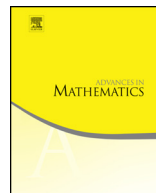




ELSEVIER

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

Embedding Riemannian manifolds by the heat kernel of the connection Laplacian



Hau-Tieng Wu

Department of Mathematics, University of Toronto, Toronto, Ontario, Canada

ARTICLE INFO

Article history:

Received 27 April 2013

Received in revised form 26 October 2015

Accepted 1 May 2016

Available online 22 September 2016

Communicated by Charles Fefferman

Keywords:

Vector diffusion map

Spectral geometry

Precompactness

Graph connection Laplacian

Spectral graph theory

ABSTRACT

Given a class of closed Riemannian manifolds with prescribed geometric conditions, we introduce an embedding of the manifolds into ℓ^2 based on the heat kernel of the Connection Laplacian associated with the Levi-Civita connection on the tangent bundle. As a result, we can construct a distance in this class which leads to a pre-compactness theorem on the class under consideration.

© 2016 Published by Elsevier Inc.

1. Introduction

In [6], the following class of closed Riemannian manifolds $\mathcal{M}_{d,k,D}$ with prescribed geometric constraints is considered:

$$\mathcal{M}_{d,k,D} = \{(M, g) \mid \dim(M) = d, \text{Ric}(g) \geq (d-1)kg, \text{diam}(M) \leq D\}, \quad (1)$$

where Ric is the Ricci curvature and diam is the diameter. The authors embed $M \in \mathcal{M}_{d,k,D}$ into the space ℓ^2 of real-valued, square integrable series by considering the heat

E-mail address: hauwu@math.toronto.edu.

kernel of the Laplace–Beltrami operator of M . A distance on $\mathcal{M}_{d,k,D}$, referred to as the *spectral distance* (SD), is then introduced based on the embedding so that the set of isometric classes in $\mathcal{M}_{d,k,D}$ is precompact. Over the past decades many works in the manifold learning field benefit from this embedding scheme, for example, the eigenmaps [3], diffusion maps (DM) [10] and the manifold parameterizations [17,18]. An important practical and theoretical question about how many eigenfunctions one needs to embed the manifold via DM has been answered in [2,21].

Recently, a new mathematical framework, referred to as *vector diffusion maps* (VDM), for organizing and analyzing massive high dimensional data sets, images and shapes was introduced in [23]. In brief, VDM is a mathematical generalization of DM, which numerically is based on a generalization of graph Laplacian, referred to as *graph connection Laplacian* (GCL). While DM are based on the heat kernel of the Laplace–Beltrami operator over the manifold, VDM is based on the heat kernel of the connection Laplacian associated with the Levi-Civita connection on the tangent bundle of the manifold. The introduction of VDM was motivated by the problem of finding an efficient way to organize complex data sets, embed them in a low dimensional space, and interpolate and regress vector fields over the data. In particular, it equips the data with a metric, which we refer to as the *vector diffusion distance* (VDD). VDM has been applied to several different problems, ranging from the cryo-electron microscopy problem [16,25,27] to graph realization problems [11,12] and modern light source imaging technique [1,19]. The application of VDM to the cryo-electron microscopy problem, which is aimed to reconstruct the three dimensional geometric structure of the macromolecule, provides a better organization of the given noisy projection images, and hence a better reconstruction result [16,25,27]. In addition, the GCL can be slightly modified to determine the orientability of a manifold and obtain its orientable double covering if the manifold is non-orientable [22].

In this paper, we consider the same class of closed Riemannian manifolds $\mathcal{M}_{d,k,D}$ and focus on the connection Laplacian associated with the Levi-Civita connection on the tangent bundle. We analyze how the VDM embeds the manifold $M \in \mathcal{M}_{d,k,D}$ into ℓ^2 based on the heat kernel of the connection Laplacian of the tangent bundle. Based on the VDM and VDD, we introduce a new distance on $\mathcal{M}_{d,k,D}$, referred to as *vector spectral distance* (VSD), which leads to the pre-compactness of the set $\mathcal{M}_{d,k,D}$.

1.1. Summary of the work and organization of the paper

The results of this paper are organized in the following way. In Section 2 the background material and notations are provided.

In Section 3, the VDM and VDD are defined and we discuss the embedding property of VDM in Theorem 3.1 and the local geodesic estimation property of VDD in Theorem 3.2. Indeed, VDM is a diffeomorphic embedding of a manifold $M \in \mathcal{M}_{d,k,D}$ into the Hilbert space ℓ^2 , and if $x, y \in M$ are close enough in the sense of geodesic distance, then the VDD of these two points are closely related to the geodesic distance.

In Section 4 we define VSD in the manifold set $\mathcal{M}_{d,k,D}$. The key ingredients in this section are the generalized Kato's type inequality comparing the trace of the heat kernel of the Laplace–Beltrami operator and the trace of the heat kernel of the connection Laplacian and a nice isoperimetric inequality for heat kernel comparisons. In brief, although the behavior of the heat kernel of the connection Laplacian might be intricate, we may control its trace simply by the trace of the heat kernel of the Laplace–Beltrami operator. With these key ingredients, in Theorem 4.6 we show that the newly defined metric, the VSD, is really a distance between isometry classes of Riemannian manifolds in $\mathcal{M}_{d,k,D}$. In Section 5, with the help of the VSD, the pre-compactness of the manifold set $\mathcal{M}_{d,k,D}$ is derived in Theorem 5.1 from the Rellich's Theorem. To show this Theorem, we also need the following Lemma:

Lemma 1.1. [6, Lemma 15] *Let (E, δ) be a metric space. Let $\mathcal{F}(E)$ denote the set of non-empty closed subsets of E , equipped with the Hausdorff distance h_δ associated with δ . If the metric space (E, δ) is precompact, so is the metric space $(\mathcal{F}(E), h_\delta)$.*

In fact, we view the VDM of a given manifold in $\mathcal{M}_{d,k,D}$ as a point in the set consisting of embeddings of all manifolds in $\mathcal{M}_{d,k,D}$, and then apply Lemma 1.1 to show the pre-compactness of $\mathcal{M}_{d,k,D}$ with related to the VSD. The main theorems are summarized here for the reader's convenience:

- Theorem 4.6: For any fixed $t > 0$, d_t , the VSD defined in (43), is a distance between isometry classes of Riemannian manifolds in $\mathcal{M}_{d,k,D}$. In particular, two Riemannian manifolds $(M, g), (M', g') \in \mathcal{M}_{d,k,D}$ satisfy $d_t(M, M') = 0$ if and only if M and M' are isometric.
- Theorem 5.1: For any $t > 0$, the space of the isometry classes in $\mathcal{M}_{d,k,D}$ is d_t -precompact.

In Section 6, we discuss the application of VDM and this new pre-compactness result from the viewpoint of manifold learning.

2. Background material

Let (M, g) be a closed Riemannian manifold and TM the tangent bundle. Denote $C^\infty(TM)$ the smooth vector fields and $L^2(TM)$ the vector fields satisfying

$$\int_M \langle X, X \rangle(x) dV(x) \leq \infty, \quad (2)$$

where dV is the volume form associated with g and $\langle X, X \rangle(x) := g(X(x), X(x))$. Denote ∇ the Levi-Civita connection of \mathcal{M} and $P_{x,y}$ the parallel transport from y to x via the geodesic linking them. Denote Δ^{TM} the connection Laplacian associated with ∇ on

the tangent bundle TM [15]. The connection Laplacian Δ^{TM} is a self-adjoint, second order elliptic operator [15]. From the classical elliptic theory [15] we know that the heat semigroup, $e^{t\Delta^{\text{TM}}}$, $t > 0$, with the infinitesimal generator Δ^{TM} is a family of self-adjoint operators with the heat kernel $k_{TM}(t, x, y)$ so that

$$e^{t\Delta^{\text{TM}}}X(x) = \int_M k_{TM}(t, x, y)X(y)dV(y). \quad (3)$$

The heat kernel $k_{TM}(t, x, y)$ is smooth in x and y and analytic in t [15].

It is well known [15] that the spectrum of Δ^{TM} is discrete inside \mathbb{R}^- , the non-positive real numbers, and the only possible accumulation point is $-\infty$. We will denote the spectrum of Δ^{TM} as $\{-\lambda_k\}_{k=1}^\infty$, where $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \dots$, and its eigen-vector fields as $\{X_k\}_{k=1}^\infty$. Notice that λ_0 may not exist due to the topological obstruction. For example, we can not find a nowhere non-vanishing vector field on S^2 . In other words, $\Delta^{\text{TM}}X_k = -\lambda_k X_k$ for all $k = 1, 2, \dots$. It is also well known [15] that $\{X_k\}_{k=1}^\infty$ form an orthonormal basis for $L^2(TM)$. Denote the heat kernel of Δ^{TM} by $k_{TM}(t, x, y)$, which can be expressed as [15]:

$$k_{TM}(t, x, y) = \sum_{n=1}^\infty e^{-\lambda_n t} X_n(x) \otimes \overline{X_n(y)}. \quad (4)$$

A calculation of the Hilbert–Schmidt norm of the heat kernel at (t, x, y) gives

$$\begin{aligned} \|k_{TM}(t, x, y)\|_{HS}^2 &= \text{Tr} \left[k_{TM}(t, x, y) \overline{k_{TM}(t, x, y)} \right] \\ &= \sum_{n,m=1}^\infty e^{-(\lambda_n + \lambda_m)t} \langle X_n(x), X_m(x) \rangle \overline{\langle X_n(y), X_m(y) \rangle}. \end{aligned} \quad (5)$$

On the other hand, the classical elliptic theory [15] allows us to decompose $L^2(TM)$ as $L^2(TM) = \oplus_{k=1}^\infty E_k$, where E_k is the eigenspace of Δ^{TM} corresponding to the eigenvalue $-\lambda_k$. Denote by $m(-\lambda_k)$ the multiplicity of $-\lambda_k$. It is also well known that $m(-\lambda_k)$ is finite. Denote $\mathcal{B}(E_k)$ the set of bases of E_k , which is identical to the orthogonal group $O(m(-\lambda_k))$. Denote the set of the corresponding orthonormal bases of $L^2(TM)$ by

$$\mathcal{B}(M, g) = \Pi_{k=1}^\infty \mathcal{B}(E_k). \quad (6)$$

By Tychonoff's theorem, we know $\mathcal{B}(M, g)$ is compact since $O(m(-\lambda_k))$ is compact for all $k \in \mathbb{N}$. Also note that the dot products $\langle X_n(x), X_m(x) \rangle$, where $n, m \in \mathbb{N}$, are invariant to the choice of basis for $T_x M$.

