f \rangle - \langle \mathrm{grad}_g f, \mathrm{grad}_g I_{\phi_s} f \rangle \\ &- \langle \mathrm{grad}_g I_{\phi_s} f, \mathrm{grad}_g f \rangle + \langle \mathrm{grad}_g I_{\phi_s} f, \mathrm{grad}_g I_{\phi_s} f \rangle. \end{aligned}$$

Grouping the first two and last two terms, the desired result is immediate. ■

We now prove Lemma 4.2 which states that $\Delta_M f$ can be weakly approximated.

Proof of Lemma 4.2 We again add and subtract a mixed term.

$$\begin{aligned} & \langle \Delta_M f, h \rangle - \left\langle \mathrm{grad}_g I_{\phi_s} f, \mathrm{grad}_g I_{\phi_s} h \right\rangle \\ = & \langle \Delta_M f, h \rangle - \langle \mathrm{grad}_g f, \mathrm{grad}_g I_{\phi_s} h \rangle + \langle \mathrm{grad}_g f, \mathrm{grad}_g I_{\phi_s} h \rangle - \left\langle \mathrm{grad}_g I_{\phi_s} f, \mathrm{grad}_g I_{\phi_s} h \right\rangle. \end{aligned}$$

The first two terms are bounded by $\epsilon_h \|\Delta_M f\|_{L^2(M)}$ while the second two terms are bounded by $\epsilon_f \|\Delta_M I_{\phi_s} h\|_{L^2(M)}$ with a total probability higher than $1-\delta_f-\delta_h$. Two repeated uses of Lemma 4.1 yield the final result. ■

Using similar arguments as above, we now have a convergence in L^2 norm result.

Lemma C.3 *Let $f \in C^\infty(M)$, and denote $(\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) f$ by h. With probability higher than $1-\delta_f-\delta_{\Delta_M f}-\delta_h$,*

$$\begin{aligned} \|\Delta_M f - h\|^2_{L^2(M)} &\le \epsilon_f \|\Delta_M^2 f\|_{L^2(M)} + \epsilon_{\Delta_M f} \|\Delta_M f\|_{L^2(M)} \\ &+ \epsilon_h \|\Delta_M f\|_{L^2(M)} + \epsilon_f \|\Delta_M I_{\phi_s} h\|_{L^2(M)}. \end{aligned}$$

Proof Expanding $\|\Delta_M f - (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) f\|^2_{L^2(M)}$ yields

$$\begin{aligned} & \|\Delta_M f - (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) f\|^2_{L^2(M)} \\ = & \langle \Delta_M f, \Delta_M f \rangle - \langle \Delta_M f, (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) f \rangle - \langle (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) f, \Delta_M f \rangle \\ + & \langle (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) f, (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) f \rangle. \end{aligned}$$

We use the previous Lemma twice, once on the first two terms (which gives an error of $\epsilon_{\Delta_M f}\|\Delta_M f\|_{L^2(M)} + \epsilon_f \|\Delta_M I_{\phi_s}\Delta_M f\|_{L^2(M)}$) and once on the last two terms (which gives an error of $\epsilon_h\|\Delta_M f\|_{L^2(M)} + \epsilon_f\|\Delta_M I_{\phi_s} h\|_{L^2(M)}$). This completes the proof. ■

Remark 14 *Note that since each estimated function in this section lies within the RKHS space, it follows from Lemma 4.1 that the error (denoted by ϵ with a subscript) converges with specified rate as $N \to \infty$. Moreover, Lemma B.3 ensures that no norms on the right-hand-side of the above estimates blow-up as $N \to \infty$.*

C.2 Pointwise and Weak Convergence Results: Empirical Error

We now quantify the error obtained when discretizing our estimators on the data set X. The results of this section are primarily based on the law of large numbers. First, we prove Lemma 4.3.

Proof of Lemma 4.3 Using the fact that $\text{grad}_g I_{\phi_s} f = (\mathcal{G}_1 I_{\phi_s} f, \mathcal{G}_2 I_{\phi_s} f, \dots, \mathcal{G}_n I_{\phi_s} f)^\top$ as defined in (4), it is clear that,

$$\langle \text{grad}_g I_{\phi_s} f, \text{grad}_g I_{\phi_s} f\rangle_{L^2(\mathfrak{X}(M))} = \int_M \sum_{i=1}^n \left(\mathcal{G}_i I_{\phi_s} f(x)\right)\left(\mathcal{G}_i I_{\phi_s} h(x)\right) d\text{Vol}(x),$$

and we can see immediately that the result follows from a concentration inequality on the random variable $\sum_{i=1}^n \left(\mathcal{G}_i I_{\phi_s} f(x)\right)\left(\mathcal{G}_i I_{\phi_s} h(x)\right)$. The range of I_{ϕ_s} is in $C^{\alpha-\frac{(n-d)}{2}}(M)$, and by the Assumption 4.1 along with Lemma B.3, since $\mathcal{G}_i$ are simply differential operators, we see that the random variable is bounded by a constant C (depending on the kernel ϕ_s, as well as f). By Hoeffding's inequality, we obtain,

$$\mathbb{P}_X\left(\left|\langle \mathbf{G}^\top \mathbf{G}\mathbf{f}, \mathbf{h}\rangle_{L^2(\mu_N)} - \langle \text{grad}_g I_{\phi_s} f, \text{grad}_g I_{\phi_s} h\rangle_{L^2(\mathfrak{X}(M))}\right| \geq \epsilon\right) \leq 2\text{exp}\left(\frac{-2\epsilon^2 N}{c}\right),$$

for some constant $c > 0$. Take $N^{-1} = \exp\left(\frac{-2\epsilon^2 N}{c}\right)$, solve for ϵ, the proof is complete. ■

Since $\mathbf{G}$ is simply the restricted version of $\text{grad}_g I_{\phi_s}$, the same reasoning as above gives the following result, which will be needed to prove convergence of eigenvectors.

Lemma C.4 *Let $f \in C^\infty(M)$. Let Assumption 4.1 be valid. Then*

$$\mathbb{P}_X\left(\left|\|\mathbf{G}^\top\mathbf{G}\mathbf{f} - R_N\Delta_M f\|^2_{L^2(\mu_N)} - \|(\text{grad}_g I_{\phi_s})^*(\text{grad}_g I_{\phi_s})f - \Delta_M f\|^2_{L^2(M)}\right| \geq \epsilon\right) \leq 2\text{exp}\left(\frac{-2\epsilon^2 N}{C}\right),$$

for some constant $C > 0$.

C.3 Proof of Eigenvector Convergence

We now prove the convergence of eigenvectors. The outline of this proof follows the arguments in the convergence analysis found in (Calder and Trillos, 2022) and (Peoples and Harlim, 2021). It is important to note that since the matrix $\mathbf{G}^\top\mathbf{G}$ is symmetric, there

exists an orthonormal basis of eigenvectors of $\mathbf{G}^\top\mathbf{G}$, which is key in the following proof of Theorem 4.2.

Proof Fix some $\ell \in \mathbb{N}$. For convenience, we let ϵ_{λ_ℓ} denote the error in approximating the ℓ-th eigenvalue, from the previous section. Similarly, we let δ_{λ_ℓ} denote the quantity such that eigenvalue approximation occurs with probability higher than $1-\delta_{\lambda_\ell}$. Let m be the geometric multiplicity of the eigenvalue λ_ℓ, i.e., there is an i such that $\lambda_{i+1} = \lambda_{i+2} = \cdots = \lambda_\ell = \cdots = \lambda_{i+m}$. Let

$$c_\ell = \frac{1}{2}\min\left\{|\lambda_\ell - \lambda_i|, |\lambda_\ell - \lambda_{i+m+1}|\right\}.$$

By Theorem 4.1, if $\epsilon_{\lambda_i}, \epsilon_{\lambda_{i+m+1}} < c_\ell$, then with probability $1-\delta_{\lambda_i}-\delta_{\lambda_{i+m+1}}$,

$$|\hat{\lambda}_i - \lambda_i| < c_\ell, \quad |\hat{\lambda}_{i+m+1} - \lambda_{i+m+1}| < c_\ell.$$

Let $\hat{\mathbf{u}}_1, \ldots \hat{\mathbf{u}}_N$ be an orthonormal basis of $L^2(\mu_N)$ consisting of eigenvectors of $\mathbf{G}^\top\mathbf{G}$, where $\hat{\mathbf{u}}_j$ has eigenvalue $\hat{\lambda}_j$. Let S be the m dimensional subspace of $L^2(\mu_N)$ corresponding to the span of $\{\hat{\mathbf{u}}_j\}_{j=i+1}^{i+m}$, and let P_S (resp. $P_S^\perp$) denote the projection onto S (resp. orthogonal complement of S). Let f be a norm 1 eigenfunction of Δ_M corresponding to eigenvalue λ_ℓ. Notice that

$$P_S^\perp R_N \Delta_M f = \lambda_\ell P_S^\perp R_N f = \lambda_\ell \sum_{j\neq i+1,\ldots,i+m} \langle R_N f, \hat{\mathbf{u}}_j\rangle_{L^2(\mu_N)} \hat{\mathbf{u}}_j.$$

Similarly,

$$P_S^\perp \mathbf{G}^\top\mathbf{G} R_N f = \sum_{j\neq i+1,\ldots,i+m} \hat{\lambda}_j \langle R_N f, \hat{\mathbf{u}}_j\rangle_{L^2(\mu_N)} \hat{\mathbf{u}}_j.$$

Hence,

$$\begin{aligned}
\left\| P_S^\perp R_N \Delta_M f \;-\; P_S^\perp \mathbf{G}^\top\mathbf{G} R_N f \right\|_{L^2(\mu_N)} &= \left\| \sum_{j\neq i+1,\ldots,i+m} (\lambda_\ell - \hat{\lambda}_j)\langle R_N f, \hat{\mathbf{u}}_j\rangle_{L^2(\mu_N)} \hat{\mathbf{u}}_j \right\|_{L^2(\mu_N)} \\
&\geq \min\left\{|\lambda_\ell - \hat{\lambda}_i|, |\lambda_\ell - \hat{\lambda}_{i+m+1}|\right\} \left\| \sum_{j\neq i+1,\ldots,i+m} \langle R_N f, \hat{\mathbf{u}}_j\rangle_{L^2(\mu_N)} \hat{\mathbf{u}}_j \right\|_{L^2(\mu_N)} \\
&\geq \min\left\{|\lambda_\ell - \hat{\lambda}_i|, |\lambda_\ell - \hat{\lambda}_{i+m+1}|\right\} \left\| P_S^\perp R_N f \right\|_{L^2(\mu_N)}.
\end{aligned}$$

But $P_S^\perp$ is an orthogonal projection, so

$$\begin{aligned}
\min\{|\lambda_\ell - \hat{\lambda}_i|, |\lambda_\ell - \hat{\lambda}_{i+m+1}|\} \left\| P_S^\perp R_N f \right\|_{L^2(\mu_N)} &\leq \left\| P_S^\perp R_N \Delta_M f - P_S^\perp \mathbf{G}^\top\mathbf{G} R_N f \right\|_{L^2(\mu_N)} \\
&\leq \left\| R_N \Delta_M f - \mathbf{G}^\top\mathbf{G} R_N f \right\|_{L^2(\mu_N)}.
\end{aligned}$$

Without loss of generality, assume that $\min\{|\lambda_\ell - \hat{\lambda}_i|, |\lambda_\ell - \hat{\lambda}_{i+m+1}|\} = |\lambda_\ell - \hat{\lambda}_i|$. Notice that

$$|\lambda_\ell - \hat{\lambda}_i| \geq \Big| |\lambda_\ell - \lambda_i| - |\lambda_i - \hat{\lambda}_i| \Big| > c_\ell,$$

by the hypothesis. Hence,

$$\|P_S^\perp R_N f\|^2_{L^2(\mu_N)} \leq \frac{1}{c_\ell^2} \|R_N \Delta_M f - \mathbf{G}^\top\mathbf{G} R_N f\|^2_{L^2(\mu_N)}.$$

By Lemma C.4 paired with Lemma C.3, this upper bound is smaller than

$$\epsilon_f \|\Delta_M^2 f\|_{L^2(M)} + \epsilon_{\Delta_M f}\|\Delta_M f\|_{L^2(M)} + \epsilon_h \|\Delta_M f\|_{L^2(M)} + \epsilon_f \|\Delta_M I_{\phi_s} h\|_{L^2(M)} + O\big(N^{-\frac{1}{2}}\big),$$

with probability higher than $1 - \frac{2}{N} - \delta_f - \delta_{\Delta_M f} - \delta_h$, where h is defined as in Lemma C.3. Notice that $P_S^{\perp} R_N f = R_N f - P_S R_N f$. Hence, if $\{f_1, f_2, \ldots, f_m\}$ are an orthonormal basis for the eigenspace corresponding to λ_ℓ, applying the above reasoning m times, we see that with a total probability of $1 - \frac{2}{N} - \delta_{f_1} - \delta_{\Delta_M f_1} - \delta_{h_1} - \cdots - \frac{2}{N} - \delta_{f_m} - \delta_{\Delta_M f_m} - \delta_{h_m}$,

$$\begin{aligned} \|R_N f_j - P_S R_N f_j\|^2_{L^2(\mu_N)} &\leq \frac{1}{c_\ell^2}\Big(\epsilon_{f_j} C_{f_j} \|\Delta_M^2 f_j\|_{L^2(M)} + 2\epsilon_{\Delta_M f_j}\|\Delta_M f_j\|_{L^2(M)} \\ &+ \epsilon_{h_j}\|\Delta_M f_j\|_{L^2(M)} + \epsilon_{f_j}\|\Delta_M I_{\phi_s} h_j\|_{L^2(M)} + O\big(N^{-\frac{1}{2}}\big)\Big) \\ &:= \text{Error}(j), \end{aligned} \tag{65}$$

for $j = 1, 2, \ldots m$. Let C_ℓ denote an upper bound on the essential supremum of the eigenvectors $\{f_1, f_2, \ldots, f_m\}$. For any i, j,

$$\big|f_i(x) f_j(x) - \int_M f_j(y) f_i(y) d\mu(y)\big| \leq C_\ell^2 (1 + \text{Vol}(M)).$$

Hence, using Hoeffding's inequality with $\alpha = 2\sqrt{2} C_\ell \sqrt{\frac{\log(N)}{N}}$, with probability $1 - \frac{2}{N}$,

$$\left| \frac{1}{N} \sum_{l=1}^{N} f_i(x_l) f_j(x_l) - \int_M f_i(y) f_j(y) d\mu(y) \right| < \alpha.$$

Since $\{f_1, \ldots, f_m\}$ are orthonormal in $L^2(M)$, by Hoeffding's inequality used m^2 times, we see that probability higher than $1 - \frac{2m^2}{N}$,

$$\langle R_N f_i, R_N f_j \rangle_{L^2(\mu_N)} = \delta_{ij} + O\left(\frac{1}{\sqrt{N}}\right).$$

Hence, with a total probability higher than $1 - \delta_{\lambda_i} - \delta_{\lambda_{i+m+1}} - \frac{2m^2}{N} - \frac{2}{N} - \delta_{f_1} - \delta_{\Delta_M f_1} - \delta_{h_1} - \cdots - \frac{2}{N} - \delta_{f_m} - \delta_{\Delta_M f_m} - \delta_{h_m}$, we have that

$$\begin{aligned} \langle P_S R_N f_i, P_S R_N f_j \rangle_{L^2(\mu_N)} &= \langle R_N f_i, R_N f_j \rangle_{L^2(\mu_N)} - \langle R_N f_i - P_S R_N f_i, R_N f_j - P_S R_N f_j \rangle_{L^2(\mu_N)} \\ &= \delta_{ij} + O\left(\frac{1}{\sqrt{N}}\right) + \sqrt{\text{Error}(i)}\sqrt{\text{Error}(j)}, \end{aligned}$$

where Error(j) is as defined in (65). Letting $\mathbf{v}_1 = \frac{P_S R_N f_1}{\|P_S R_N f_1\|_{L^2(\mu_N)}}$, we see that

$$\|P_S R_N f_1 - \mathbf{v}_1\|^2_{L^2(\mu_N)} = O\big(1/\sqrt{N}\big) + O\big(\text{Error}(1)\big).$$