3. Vector diffusion maps

Based on these observations, given $a \in \mathcal{B}(M, g)$ and $t > 0$, the authors in [23] define the VDM V_t^a which maps $x \in M$ to the Hilbert space ℓ^2 by¹:

$$V_t^a : x \mapsto \text{Vol}(M) \left(e^{-(\lambda_n + \lambda_m)t/2} \langle X_n(x), X_m(x) \rangle \right)_{n,m=1}^{\infty}, \quad (7)$$

where $a = \{X_n\}_{n=1}^{\infty}$. A direct calculation shows that

$$\|k_{TM}(t, x, y)\|_{HS}^2 = \frac{1}{\text{Vol}(M)^2} \langle V_t^a(x), V_t^a(y) \rangle_{\ell^2}. \quad (8)$$

Fix $a \in \mathcal{B}(M, g)$. For all $t > 0$, the following Theorem states that the VDM V_t^a is an embedding of the compact Riemannian manifold M into ℓ^2 . The proof of the theorem is given in [23, Theorem 8.1].

Theorem 3.1. *Given a d -dim closed Riemannian manifold (M, g) and an orthonormal basis $a = \{X_k\}_{k=1}^{\infty}$ of $L^2(TM)$ composed of the eigenvector-fields of the connection Laplace Δ^{TM} , then for any $t > 0$, the VDM V_t^a is a diffeomorphic embedding of M into ℓ^2 .*

It is Theorem 3.1 that allows the authors to define the VDD between $x, y \in M$, denoted as $d_{\text{VDM},t}(x, y)$, in [23]:

$$d_{\text{VDM},t}(x, y) := \|V_t^a(x) - V_t^a(y)\|_{\ell^2}, \quad (9)$$

which is clearly a distance function over M . We mention that by the following expansion

$$\begin{aligned} d_{\text{VDM},t}^2(x, y) &= \|V_t^a(x) - V_t^a(y)\|_{\ell^2}^2 \\ &= \text{Vol}(M)^2 \sum_{n,m=1}^{\infty} e^{-(\lambda_n + \lambda_m)t} (\langle X_n(x), X_m(x) \rangle - \langle X_n(y), X_m(y) \rangle)^2 \\ &= \text{Vol}(M)^2 [\text{Tr}(k_{TM}(t, x, x)k_{TM}(t, x, x)^*) + \\ &\quad \text{Tr}(k_{TM}(t, y, y)k_{TM}(t, y, y)^*) - 2\text{Tr}(k_{TM}(t, x, y)k_{TM}(t, x, y)^*)], \end{aligned} \quad (10)$$

the defined VDD $d_{\text{VDM},t}$ does not depend on the choice of the basis a . The following theorem shows that in this asymptotic limit the vector diffusion distance behaves like the geodesic distance. The proof of the theorem is given in [23, Theorem 8.2].

Theorem 3.2. *Let (M, g) be a smooth d -dim closed Riemannian manifold. Suppose $x, y \in M$ so that $x = \exp_y v$, where $v \in T_y M$. For any $t > 0$, when $\|v\|^2 \ll t \ll 1$ we have the following asymptotic expansion of the VDD:*

¹ Note that the basis a and the volume of M are not taken into consideration in the definition of the VDM in [23]. To prove the precompactness theorem, we need to take them into consideration.

$$d_{\text{VDM},t}^2(x, y) = d\text{Vol}(M)^2(4\pi)^{-d} \frac{\|v\|^2}{t^{d+1}} + O(t^{-d}\|v\|^2). \quad (11)$$

4. Vector spectral distances

In this section and the next, we show that based on the VDM V_t^a , we can define a family of VSD d_t , $t > 0$, on the space of the isometry classes in $\mathcal{M}_{d,k,D}$ so that for any $t > 0$ the space of the isometry classes in $\mathcal{M}_{d,k,D}$ is d_t -precompact.

Denote the Laplace–Beltrami operator over (M, g) by Δ_M and its eigenvalues and eigenfunctions by $-\mu_k$ and ϕ_k , where $k \in \{0\} \cup \mathbb{N}$, that is, $\Delta_M \phi_k = -\mu_k \phi_k$, so that $\mu_0 = 0 < \mu_1 \leq \mu_2 \dots$. Define the following partition functions

$$Z_{TM}(t) := \sum_{j=1}^{\infty} e^{-\lambda_j t} \quad (12)$$

and

$$Z_M(t) := \sum_{j=0}^{\infty} e^{-\mu_j t}, \quad (13)$$

which are related by the following generalized Kato's type inequality [5, p. 135]:

$$Z_{TM}(t) \leq dZ_M(t) \text{ for all } t > 0. \quad (14)$$

Then recall the following result:

Theorem 4.1. [5, p. 108 C. 26] *Let (M, g) be an d -dimensional closed Riemannian manifold. Define*

$$r_{\min}(M) = \inf\{\text{Ric}(v, v) : \|v\| = 1\} \quad (15)$$

and the diameter of M by $D(M)$. If (M, g) satisfies $r_{\min}(M)D(M)^2 \geq (d-1)\epsilon\alpha^2$ for $\epsilon \in \{-1, 0, 1\}$ and $\alpha > 0$, then

$$\text{Vol}(M)k_M(t, x, x) \leq \text{Vol}(S^d(R))k_{S^d}(t, y, y) = Z_{S^d(R)}(t) = Z_{S^d(1)}(t/R^2), \quad (16)$$

where $y \in S^d(R)$, $S^d(R)$ is the standard d -dim sphere in \mathbb{R}^{d+1} with center 0 and radius R , $R = D(M)/a(d, \epsilon, \alpha)$ and

$$a(d, \epsilon, \alpha) = \begin{cases} \alpha \omega_d^{1/d} \left(2 \int_0^{\alpha/2} \cos^{d-1}(t) dt \right)^{-1/d} & \text{if } \epsilon = 1 \\ (1 + d\omega_d)^{1/d} - 1 & \text{if } \epsilon = 0 \\ \alpha c(\alpha) & \text{if } \epsilon = -1, \end{cases} \quad (17)$$

where $\omega_d = \text{Vol}(S^d)/\text{Vol}(S^{d-1})$ and $c(\alpha)$ is the unique positive root $z > 0$ of the equation

$$z \int_0^\alpha (\cosh(t) + z \sinh(t))^{d-1} dt = \omega_d. \quad (18)$$

With inequalities (14) and (16), we prove the following lemmas, which is essential in showing the pre-compactness result. We omit the dependence on M to simplify the statement of the lemmas and the proof.

Lemma 4.2. *With the above notations, there exist positive constants $A(d, k, D)$, $B(d, k, D)$ and $E(d, k, D)$ that depend only on d, k and D such that for any $(M, g) \in \mathcal{M}_{d,k,D}$:*

- (a) $\lambda_j \geq A(d, k, D)j^{2/d}$;
- (b) $N(\lambda) := \#\{j \mid j \geq 0, \lambda_j \leq \lambda\} \leq ed + B(d, k, D)\lambda^{d/2}$;
- (c) *for all $x \in M$ and $\alpha \geq 0$, we have*

$$\sum_{n,m \geq 1} (\lambda_n + \lambda_m)^\alpha e^{-t(\lambda_n + \lambda_m)} \langle X_n(x), X_m(x) \rangle^2 \leq \frac{E(d, k, D)}{\text{Vol}(M)^2} F(\alpha, d) t^{-\alpha-d}, \quad (19)$$

where

$$F(\alpha, d) := \int_0^\infty \int_0^\infty [(x+y)^2 - 2\alpha(x+y) + \alpha(\alpha-1)](x+y)^{\alpha+d-2} e^{-(x+y)} dx dy. \quad (20)$$

Proof. The proofs for (a) and (b) are almost the same as the proofs of Theorem 3 in [6] except that we apply (14). We provide the proofs here for completion. If $k \geq 0$, $r_{\min} D^2 \geq 0$ and if $k < 0$, $r_{\min} D^2 \geq (d-1)kD$. Thus we can apply Theorem 4.1 with ϵ and α depending only on k and D . Thus,

$$\begin{aligned} Z_{TM}(t) &\leq dZ_M(t) = d \int_M k_M(t, x, x) dx \leq d \text{Vol}(M) \sup_{x \in M} k_M(t, x, x) \\ &\leq d \text{Vol}(S^d(R)) k_{S^d(R)}(t, y, y) = dZ_{S^d(R)}(t) = dZ_{S^d(1)}(t/R^2), \end{aligned} \quad (21)$$

where $y \in S^d(R)$. The trace of the heat kernel Z_{TM} is thus uniformly bounded on the set $\mathcal{M}_{d,k,D}$. Note that for $S^d(1)$, the j -th eigenvalue is $-j(j+d-1)$ with multiplicity $\binom{d+j}{d} - \binom{d+j-2}{d}$, where $\binom{m}{n}$ is the number of ways of choosing n from m , so by a direct calculation, there exists a constant $b(d)$, depending on d only, such that for any $t > 0$,

$$Z_{S^d(1)}(t) - 1 \leq b(d)t^{-d/2}. \quad (22)$$

Also note that $j \leq N(\lambda_j)$, $j \in \mathbb{N} \cup \{0\}$, in general, and $j = N(\lambda_j)$ when all eigenvalues are simple. As a consequence, we have

$$\begin{aligned}
 j \leq N(\lambda_j) &\leq e \sum_{0 \leq \lambda_i \leq \lambda_j} e^{-\lambda_i/\lambda_j} \leq e Z_{TM}(1/\lambda_j) \\
 &\leq e d Z_{S^d(1)} \left(\frac{1}{\lambda_j R^2} \right) \leq e d + e d b(d) R^d \lambda_j^{d/2}
 \end{aligned} \tag{23}$$

and hence

$$\lambda_j \geq \left(\frac{j - ed}{edb(d)R^d} \right)^{2/d} \geq (edb(d)R^d)^{-2/d} j^{2/d}. \tag{24}$$

Since $R = a(d, \epsilon, \alpha)D(M) \leq a(d, \epsilon, \alpha)D$ and $a(d, \epsilon, \alpha)$ only depends on d, k and D , this proves (a) and (b).

Next we prove (c). Define the positive measure μ_x , where $x \in M$, on $\mathbb{R} \times \mathbb{R}$ by

$$d\mu_x = \sum_{n, m \geq 1} \langle X_n(x), X_m(x) \rangle^2 \delta_{\lambda_n} \times \delta_{\lambda_m} \tag{25}$$

where δ_{λ_n} is the Dirac measure at $\lambda_n \in \mathbb{R}$. The left hand side of (19) becomes:

$$\begin{aligned}
 &\sum_{n, m \geq 1} (\lambda_n + \lambda_m)^\alpha e^{-t(\lambda_n + \lambda_m)} \langle X_n(x), X_m(x) \rangle^2 \\
 &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\lambda + \nu)^\alpha e^{-t(\lambda + \nu)} d\mu_x(\lambda, \nu),
 \end{aligned} \tag{26}$$