Similarly, letting $\tilde{\mathbf{v}}_2 = P_S R_N f_2 - \frac{\langle P_S R_N f_1, P_S R_N f_2 \rangle_{L^2(\mu_N)}}{\|P_S R_N f_1\|^2_{L^2(\mu_N)}} P_S R_N f_1$ and $\mathbf{v}_2 = \frac{\tilde{\mathbf{v}}_2}{\|\tilde{\mathbf{v}}_2\|_{L^2(\mu_N)}}$, it is easy to see that

$$\|P_S R_N f_2 - \tilde{\mathbf{v}}_2\|^2_{L^2(\mu_N)} = O\big(1/\sqrt{N}\big) + O\big(\text{Error}(2)\big),$$

and hence,

$$\|P_S R_N f_2 - \mathbf{v}_2\|^2_{L^2(\mu_N)} = O\big(1/\sqrt{N}\big) + O\big(\text{Error}(2)\big).$$

Continuing in this way, we see that the Gram-Schmidt procedure on $\{P_S R_N f_j\}_{j=1}^m$, yields an orthonormal set of m vectors $\{\mathbf{v}_j\}_{j=1}^m$ spanning S such that

$$\|P_S R_N f_j - \mathbf{v}_j\|^2_{L^2(\mu_N)} = O\big(1/\sqrt{N}\big) + O\big(\text{Error}(j)\big), \qquad j = 1, 2, \ldots, m,$$

and therefore

$$\begin{aligned} \|R_N f_j - \mathbf{v}_j\|_{L^2(\mu_N)} &\leq \|R_N f_j - P_S R_N f_j\|_{L^2(\mu_N)} + \|P_S R_N f_j - \mathbf{v}_j\|_{L^2(\mu_N)} \\ &= 2\sqrt{O\big(1/\sqrt{N}\big) + O\big(\text{Error}(j)\big)}, \qquad j = 1, 2, \ldots, m. \end{aligned}$$

Therefore, for any eigenvector $\mathbf{u} = \sum_{j=1}^m b_j \mathbf{v}_j$ with $L^2(\mu_N)$ norm 1, notice that $f = \sum_{j=1}^m b_j f_j$ is a $L^2(M)$ norm 1 eigenfunction of Δ_M with eigenvalue λ_ℓ. Indeed,

$$\Delta_M f = \sum_{j=1}^m b_j \Delta_M f_j = \lambda_\ell f,$$

and

$$\|f\|^2_{L^2(M)} = \langle f, f\rangle_{L^2(M)} = \sum_{i=1}^m \sum_{j=1}^m b_i b_j \langle f_i, f_j\rangle_{L^2(M)} = \sum_{j=1}^m b_j^2 = 1,$$

where the last equality follows from the fact that

$$\|\mathbf{u}\|^2_{L^2(\mu_N)} = \|\sum_{i=1}^m b_i \mathbf{v}_i\|^2_{L^2(\mu_N)} = \sum_{i=1}^m b_i^2 = 1.$$

Moreover, the function f also satisfies

$$\|R_N f - \mathbf{u}\|^2_{L^2(\mu_N)} \leq \sum_{j=1}^m |b_j|^2 \|R_N f_j - \mathbf{v}_j\|^2_{L^2(\mu_N)} = O\big(1/\sqrt{N}\big) + \sum_{j=1}^m O\big(\text{Error}(j)\big).$$

Using Lemma 4.1 and collecting the probabilities, the above holds with probability higher than $1 - \left(\frac{2m^2+5m+24}{N}\right)$. Moreover, each Error$(j)$ is on the order of $O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right)$. Taking the square root yields the final result. This completes the proof. ∎

Appendix D. Proof of Spectral Convergence: the Bochner Laplacian

In this appendix, we discuss theoretical results concerning the Bochner Laplacian approximated by the symmetric matrix

$$\mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes} = \sum_{i=1}^n \mathbf{P}^{\otimes}\mathbf{H}_i^\top\mathbf{H}_i\mathbf{P}^{\otimes}.$$

This appendix is organized analogously to Appendix C which studies the spectral convergence of the symmetric approximation to the Laplace-Beltrami operator. Since the discussion of spectral convergence involves interpolating $(2,0)$ tensor fields and approximating the corresponding continuous inner products with discrete ones, we begin by giving a brief discussion of these details. We then investigate the continuous counterpart of the above discrete estimator, and prove its convergence in terms of interpolation error to the Bochner Laplacian in the weak sense, as well as in $L^2(\mathfrak{X}(M))$ sense. After applying law of large numbers results to quantify the error obtained by discretizing to the data set, we prove spectral convergence results.

We note that Lemma B.4, which is the analogue of Lemma 4.1 for vector fields, is especially useful for this section. Its proof is simply an application of Lemma 4.1 n times. Similarly, an interpolation of $(2,0)$ tensor-fields result is needed. This can also be found in Lemma B.6, and is essentially an application of Lemma 4.1 n^2 times. We also need the following definition, analogous to Definition 13, but in the setting of vector fields.

Definition 15 *Denote the L^2 norm error between the vector fields $I_{\phi_s}v$ and v by ϵ_v, i.e.,*

$$\epsilon_v := \|I_{\phi_s}v - v\|_{L^2(\mathfrak{X}(M)}.$$

We will also define a parameter $0 \leq \delta_v \leq 1$ to probabilistically characterize an upper bound for ϵ_v. For example, Lemma B.4 states that $\epsilon_v = O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right)$ with probability higher than $1-\delta_v$, where $\delta_v = n/N$.

D.1 Interpolation of (2,0)-Tensor Fields and Approximation of Inner Products

In the Bochner Laplacian discussion, we need to interpolate $(2,0)$ tensor fields, and approximate the corresponding continuous inner product with a discrete inner product. We outline the strategy for doing so presently. Given a $(2,0)$ tensor field $a = a_{ij}\frac{\partial}{\partial\theta^i}\otimes\frac{\partial}{\partial\theta^j}$, we can extend a to $A = A_{ij}\frac{\partial}{\partial X^i}\otimes\frac{\partial}{\partial X^j}$, defined on a neighborhood of M in $\mathbb{R}^n$, and write it as an $n\times n$ matrix in the basis $\frac{\partial}{\partial X^i}\otimes\frac{\partial}{\partial X^j}$. We define $I_{\phi_s}A$ to be the $(2,0)$ tensor field with components $I_{\phi_s}A_{ij}$ in the above ambient space basis. Recall that if $a,b\in T^{(2,0)}TM$, then by definition to the Riemannian inner product at x is given by

$$\langle a,b\rangle_x = \sum_{i,j,k,l} a_{ij}b_{kl}\left\langle\frac{\partial}{\partial\theta^i},\frac{\partial}{\partial\theta^k}\right\rangle_x\left\langle\frac{\partial}{\partial\theta^j},\frac{\partial}{\partial\theta^l}\right\rangle_x.$$

Performing a change of basis to the ambient space coordinates, a computation shows that

$$\langle a,b\rangle_x = \mathrm{tr}(A(x)^\top B(x)),$$

where A,B are the extensions of a,b and thought of as $n\times n$ matrices written w.r.t. the ambient space coordinates. Hence, defining the restriction of a $(2,0)$ tensor field $A = A_{ij}\frac{\partial}{\partial X^i}\otimes\frac{\partial}{\partial X^i}$ to be the tensor $R_NA\in\mathbb{R}^{n\times n\times N}$ with components,

$$(R_NA)_{ijk} = A_{ij}(x_k),$$

it follows that the inner product on $\mathbb{R}^{n\times n\times N}$ given by

$$\langle R_N A, R_N B\rangle_{L^2(\mu_{N,n\times n})} := \frac{1}{N}\sum_{k=1}^{N} \mathrm{tr}\left(A(x_k)^\top B(x_k)\right),$$

approximates the continuous inner product on $(2,0)$ tensor fields over M. In a previous notion of discrete $(2,0)$ tensor fields $[\mathbf{U}_1,\ldots,\mathbf{U}_n]\in\mathbb{R}^{nN\times n}$, each column $\mathbf{U}_i$ was thought of as a restricted vector field, equipped with the inner product,

$$\langle[\mathbf{U}_1,\ldots,\mathbf{U}_n],[\mathbf{V}_1,\ldots,\mathbf{V}_n]\rangle_{L^2(\mu_{N,N\times n})} := \frac{1}{N}\sum_{i=1}^{n}\mathbf{U}_i\cdot\mathbf{V}_i.$$

These two notions are equivalent in the following sense. Define $\Phi:\mathbb{R}^{nN\times n}\to\mathbb{R}^{n\times n\times N}$ by

$$\Phi\mathbf{E}_{(i-1)N+k,j} = \mathbf{E}_{i,j,k} \qquad i,j=1,2,\ldots,n \text{ and } k=1,\ldots,N,$$

where $\mathbf{E}_{(i-1)N+k,j}$ denotes the $nN\times n$ matrix with a one in entry $((i-1)N+k,j)$ and zeros elsewhere. Similarly for $\mathbf{E}_{i,j,k}$. One can easily check that Φ is an isometric isomorphism between the two inner product spaces defined above. In what follows, we will use this identification to consider $\mathbf{H}:\mathbb{R}^{nN}\to\mathbb{R}^{nN\times n}$ as a map with range in $\mathbb{R}^{n\times n\times N}$. This will be useful for computing the adjoint. Denote by $L^2(\mu_{N,n})$ the inner product on $\mathbb{R}^{nN}$ which approximates $L^2(\mathfrak{X}(M))$:

$$\langle\mathbf{U},\mathbf{V}\rangle_{L^2(\mu_{N,n})} = \frac{1}{N}\sum_{j=1}^{n}\mathbf{U}^j\cdot\mathbf{V}^j,$$

where $\mathbf{U}^j\in\mathbb{R}^N$, and $\mathbf{U}=((\mathbf{U}^1)^\top,\ldots,(\mathbf{U}^n)^\top)^\top$. Using the identification above, it is simple to check that the transpose of $\Phi\mathbf{H}:\mathbb{R}^{nN}\to\mathbb{R}^{n\times n\times N}$ with respect to inner products defined above is given by,

$$[\tilde{\mathbf{U}}^1,\ldots,\tilde{\mathbf{U}}^n]\mapsto[\mathbf{U}^1,\ldots,\mathbf{U}^n]\mapsto\sum_i\mathbf{H}_i^\top\mathbf{U}^i,$$

where $[\tilde{\mathbf{U}}^1,\ldots,\tilde{\mathbf{U}}^n]\in\mathbb{R}^{n\times n\times N}$. Since Φ is an isometric isomorphism, $\mathbf{H}^\top\Phi^\top\Phi\mathbf{H}=\mathbf{H}^\top\mathbf{H}$. This shows that $\mathbf{H}^\top\mathbf{H}:\mathbb{R}^{nN}\to\mathbb{R}^{nN}$ is indeed given by the formula,

$$\mathbf{H}^\top\mathbf{H}=\sum_i\mathbf{H}_i^\top\mathbf{H}_i,$$

which will be used extensively in the following calculations.

D.2 Pointwise and Weak Convergence Results: Interpolation Error

Using Cauchy-Schwarz, along with the fact that the formal adjoint of grad_g acting on vector fields is $-\mathrm{div}_1^1$ as defined in (19), we immediately have the following result.

Lemma D.1 *Let $u\in\mathfrak{X}(M)$, and let $a\in T^{(2,0)}TM$. Then with probability higher than $1-\delta_u$,*

$$\left|\langle\mathrm{grad}_g u-\mathrm{grad}_g I_{\phi_s}u,a\rangle_{L^2(T^{(2,0)}TM)}\right|\le\epsilon_u\|\mathrm{div}_1^1(a)\|_{L^2(\mathfrak{X}(M))}.$$

We also have the following norm result.

Corollary D.1 *Let $u \in \mathfrak{X}(M)$. Then with probability higher than $1 - \delta_u$,*

$$\|\mathrm{grad}_g u - \mathrm{grad}_g I_{\phi_s} u\|^2_{L^2(T^{(2,0)}TM)} \leq \epsilon_u \left(\|\Delta_B u\|_{L^2(\mathfrak{X}(M))} + \|\Delta_B I_{\phi_s} u\|_{L^2(\mathfrak{X}(M))} \right).$$

Proof Note that,

$$\begin{aligned} \|\mathrm{grad}_g u - \mathrm{grad}_g I_{\phi_s} u\|^2 &= \langle \mathrm{grad}_g u, \mathrm{grad}_g u \rangle - \langle \mathrm{grad}_g u, \mathrm{grad}_g I_{\phi_s} u \rangle - \langle \mathrm{grad}_g I_{\phi_s} u, \mathrm{grad}_g u \rangle \\ &+ \langle \mathrm{grad}_g I_{\phi_s} u, \mathrm{grad}_g I_{\phi_s} u \rangle. \end{aligned}$$

Grouping the first two and last two terms, the desired result is immediate. ■

Using the same reasoning as before, we can deduce a weak convergence result.

Lemma D.2 *Let $v, w \in \mathfrak{X}(M)$. Then with probability higher than $1 - \delta_v - \delta_w$,*

$$\begin{aligned} & \left| \langle \Delta_B v, w \rangle_{L^2(\mathfrak{X}(M))} - \langle \mathrm{grad}_g I_{\phi_s} v, \mathrm{grad}_g I_{\phi_s} w \rangle_{L^2(T^{(2,0)}TM)} \right| \\ \leq \; & \epsilon_w \|\Delta_B v\|_{L^2(\mathfrak{X}(M))} + \epsilon_v \|\Delta_B I_{\phi_s} w\|_{L^2(\mathfrak{X}(M))}. \end{aligned}$$

Proof We again add and subtract a mixed term,

$$\langle \Delta_B v, w \rangle - \langle \mathrm{grad}_g v, \mathrm{grad}_g I_{\phi_s} w \rangle + \langle \mathrm{grad}_g v, \mathrm{grad}_g I_{\phi_s} w \rangle - \left\langle \mathrm{grad}_g I_{\phi_s} v, \mathrm{grad}_g I_{\phi_s} w \right\rangle.$$

The first two terms are bounded by $\epsilon_w \|\Delta_B v\|_{L^2(\mathfrak{X}(M))}$ while the second two terms are bounded by $\epsilon_v \|\Delta_B I_{\phi_s} w\|_{L^2(\mathfrak{X}(M))}$. ■

Similarly, we can derive an L^2 convergence result. First, similar to the previous section, let us denote the formal adjoint of $\mathrm{grad}_g I_{\phi_s} : \mathfrak{X}(M) \to T^{(2,0)}TM$ by $(\mathrm{grad}_g I_{\phi_s})^* : T^{(2,0)}TM \to \mathfrak{X}(M)$. The exact same proof as before yields the following Lemma.

Lemma D.3 *Let $v \in \mathfrak{X}(M)$, and denote by w the vector field $(\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) v$. With probability higher than $1 - \delta_v - \delta_{\Delta_B v} - \delta_w$,*

$$\begin{aligned} \|\Delta_B v - w\|^2 \leq \epsilon_{\Delta_B v} \|\Delta_B v\| &+ \epsilon_v \|\Delta_B I_{\phi_s} \Delta_B v\| \\ \epsilon_w \|\Delta_B v\| &+ \epsilon_v \|\Delta_B I_{\phi_s} \mathrm{grad}_g w\|. \end{aligned}$$

Proof Expanding $\|\Delta_B v - (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) v\|^2$ yields

$$\begin{aligned} \|\Delta_B v - (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) v\|^2 &= \langle \Delta_B v, \Delta_B v \rangle - \langle \Delta_B v, (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) v \rangle \\ &- \langle (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) v, \Delta_B v \rangle \\ &+ \langle (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) v, (\mathrm{grad}_g I_{\phi_s})^* (\mathrm{grad}_g I_{\phi_s}) v \rangle. \end{aligned}$$

We use the previous Lemma twice, once on the first two terms (which gives an error of $\epsilon_{\Delta_B v} \|\Delta_B v\| + \epsilon_v \|\Delta_B I_{\phi_s} \Delta_B v\|$) and once on the last two terms (which gives an error of $\epsilon_w \|\Delta_B v\| + \epsilon_v \|\Delta_B I_{\phi_s} \mathrm{grad}_g w\|$). This completes the proof. ■

Remark 16 *Similar to Remark 14 for the Laplace-Beltrami operator, we note that Lemma B.4 guarantees all estimation error terms (errors indicated by ϵ with corresponding subscript) converge as $N \to \infty$. Additionally, it follows from Lemma B.5, since the Bochner Laplacian is obtained from second order differential operators on the ambient space coefficients of each vector field, that no norm terms on the right-hand-side of each estimate above blow up as $N \to \infty$.*

D.3 Pointwise and Weak Convergence Results: Empirical Error

The following is a simple concentration result.