where the equality holds since when $t > 0$, $(\lambda + \nu)e^{-t(\lambda + \nu)}$ is smooth and decays fast enough; that is, $(\lambda + \nu)e^{-t(\lambda + \nu)} \chi_{[0, \infty) \times [0, \infty)}$ is a Schwartz function defined on \mathbb{R}^2 . To study (26), consider the following $L^1_{loc}(\mathbb{R}^2)$ function:

$$\psi(\lambda, \nu) := \sum_{n: 0 \leq \lambda_n \leq \lambda, m: 0 \leq \lambda_m \leq \nu} \langle X_n(x), X_m(x) \rangle^2, \tag{27}$$

which is supported on $[0, \infty) \times [0, \infty)$. Note that by the definition of the derivative in the sense of distribution, for any Schwartz function ϕ we have

$$\int \phi(\lambda, \nu) \sum_{n \geq 1, m \geq 1} \langle X_n(x), X_m(x) \rangle^2 \delta_{\lambda_n} \times \delta_{\lambda_m} = \int \partial_\lambda \partial_\nu \phi(\lambda, \nu) \psi(\lambda, \nu) d\lambda d\nu. \tag{28}$$

Thus, (26) becomes

$$\begin{aligned}
 &\int_{\mathbb{R}} \int_{\mathbb{R}} (\lambda + \nu)^\alpha e^{-t(\lambda + \nu)} d\mu_x(\lambda, \nu) \\
 &= \int_0^\infty \int_0^\infty \partial_\lambda \partial_\nu \left[(\lambda + \nu)^\alpha e^{-t(\lambda + \nu)} \right] \psi(\lambda, \nu) d\lambda d\nu
 \end{aligned}$$

$$\begin{aligned}
&= \int_0^\infty \int_0^\infty [t^2(\lambda + \nu)^2 - 2\alpha t(\lambda + \nu) + \alpha(\alpha - 1)] \times \\
&\quad (\lambda + \nu)^{\alpha-2} e^{-t(\lambda+\nu)} \psi(\lambda, \nu) d\lambda d\nu.
\end{aligned} \tag{29}$$

To finish the proof, we bound $\psi(\lambda, \nu)$ when $\lambda, \nu > 0$. By (16) and the fact that $\|k_{TM}(t, x, x)\|_{HS} \leq k_M(t, x, x)$ [6, p. 137] for all $t > 0$ and $x \in M$, we have

$$\begin{aligned}
\psi(\lambda, \nu) &= \sum_{n: 0 \leq \lambda_n \leq \lambda, m: 0 \leq \lambda_m \leq \nu} \langle X_n(x), X_m(x) \rangle^2 \\
&\leq e \sum_{n: 0 \leq \lambda_n \leq \lambda, m: 0 \leq \lambda_m \leq \nu} e^{-(\lambda_n + \lambda_m)/(\lambda + \nu)} \langle X_n(x), X_m(x) \rangle^2 \\
&\leq e \sum_{n=1, \dots, \infty, m=1, \dots, \infty} e^{-(\lambda_n + \lambda_m)/(\lambda + \nu)} \langle X_n(x), X_m(x) \rangle^2 \\
&= e \left\| k_{TM} \left(\frac{1}{\lambda + \nu}, x, x \right) \right\|_{HS}^2 \leq e k_M \left(\frac{1}{\lambda + \nu}, x, x \right)^2 \\
&\leq \frac{e}{\text{Vol}(M)^2} Z_{S^n} \left(\frac{1}{(\lambda + \nu)R^2} \right) \leq \frac{C(d, k, D)}{\text{Vol}(M)^2} (\lambda + \nu)^d,
\end{aligned} \tag{30}$$

where R is defined in Theorem 4.1 and $C(d, k, D)$ is a universal constant depending on d, k and D . Thus we conclude that the left hand side of (19) is bounded by

$$\begin{aligned}
&\sum_{n, m \geq 1} (\lambda_n + \lambda_m)^\alpha e^{-t(\lambda_n + \lambda_m)} \langle X_n(x), X_m(x) \rangle^2 \\
&\leq \frac{C(d, k, D)^2}{\text{Vol}(M)^2} \int_0^\infty \int_0^\infty [t^2(\lambda + \nu)^2 - 2\alpha t(\lambda + \nu) + \alpha(\alpha - 1)] \times \\
&\quad (\lambda + \nu)^{\alpha-2} e^{-t(\lambda+\nu)} (\lambda + \nu)^d d\lambda d\nu \\
&\leq \frac{E(d, k, D)}{\text{Vol}(M)^2} F(\alpha, d) t^{-\alpha-d},
\end{aligned} \tag{31}$$

where $E(d, k, D)$ is a universal positive constant that depends only on d, k and D , and F is defined in (20). \square

Recall that the VDM V_t^a depends on the choice of an orthonormal basis a of eigenvector fields. Given a finite dimensional Euclidean space E , we can define the distance between $R_1, R_2 \in O(\dim E)$ by

$$d_E(R_1, R_2) = \|R_1^{-1}R_2 - I\|_{HS}. \tag{32}$$

It is clear that $d_E(R_1, R_2) \leq 2\sqrt{\dim E}$. As we discussed above, $\mathcal{B}(M, g)$ is a compact set with respect to the product topology. This topology can be described by the distance $d_{\mathcal{B}(M, g)}$ between $a, b \in \mathcal{B}(M, g)$ defined as

$$d_{\mathcal{B}(M,g)}(a,b)^2 := \sum_{i=1}^{\infty} \lambda_i^{-N} d_{E_i}(a|_{E_i}, b|_{E_i})^2, \quad (33)$$

where $a|_{E_i} \in \mathcal{B}(E_i)$ means a basis of the eigenspace associated with the i -th eigenvalue and $d_{E_i}(a|_{E_i}, b|_{E_i})$ is defined by (32). The series on the right hand side of (33) converges when $N > d/2$ due to Lemma 4.2(a). The following Lemma is a generalization of Theorem 8 in [6] to the tangent bundle we have interest.

Lemma 4.3. *Let (M, g) be a smooth d -dim closed Riemannian manifold. The map $V : \mathbb{R}_+ \times \mathcal{B}(M, g) \times M \rightarrow \ell^2$ defined by $V(t, a, x) := V_t^a(x)$ is continuous and satisfies:*

$$\begin{aligned} \|V_t^a(x) - V_s^b(y)\|_{\ell^2}^2 &\leq \text{Vol}(M)^2 \left\{ \|k_{TM}(t, x, x)\|_{HS} + \|k_{TM}(s, y, y)\|_{HS} \right. \\ &\quad \left. - 2 \left\| k_{TM} \left(\frac{t+s}{2}, x, y \right) \right\|_{HS} + 2d_{\mathcal{B}(M,g)}(a, b) k_{TM}^{(N)}(t, x, x)^{1/2} k_{TM}^{(N)}(s, y, y)^{1/2} \right\} \end{aligned} \quad (34)$$

where $t > 0$, $a, b \in \mathcal{B}(M, g)$, $N > d/2$, $x, y \in M$, and

$$k_{TM}^{(N)}(t, x, x) = \sum_{n,m=1}^{\infty} (\lambda_n^{N/2} + \lambda_m^{N/2}) e^{-t(\lambda_n + \lambda_m)} \langle X_n^a(x), X_m^a(x) \rangle^2. \quad (35)$$

Proof. We denote the basis $a \in \mathcal{B}(M, g)$ by $\{X_n^a\}_{n=1}^{\infty}$. First we have

$$\begin{aligned} &\|V_t^a(x) - V_s^b(y)\|_{\ell^2}^2 \\ &= \text{Vol}(M)^2 \sum_{n,m=1}^{\infty} \left(e^{-t(\lambda_n + \lambda_m)/2} \langle X_n^a(x), X_m^a(x) \rangle - e^{-s(\lambda_n + \lambda_m)/2} \langle X_n^b(y), X_m^b(y) \rangle \right)^2 \\ &= \text{Vol}(M)^2 \left\{ \|k_{TM}(t, x, x)\|_{HS}^2 + \|k_{TM}(s, y, y)\|_{HS}^2 \right. \\ &\quad \left. - 2 \sum_{n,m=1}^{\infty} e^{-(t+s)(\lambda_n + \lambda_m)/2} \langle X_n^a(x), X_m^a(x) \rangle \langle X_n^b(y), X_m^b(y) \rangle \right\} \\ &= \text{Vol}(M)^2 \left\{ \|k_{TM}(t, x, x)\|_{HS}^2 + \|k_{TM}(s, y, y)\|_{HS}^2 - \left\| k_{TM} \left(\frac{t+s}{2}, x, y \right) \right\|_{HS}^2 \right. \\ &\quad \left. - 2 \sum_{n,m=1}^{\infty} e^{-(t+s)(\lambda_n + \lambda_m)/2} \langle X_n^a(x), X_m^a(x) \rangle \left(\langle X_n^b(y), X_m^b(y) \rangle - \langle X_n^a(y), X_m^a(y) \rangle \right) \right\}, \end{aligned}$$

where we denote the last summation on the right hand side as A . Denote the eigen-vector fields inside the eigenspace E_n by $X_{n(j)}^b(y)$, where $j = 1, \dots, m(\nu_n)$, when the basis of $L^2(TM)$ is chosen to be $b \in \mathcal{B}(M, g)$. By definition, we have the following relationship:

$$X_{n(j)}^b(y) = \sum_{k=1}^{m(\nu_n)} \alpha_{j,k}(b, a) X_{n(k)}^a(y) \quad (36)$$

where the matrix $[\alpha_{j,k}(b, a)]_{k,j=1}^{m(\nu_j)} \in O(m(\nu_n))$. Here the orthogonal matrix $[\alpha_{j,k}(b, a)]_{k,j=1}^{m(\nu_j)}$ is the relationship between two bases, $a|_{E_j}$ and $b|_{E_j}$, associated with the eigenspace of the j -th eigenvalue. Thus we can rewrite A as

$$\begin{aligned} A &= \sum_{n,m=1}^{\infty} e^{-(t+s)(\nu_n+\nu_m)/2} \sum_{k=1}^{m(\nu_n)} \sum_{l=1}^{m(\nu_m)} \langle X_{n(k)}^a(x), X_{m(l)}^a(x) \rangle \times \\ &\quad \left(\langle X_{n(k)}^a(y), X_{m(l)}^a(y) \rangle - \sum_{i=1}^{m(\nu_n)} \sum_{j=1}^{m(\nu_m)} \alpha_{k,i}(b, a) \alpha_{l,j}(b, a) \langle X_{n(i)}^a(y), X_{m(j)}^a(y) \rangle \right) \\ &= \sum_{n,m=1}^{\infty} e^{-(t+s)(\nu_n+\nu_m)/2} \sum_{i,k=1}^{m(\nu_n)} \sum_{j,l=1}^{m(\nu_m)} \langle X_{n(k)}^a(x), X_{m(l)}^a(x) \rangle \times \\ &\quad \langle X_{n(i)}^a(y), X_{m(j)}^a(y) \rangle \left(\delta_{k,i} \delta_{l,j} - \alpha_{k,i}(b, a) \alpha_{l,j}(b, a) \right), \end{aligned} \quad (37)$$