Lemma D.4 *Let $u, v \in \mathfrak{X}(M)$. Additionally, let Assumption 4.1 be valid. Then*

$$\mathbb{P}\left(\left|\langle \mathbf{H}\mathbf{P}^{\otimes} R_N u, \mathbf{H}\mathbf{P}^{\otimes} R_N v\rangle_{L^2(\mu_{N,n\times n})} - \langle \mathrm{grad}_g I_{\phi_s} u, \mathrm{grad}_g I_{\phi_s} v\rangle_{L^2(T^{(2,0)}TM)}\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-2\epsilon^2 N}{C}\right),$$

for some constant $C > 0$.

Proof We note that

$$\begin{aligned}\langle \mathrm{grad}_g I_{\phi_s} u, \mathrm{grad}_g I_{\phi_s} v\rangle &= \int_M \mathrm{tr}\left([\mathcal{H}_1 U, \dots \mathcal{H}_n U]^\top [\mathcal{H}_1 V, \dots \mathcal{H}_n V]\right) d\mathrm{Vol}(x) \\ &= \int_M \sum_{i=1}^n \langle \mathcal{H}_i U, \mathcal{H}_i V\rangle_x \, d\mathrm{Vol}(x) = \int_M \sum_{i=1}^n \langle \mathcal{H}_i \mathbf{P} U, \mathcal{H}_i \mathbf{P} V\rangle_x \, d\mathrm{Vol}(x),\end{aligned}$$

where U and V are extensions of u, v onto an $\mathbb{R}^n$ neighborhood of M. The last equality comes from the fact that since U, V extend u, v, then at each point $x \in M$, we have $\mathbf{P}U = U, \mathbf{P}V = v$. We can see immediately that the result follows by using a concentration inequality on the random variable $\sum_{i=1}^n \langle \mathcal{H}_i \mathbf{P} U, \mathcal{H}_i \mathbf{P} V\rangle_x$. Plugging in to Hoeffding's inequality and using smoothness assumptions, along with Assumption 4.1 and Lemma B.5, gives the desired result. ■

Again, $\mathbf{H}$ is simply the restricted version of $\mathrm{grad}_g I_{\phi_s}$, as defined in (17). Hence, the same reasoning as above yields the norm convergence result, which plays a part in proving the convergence of eigenvectors.

Lemma D.5 *Let $v \in \mathfrak{X}(M)$. Additionally, let Assumption 4.1 be valid. Then*

$$\mathbb{P}\left(\left|\|\mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes} v - R_N \Delta_B v\|^2_{L^2(\mu_N,n)} - \|(\mathrm{grad}_g I_{\phi_s})^*(\mathrm{grad}_g I_{\phi_s}) v - \Delta_B v\|^2_{L^2(\mathfrak{X}(M))}\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-2\epsilon^2 N}{C}\right),$$

for some constant $C > 0$.

D.4 Proof of Spectral Convergence

Here, we prove Theorem 4.3, the convergence of eigenvalue result for the Bochner Laplacian operator.

Proof Enumerate the eigenvalues of $\mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes}$ and label them $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_N$. Let $\mathcal{S}_i' \subseteq \mathfrak{X}(M)$ denote an i-dimensional subspace of smooth functions on which the quantity $\max_{v\in\mathcal{S}_i} \frac{\langle \mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes} R_N v, R_N v\rangle_{L^2(\mu_{N,n})}}{\|R_N v\|_{L^2(\mu_{N,n})}}$ achieves its minimum. Let $\tilde{v} \in \mathcal{S}_i'$ be the smooth vector field on which the maximum $\max_{v\in\mathcal{S}_i'}\langle \Delta_B v, v\rangle$ occurs. WLOG, assume that $\|\tilde{v}\|^2_{L^2(\mathfrak{X}(M))} = 1$. Assume that N is sufficiently large so that by Hoeffding's inequality $\left|\|R_N\tilde{v}\|^2_{L^2(\mu_{N,n})} - 1\right| \leq \frac{\mathrm{Const}}{\sqrt{N}} \leq 1/2$, with probability $1 - \frac{2}{N}$, and thus $\|R_N\tilde{v}\|^2_{L^2(\mu_{N,n})}$ is bounded away from zero. Hence, we can Taylor expand the denominator term of $\frac{\langle \mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes} R_N\tilde{v}, R_N\tilde{v}\rangle_{L^2(\mu_{N,n})}}{\|R_N\tilde{v}\|^2_{L^2(\mu_{N,n})}}$ to obtain

$$\frac{\langle \mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes} R_N\tilde{v}, R_N\tilde{v}\rangle_{L^2(\mu_{N,n})}}{\|R_N\tilde{v}\|^2_{L^2(\mu_{N,n})}} = \langle \mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes} R_N\tilde{v}, R_N\tilde{v}\rangle_{L^2(\mu_{N,n})} - \frac{\mathrm{Const}\langle \mathbf{P}^{\otimes}\mathbf{H}^\top\mathbf{H}\mathbf{P}^{\otimes} R_N\tilde{v}, R_N\tilde{v}\rangle_{L^2(\mu_{N,n})}}{\sqrt{N}}.$$

Note that with probability higher than $1 - \frac{2}{N}$, we have that

$$\left| \langle \mathbf{P}^{\otimes} \mathbf{H}^{\top} \mathbf{H} \mathbf{P}^{\otimes} R_N \tilde{v}, R_N \tilde{v} \rangle_{L^2(\mu_{N,n})} - \langle \mathrm{grad}_g I_{\phi_s} \tilde{v}, \mathrm{grad}_g I_{\phi_s} \tilde{v} \rangle_{L^2(T^{(1,1)}(TM))} \right| = O\left(N^{-\frac{1}{2}}\right),$$

by Lemma D.4, choosing $\epsilon = \sqrt{\frac{\log(N)}{N^{1/2}}}$ and ignoring log factors. Combining with Lemma D.2 and plugging in the result from Lemma B.4, we obtain that

$$\langle \Delta_B \tilde{v}, \tilde{v} \rangle_{L^2(\mathfrak{X}(M))} \leq \frac{\langle \mathbf{P}^{\otimes} \mathbf{H}^{\top} \mathbf{H} \mathbf{P}^{\otimes} R_N \tilde{v}, R_N \tilde{v} \rangle_{L^2(\mu_{N,n})}}{\| R_N \tilde{v} \|^2_{L^2(\mu_{N,n})}} + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right),$$

with total probability higher than $1 - \frac{2n+2}{N}$. Since $\tilde{v} = \arg\max_{v \in \mathcal{S}_i'} \langle \Delta_B v, v \rangle_{L^2(\mathfrak{X}(M))}$, and since,

$$\frac{\langle \mathbf{P}^{\otimes} \mathbf{H}^{\top} \mathbf{H} \mathbf{P}^{\otimes} R_N \tilde{v}, R_N \tilde{v} \rangle_{L^2(\mu_{N,n})}}{\| R_N v \|_{L^2(\mu_{N,n})}} \leq \max_{v \in \mathcal{S}_i'} \frac{\langle \mathbf{P}^{\otimes} \mathbf{H}^{\top} \mathbf{H} \mathbf{P}^{\otimes} R_N v, R_N v \rangle_{L^2(\mu_{N,n})}}{\| R_N v \|_{L^2(\mu_{N,n})}},$$

we have the following:

$$\max_{v \in \mathcal{S}_i'} \langle \Delta_B v, v \rangle_{L^2(\mathfrak{X}(M))} \leq \max_{v \in \mathcal{S}_i'} \frac{\langle \mathbf{P}^{\otimes} \mathbf{H}^{\top} \mathbf{H} \mathbf{P}^{\otimes} R_N v, R_N v \rangle_{L^2(\mu_{N,n})}}{\| R_N v \|_{L^2(\mu_{N,n})}} + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{-\frac{2\alpha+(n-d)}{2d}}\right).$$

Since $\mathcal{S}_i'$ is the exact subspace on which $\max_{v \in \mathcal{S}_i} \frac{\langle \mathbf{P}^{\otimes} \mathbf{H}^{\top} \mathbf{H} \mathbf{P}^{\otimes} R_N v, R_N v \rangle_{L^2(\mu_{N,n})}}{\| R_N v \|_{L^2(\mu_{N,n})}}$ achieves its minimum, then

$$\max_{v \in \mathcal{S}_i'} \langle \Delta_B v, v \rangle_{L^2(\mathfrak{X}(M))} \leq \hat{\lambda}_i + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

But the left-hand-side certainly bounds above the minimum of $\max_{v \in \mathcal{S}_i} \langle \Delta_B v, v \rangle_{L^2(\mathfrak{X}(M))}$ over all i-dimensional smooth subspaces $\mathcal{S}_i$. Hence,

$$\lambda_i \leq \hat{\lambda}_i + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right).$$

The same argument yields that $\hat{\lambda}_i \leq \lambda_i + O\left(N^{-\frac{1}{2}}\right) + O\left(N^{\frac{-2\alpha+(n-d)}{2d}}\right)$, with probability higher than $1 - \frac{2+2n}{N}$. This completes the proof. ■

Remark 17 *Since the exact same proof as the eigenvector convergence for the Laplace Beltrami operator is valid, we do not rewrite it in full detail. Instead, we simply note that while Theorem 4.2 is proved using Theorem 4.1, as well as Lemmas 4.1, C.3, and C.4, similarly, we have that Theorem 4.4 can be proved with the exact same argument, using instead Theorem 4.3, along with Lemmas B.4, D.3, and D.5. Statements and proofs of the lemmas mentioned above can be found in the present section.*

Appendix E. Additional Numerical Results with Diffusion Maps

In this section, we report additional numerical results with diffusion maps on the 2D torus example to support the conclusion in Section 5.2. Particularly, we would like to demonstrate that other choices of nearest neighbor parameter, K, and kernel bandwidth parameter, ϵ, do not improve the accuracy of the estimates, and thus, the particular choice of K and ϵ in Section 5.2 is representative.

Figure 16 shows the eigenvector estimates for mode $k = 13$. Each row in these figures corresponds to manually tuned ϵ for a fixed K. In the last row, we also include the auto-tuned ϵ (using the method that was originally proposed in Coifman et al. (2008)). One can see that the auto-tuned ϵ seems to only work for smaller values of K. The denser the matrix, the estimates seem to be more sensitive to the choice of ϵ. On the other hand, one can also see that qualitatively the eigenvector estimates for various choices of K and ϵ (including $K = N$) are not very different from those with smaller values of K. In Figure 17, we give further details on errors in the estimation of eigenvalues and eigenvectors for each K and ϵ. As a baseline, we plot the errors corresponding to SRBF $\hat{\mathbf{P}}$. Overall, one can see that when K is small, the leading eigenvalues can be more accurately estimated by DM with an appropriate choice of ϵ. For larger K (e.g. $K = 400, 2500$, see panel (d)), one can see $\epsilon = 0.05$ leads to accurate estimation of only the leading eigenvalues (red curves) but for $\epsilon = 0.2$ the non-leading spectra becomes more accurate in the expense of less accurate leading eigenvalue estimation. At the same time, the accuracy in the estimation of the corresponding eigenfunction is more or less similar for any choice of K, confirming the qualitative results shown in Figure 16. Indeed if one chooses $K = 400$ and $\epsilon = 0.2$, while the errors of the estimation of eigenvalues are smaller than those of SRBF $\hat{\mathbf{P}}$ (compare grey and cyan curves in panel (d) in Figure 17), the corresponding eigenvector estimate for mode $k = 13$ does not qualitatively reflect on an accurate estimation at all (see row 4, column 5 in Figure 16).

Based on the results in Figure 17 panel (f), Notice that the best estimate corresponds to the case of $K = 100$ (red curves). Based on this empirical result, we present the case $K = 100$ with auto-tuned ϵ in Figures 4-6.

References

Douglas N Arnold, Richard S Falk, and Ragnar Winther. Finite element exterior calculus, homological techniques, and applications. *Acta numerica*, 15:1–155, 2006.

Douglas N Arnold, Richard S Falk, and Ragnar Winther. Finite element exterior calculus: from Hodge theory to numerical stability. *Bulletin of the American mathematical society*, 47(2):281–354, 2010.

Omri Azencot, Maks Ovsjanikov, Frederic Chazal, and Mirela Ben-Chen. Discrete derivatives of vector fields on surfaces - an operator approach. *Acm Transactions on Graphics*, 34(3):1–13, 2015.

Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

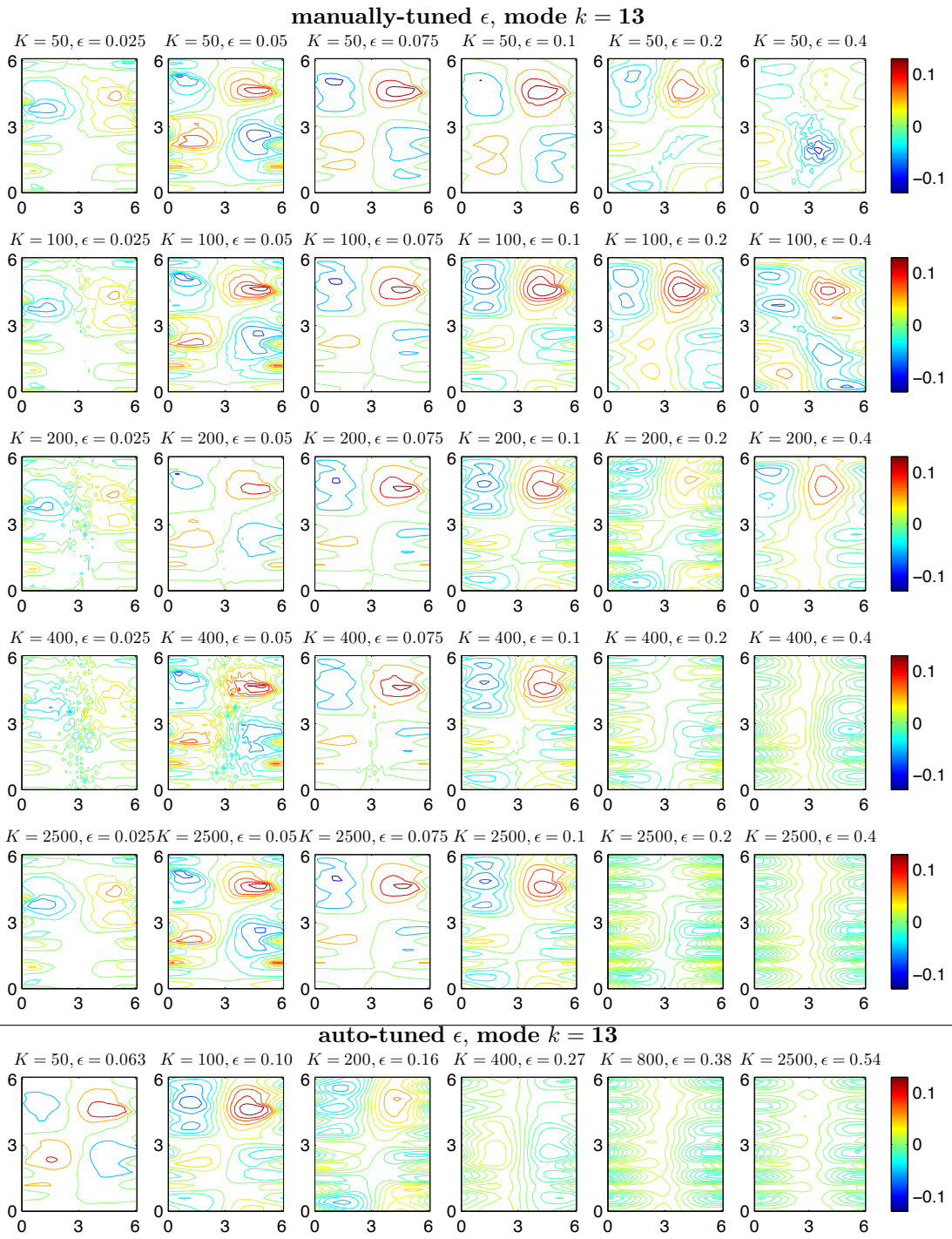

Figure 16: **2D general torus in $\mathbb{R}^{21}$.** Mode $k = 13$. Comparison of eigenfunctions of Laplace -Beltrami among various DMs with different K-nearest neighbors and bandwidth ϵ. Panels from the first row to the fifth row correspond to manually-tuned ϵ for $K = 50, 100, 200, 400, 2500$, respectively. Panels in the last row correspond to auto-tuned ϵ for various K. The randomly distributed $N = 2500$ data points on the manifold are used for computing the eigenvalue problem.