where $[\delta_{k,l}]_{k,l=1}^{m(\nu_j)}$ is an $m(\nu_j) \times m(\nu_j)$ identity matrix. Note that

$$(\delta_{k,i} \delta_{l,j} - \alpha_{k,i}(b, a) \alpha_{l,j}(b, a)) = \delta_{k,i} (\delta_{l,j} - \alpha_{l,j}(b, a)) + \alpha_{l,j}(b, a) (\delta_{k,i} - \alpha_{k,i}(b, a)) \quad (38)$$

which is bounded by $d_{E_n}(a|_{E_n}, b|_{E_n}) + d_{E_m}(a|_{E_m}, b|_{E_m})$. Hence, by the Cauchy–Schwartz inequality, A is bounded by

$$\begin{aligned} |A| &\leq \sum_{n,m=1}^{\infty} e^{-(t+s)(\nu_n+\nu_m)/2} \left(\sum_{k=1}^{m(\nu_n)} \sum_{l=1}^{m(\nu_m)} \langle X_{n(k)}^a(x), X_{m(l)}^a(x) \rangle^2 \right)^{1/2} \times \\ &\quad \left(\sum_{i=1}^{m(\nu_n)} \sum_{j=1}^{m(\nu_m)} \langle X_{n(i)}^a(y), X_{m(j)}^a(y) \rangle^2 \right)^{1/2} \left(d_{E_n}(a|_{E_n}, b|_{E_n}) + d_{E_m}(a|_{E_m}, b|_{E_m}) \right) \\ &\leq 2d_{\mathcal{B}(M,g)}(a, b) \sum_{n,m=1}^{\infty} (\nu_n^{N/2} + \nu_m^{N/2}) e^{-(t+s)(\nu_n+\nu_m)/2} \times \\ &\quad \left(\sum_{k=1}^{m(\nu_n)} \sum_{l=1}^{m(\nu_m)} \langle X_{n(k)}^a(x), X_{m(l)}^a(x) \rangle^2 \right)^{1/2} \left(\sum_{i=1}^{m(\nu_n)} \sum_{j=1}^{m(\nu_m)} \langle X_{n(i)}^a(y), X_{m(j)}^a(y) \rangle^2 \right)^{1/2} \\ &\leq 2d_{\mathcal{B}(M,g)}(a, b) \left(\sum_{n,m=1}^{\infty} (\lambda_n^{N/2} + \lambda_m^{N/2}) e^{-t(\lambda_n+\lambda_m)} \langle X_n^a(x), X_m^a(x) \rangle^2 \right)^{1/2} \times \\ &\quad \left(\sum_{n,m=1}^{\infty} (\lambda_n^{N/2} + \lambda_m^{N/2}) e^{-s(\lambda_n+\lambda_m)} \langle X_n^a(y), X_m^a(y) \rangle^2 \right)^{1/2} \\ &= 2d_{\mathcal{B}(M,g)}(a, b) k_{TM}^{(N)}(t, x, x)^{1/2} k_{TM}^{(N)}(s, y, y)^{1/2}, \end{aligned}$$

where $k_{TM}^{(N)}(t, x, x)$ is defined in (35). Due to Lemma 4.2(c), $k_{TM}^{(N)}(t, x, x)$ is bounded, and thus the proof is finished. \square

With the above preparation, we are ready to introduce the VSD. Recall the following definitions. Suppose (X, δ) is a metric space, where δ is the metric, and $A, B \subset X$. The distance between A and B is defined as:

$$h(A, B) := \inf\{\delta(a, b) : a \in A, b \in B\}. \quad (39)$$

Denote the generalized ball of radius $\epsilon > 0$ around A :

$$\mathcal{N}(A, \epsilon) := \{x \in X : h(x, A) < \epsilon\}. \quad (40)$$

Given two subsets $A, B \subset X$, we can define the Hausdorff distance, denoted as HD, associated with δ by

$$\text{HD}(A, B) = \inf\{\epsilon : A \subset \mathcal{N}(B, \epsilon), B \subset \mathcal{N}(A, \epsilon)\}, \quad (41)$$

or equivalently

$$\text{HD}(A, B) = \max\left\{\sup_{x \in A} \inf_{y \in B} \delta(x, y), \sup_{y \in B} \inf_{x \in A} \delta(x, y)\right\}. \quad (42)$$

In the following, we focus on the metric space $(\ell^2, \|\cdot\|_{\ell^2})$. Let HD denote the Hausdorff distance between compact subsets of ℓ^2 associated with $\|\cdot\|_{\ell^2}$. Given two Riemannian manifolds M and M' and $t > 0$, we define a family of functions $d_t : \mathcal{M}_{d,k,D} \times \mathcal{M}_{d,k,D} \mapsto \mathbb{R}$ by

$$d_t(M, M') := \max\left\{\sup_{a \in \mathcal{B}(M,g)} \inf_{a' \in \mathcal{B}(M',g')} \text{HD}(V_t^a(M), V_t^{a'}(M')), \sup_{a' \in \mathcal{B}(M',g')} \inf_{a \in \mathcal{B}(M,g)} \text{HD}(V_t^{a'}(M'), V_t^a(M))\right\}. \quad (43)$$

We have to justify that d_t , $t > 0$, is a distance function defined on the set consisting of the isometry classes in $\mathcal{M}_{d,k,D}$, and then call it VSD. The following lemmas are needed for the justification.