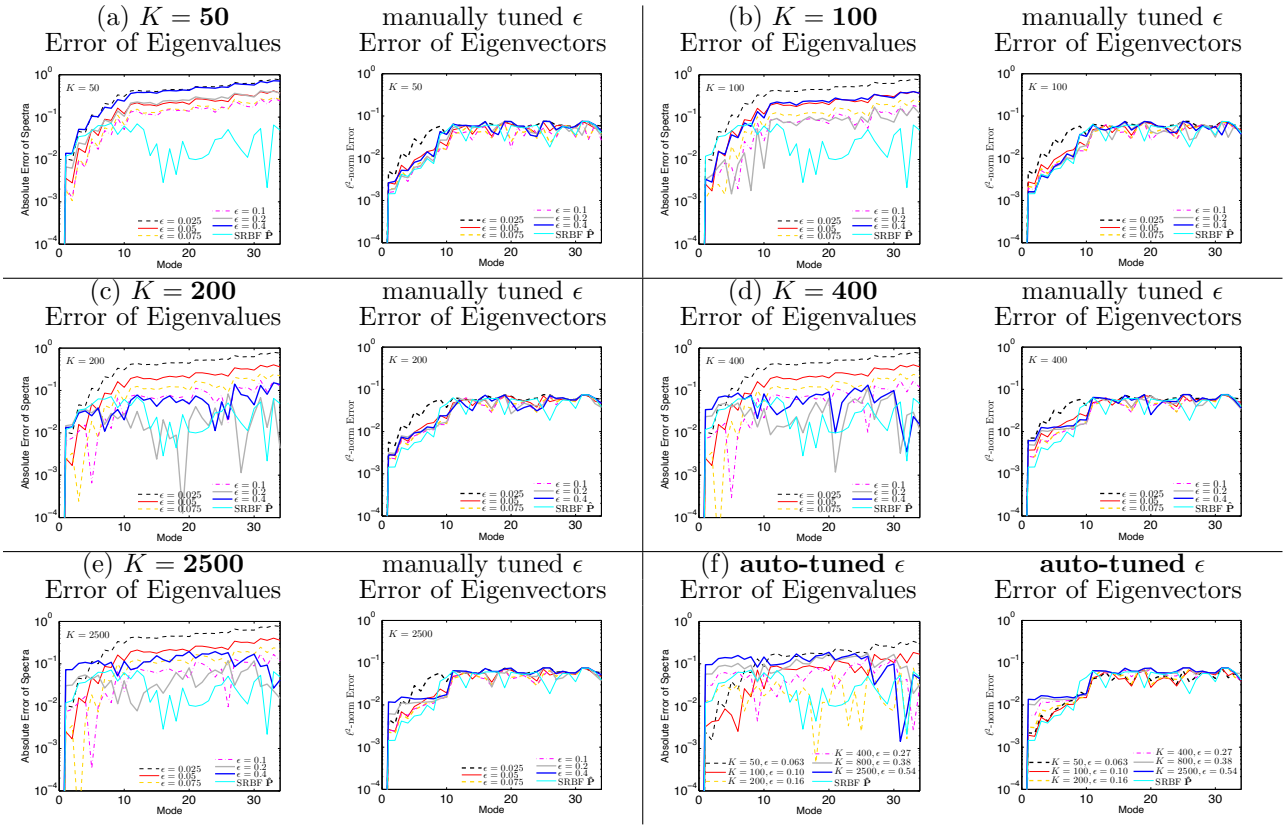

Figure 17: **2D general torus in $\mathbb{R}^{21}$.** Comparison of errors of eigenvalues and errors of eigenvectors for DM among various parameters. Panels (a)-(e) correspond to manually-tuned ϵ for $K = 50, 100, 200, 400, 2500$, respectively. Panel (f) corresponds to auto-tuned ϵ for various K. The randomly distributed $N = 2500$ data points on the manifold are used for solving the eigenvalue problem.

Mikhail Belkin and Partha Niyogi. Convergence of Laplacian eigenmaps. *Advances in Neural Information Processing Systems*, 19:129, 2007.

Tyrus Berry. MATLAB Code for Diffusion Forecast and Spectral Exterior Calculus, 2018. URL `https://math.gmu.edu/~berry/Code/SEC.zip`.

Tyrus Berry and Dimitrios Giannakis. Spectral exterior calculus. *Communications on Pure and Applied Mathematics*, 73(4):689–770, 2020.

Martin D Buhmann. *Radial basis functions: theory and implementations*, volume 12. Cambridge university press, 2003.

Dmitri Burago, Sergei Ivanov, and Yaroslav Kurylev. A graph discretization of the Laplace Beltrami operator. *Journal of Spectral Theory*, 4(4):675–714, 2014.

Jeff Calder and Nicolás García Trillos. Improved spectral convergence rates for graph Laplacians on ϵ-graphs and k-NN graphs. *Applied and Computational Harmonic Analysis*, 60:123–175, 2022.

Bennett Chow, Sun-Chin Chu, David Glickenstein, Christine Guenther, Jim Isenberg, Tom Ivey, Dan Knopf, Peng Lu, Feng Luo, and Lei Ni. *The Ricci flow: techniques and applications*. American Mathematical Society Providence, 2007.

Andreas Christmann and Ingo Steinwart. *Support vector machines*. Springer, 2008.

Fan R.K. Chung. *Spectral graph theory*. Number 92. American Mathematical Soc., 1997.

Ronald R Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006.

Ronald R Coifman, Yoel Shkolnisky, Fred J Sigworth, and Amit Singer. Graph Laplacian tomography from unknown random projections. *Image Processing, IEEE Transactions on*, 17(10):1891–1899, 2008.

Matthew J. Colbrook and Alex Townsend. Rigorous data-driven computation of spectral properties of koopman operators for dynamical systems. *Communications on Pure and Applied Mathematics*, 77(1):221–283, 2024.

Keenan Crane, Mathieu Desbrun, and Peter Schröder. Trivial connections on discrete surfaces. In *Computer Graphics Forum*, volume 29, pages 1525–1533. Wiley Online Library, 2010.

Christopher B Croke. Some isoperimetric inequalities and eigenvalue estimates. In *Annales scientifiques de l'École normale supérieure*, volume 13, pages 419–435, 1980.

Felipe Cucker and Steve Smale. On the mathematical foundations of learning. *Bulletin of the American Mathematical Society*, 39(1):1–49, 2002.

Edward B Davies and Michael Plum. Spectral pollution. *IMA journal of numerical analysis*, 24(3):417–438, 2004.

James W Demmel. *Applied numerical linear algebra*. SIAM, 1997.

Manfredo Perdigao Do Carmo and J Flaherty Francis. *Riemannian geometry*, volume 6. Springer, 1992.

David L Donoho and Carrie Grimes. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National Academy of Sciences*, 100 (10):5591–5596, 2003.

David B Dunson, Hau-Tieng Wu, and Nan Wu. Spectral convergence of graph Laplacian and heat kernel reconstruction in L^∞ from random samples. *Applied and Computational Harmonic Analysis*, 55:282–336, 2021.

Gerhard Dziuk and Charles M Elliott. Surface finite elements for parabolic equations. *Journal of Computational Mathematics*, pages 385–407, 2007.

Gerhard Dziuk and Charles M Elliott. Finite element methods for surface PDEs. *Acta Numerica*, 22:289–396, 2013.

Gregory E Fasshauer. *Meshfree approximation methods with MATLAB*, volume 6. World Scientific, 2007.

Gregory E Fasshauer and Michael J McCourt. Stable evaluation of Gaussian radial basis function interpolants. *SIAM Journal on Scientific Computing*, 34(2):A737–A762, 2012.

Natasha Flyer and Fornberg Bengt. Solving PDEs with radial basis functions. *Acta Numerica*, 2015.

Gerald B Folland. Harmonic analysis of the de Rham complex on the sphere. *Journal Für Die Reine Und Angewandte Mathematik*, 1989(398), 1989.

Bengt Fornberg and Erik Lehto. Stabilization of RBF-generated finite difference methods for convective PDEs. *Journal of Computational Physics*, 230(6):2270–2285, 2011.

Bengt Fornberg and Grady B Wright. Stable computation of multiquadric interpolants for all values of the shape parameter. *Computers & Mathematics with Applications*, 48(5-6): 853–867, 2004.

Bengt Fornberg, Elisabeth Larsson, and Natasha Flyer. Stable computations with Gaussian radial basis functions. *SIAM Journal on Scientific Computing*, 33(2):869–892, 2011.

Edward J Fuselier and Grady B Wright. Stability and error estimates for vector field interpolation and decomposition on the sphere with RBFs. *SIAM Journal on Numerical Analysis*, 47(5):3213–3239, 2009.

Edward J Fuselier and Grady B Wright. Scattered data interpolation on embedded submanifolds with restricted positive definite kernels: Sobolev error estimates. *SIAM Journal on Numerical Analysis*, 50(3):1753–1776, 2012.

Edward J Fuselier and Grady B Wright. A high-order kernel method for diffusion and reaction-diffusion equations on surfaces. *Journal of Scientific Computing*, 56(3):535–565, 2013.

Andrew Gillette, Michael Holst, and Yunrong Zhu. Finite element exterior calculus for evolution problems. *Journal of Computational Mathematics*, 35(2):187–212, 2017.

Ben J Gross, Nathaniel Trask, Paul Kuberry, and Paul J Atzberger. Meshfree methods on manifolds for hydrodynamic flows on curved surfaces: A Generalized Moving Least-Squares (GMLS) approach. *Journal of Computational Physics*, 409:109340, 2020.

Anil Nirmal Hirani. *Discrete exterior calculus*. California Institute of Technology, 2003.

Jürgen Jost and Jeurgen Jost. *Riemannian geometry and geometric analysis*, volume 42005. Springer, 2008.

Motonobu Kanagawa, Philipp Hennig, Dino Sejdinovic, and Bharath K Sriperumbudur. Gaussian processes and kernel methods: A review on connections and equivalences. *arXiv preprint arXiv:1807.02582*, 2018.

Edward J Kansa. Multiquadrics—a scattered data approximation scheme with applications to computational fluid-dynamics—ii solutions to parabolic, hyperbolic and elliptic partial differential equations. *Computers & mathematics with applications*, 19(8-9):147–161, 1990.

Tosio Kato. *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media, 2013.

John M Lee. *Introduction to Riemannian manifolds*. Springer, 2018.

Erik Lehto, Varun Shankar, and Grady B Wright. A radial basis function (RBF) compact finite difference (FD) scheme for reaction-diffusion equations on surfaces. *SIAM Journal on Scientific Computing*, 39(5):A2129–A2151, 2017.

Mathieu Lewin and Éric Séré. Spectral pollution and how to avoid it. *Proceedings of the London mathematical society*, 100(3):864–900, 2010.

Jian Liang and Hongkai Zhao. Solving partial differential equations on point clouds. *SIAM Journal on Scientific Computing*, 35(3):A1461–A1486, 2013.

Anna V Little, Jason Lee, Yoon-Mo Jung, and Mauro Maggioni. Estimation of intrinsic dimensionality of samples from noisy low-dimensional manifolds in high dimensions with multiscale svd. In *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, pages 85–88. IEEE, 2009.

X Llobet, K Appert, A Bondeson, and J Vaclavik. On spectral pollution. *Computer physics communications*, 59(2):199–216, 1990.

Don O Loftsgaarden and Charles P Quesenberry. A Nonparametric Estimate of a Multivariate Density Function. *The Annals of Mathematical Statistics*, 36(3):1049 – 1051, 1965.

Shigeyuki Morita. *Geometry of differential forms*. Number 201. American Mathematical Soc., 2001.

Francis J Narcowich and Joseph D Ward. Norms of inverses and condition numbers for matrices associated with scattered data. *Journal of Approximation Theory*, 64(1):69–94, 1991.

Andrew Y Ng, Michael I Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, pages 849–856, 2002.

Emanuel Parzen. On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076, 1962.

John W Peoples and John Harlim. Spectral Convergence of Symmetrized Graph Laplacian on manifolds with boundary. *arXiv preprint arXiv:2110.06988*, 2021.

Cécile Piret. The orthogonal gradients method: A radial basis functions method for solving partial differential equations on arbitrary surfaces. *Journal of Computational Physics*, 231(14):4662–4675, 2012.

John D Pryce. *Numerical Solution of Sturm-Liouville Problems*. Oxford University Press, 1993.

Steven J Ruuth and Barry Merriman. A simple embedding method for solving partial differential equations on surfaces. *Journal of Computational Physics*, 227(3):1943–1961, 2008.

Robert Schaback. Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics*, 3(3):251–264, 1995.

Varun Shankar, Grady B Wright, Robert M Kirby, and Aaron L Fogelson. A radial basis function (rbf)-finite difference (fd) method for diffusion and reaction–diffusion equations on surfaces. *Journal of scientific computing*, 63:745–768, 2015.

Bernard W Silverman. *Density estimation for statistics and data analysis*. Routledge, 2018.

Amit Singer and Hau-Tieng Wu. Spectral convergence of the connection Laplacian from random samples. *Information and Inference: A Journal of the IMA*, 6(1):58–123, 2017.

Bharath K Sriperumbudur, Kenji Fukumizu, and Gert RG Lanckriet. Universality, characteristic kernels and RKHS embedding of measures. *Journal of Machine Learning Research*, 12(7), 2011.

Ingo Steinwart. On the influence of the kernel on the consistency of support vector machines. *Journal of machine learning research*, 2(Nov):67–93, 2001.

Audry Ellen Tarwater. Parameter study of hardy's multiquadric method for scattered data interpolation. *Technical report UCRL-54670, Lawrence Livermore National Laboratory*, 1985.

Nicolás García Trillos, Moritz Gerlach, Matthias Hein, and Dejan Slepčev. Error estimates for spectral convergence of the graph Laplacian on random geometric graphs toward the Laplace–Beltrami operator. *Foundations of Computational Mathematics*, 20(4):827–887, 2020.

Hemant Tyagi, Elıf Vural, and Pascal Frossard. Tangent space estimation for smooth embeddings of Riemannian manifolds. *Information and Inference: A Journal of the IMA*, 2(1):69–114, 2013.

Ulrike Von Luxburg, Mikhail Belkin, and Olivier Bousquet. Consistency of spectral clustering. *The Annals of Statistics*, pages 555–586, 2008.

Max Wardetzky. *Discrete differential operators on polyhedral surfaces-convergence and approximation*. PhD thesis, 2007.

Holger Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4(1):389–396, 1995.

Holger Wendland. *Scattered Data Approximation*. Cambridge University Press, 2005.

Zong-min Wu and Robert Schaback. Local error estimates for radial basis function interpolation of scattered data. *IMA journal of Numerical Analysis*, 13(1):13–27, 1993.

Qile Yan, Shixiao W Jiang, and John Harlim. Spectral methods for solving elliptic PDEs on unknown manifolds. *Journal of Computational Physics*, 486:112132, 2023.

Zhenyue Zhang and Hongyuan Zha. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *SIAM journal on scientific computing*, 26(1):313–338, 2004.