Lemma 4.4. *Let (M, g) be a smooth d -dim closed Riemannian manifold and $\{X_n\}_{n \in \mathbb{N}}$ be an orthonormal basis of $L^2(TM)$ constituted of the eigen-vector fields of the connection Laplacian. Then*

$$\text{span}\{\langle X_n, X_m \rangle : n, m \in \mathbb{N}, n \neq m\} = L^2(M) \setminus \{\mathbb{R}\} \quad (44)$$

that is, all L^2 functions over M without the constant functions.

Proof. Fix $f \in L^2(M)$. If $\int_M \langle X_n, X_m \rangle(x) f(x) dx = \int_M \langle f X_n, X_m \rangle(x) dx = 0$ for all $n \neq m$, we show that f is constant. Since $\{X_n\}_{n \in \mathbb{N}}$ is an orthonormal basis of $L^2(TM)$, we can rewrite $f X_n = \sum_{k \in \mathbb{N}} \alpha_{n,k} X_k$, and hence we have

$$0 = \int_M \langle f X_n, X_m \rangle dx = \sum_{k \in \mathbb{N}} \alpha_{n,k} \int_M \langle X_k, X_m \rangle dx = \alpha_{n,m} \quad (45)$$

when $m \neq n$. As a result, we know $f X_n = \alpha_{n,n} X_n$, which implies that f is constant. \square

Lemma 4.5. *Let (M, g) be a smooth d -dim closed Riemannian manifold and $\{X_n\}_{n \in \mathbb{N}}$ is an orthonormal basis of $L^2(TM)$ constituted of the eigen-vector fields of the connection Laplacian. For any $x_0 \in M$, there exist d pairs of $\{(n_i, m_i)\}_{i=1}^d$, where $n_i, m_i \in \mathbb{N}$, so that the gradient vectors $\nabla \langle X_{n_i}, X_{m_i} \rangle(x_0)$ span $T_{x_0}M$.*

Proof. If not, there is a vector $v \in T_{x_0}M$ perpendicular to $\text{span}\{\nabla \langle X_n, X_m \rangle(x_0)\}_{n,m \in \mathbb{N}}$. It is well known that any function $u \in C^\infty(M)$ can be expanded by the eigenfunctions of the Laplace Beltrami operator, $\{\phi_i\}_{i=0}^\infty$, that is, $u = \sum_{j=0}^{K-1} \phi_j + \sum_{i=K}^\infty u_i \phi_i$, where K is the number of connected components, $\Delta_M \phi_j = 0$ for all $j = 0, 1, \dots, K-1$, and $u_i = \int_M u \phi_i dx$ for all $i = K, K+1, \dots$. It follows that $\nabla u(x_0) = \sum_{i=K}^\infty u_i \nabla \phi_i(x_0)$. Since any vector $v \in T_{x_0}M$ can be written as $\nabla u(x_0)$ for some smooth function u , we know

$$v = \sum_{i=K}^\infty v_i \nabla \phi_i(x_0) \quad (46)$$

for some constants v_i . On the other hand, by Lemma 4.4, ϕ_i , $i = K, K+1, \dots$, can be expanded by $\{\langle X_n, X_m \rangle\}_{n,m \in \mathbb{N}, n \neq m}$. Thus, there exist constants $w_{i,n,m}$ so that

$$\begin{aligned} v &= \sum_{i=K}^\infty v_i \sum_{n,m=1}^\infty w_{i,n,m} \nabla \langle X_n, X_m \rangle(x_0) \\ &= \sum_{n,m=1}^\infty \left(\sum_{i=K}^\infty v_i w_{i,n,m} \right) \nabla \langle X_n, X_m \rangle(x_0), \end{aligned} \quad (47)$$

which is absurd, and we finish the proof. \square

Theorem 4.6. *For any fixed $t > 0$, d_t is a distance between isometry classes of Riemannian manifolds in $\mathcal{M}_{d,k,D}$. In particular, two Riemannian manifolds $(M, g), (M', g') \in \mathcal{M}_{d,k,D}$ satisfy $d_t(M, M') = 0$ if and only if (M, g) and (M', g') are isometric.*

Proof. By definition, it is clear that $d_t(M, M') \geq 0$, $d_t(M, M') = d_t(M', M)$ and the triangular inequality holds. To finish the proof that d_t is a distance we need to show that $d_t(M, M') = 0$ if and only if M is isometric to M' . If M and M' are isometric, then it

is trivial to see that $d_t(M, M') = 0$. Now we consider the opposite direction. Fix $t > 0$. If $d_t(M, M') = 0$, we claim that M is isometric to M' . By the definition of d_t , we know

$$\inf_{a \in \mathcal{B}(M, g)} \text{HD}(V_t^a(M), V_t^{a'}(M')) = 0 \quad (48)$$

for a given $a' \in \mathcal{B}(M', g')$. Thus there exists a sequence $a_n \in \mathcal{B}(M, g)$, $n = 1, 2, \dots$, so that

$$\lim_{n \rightarrow \infty} \text{HD}(V_t^{a_n}(M), V_t^{a'}(M')) = 0. \quad (49)$$

By the compactness of $\mathcal{B}(M, g)$, a subsequence $\{a_{n_j}\}_{j=1}^\infty$ converges to $a_0 \in \mathcal{B}(M, g)$, that is,

$$\lim_{j \rightarrow \infty} d_{\mathcal{B}(M, g)}(a_{n_j}, a_0) = 0, \quad (50)$$

and it follows that $\text{HD}(V_t^{a_0}(M), V_t^{a'}(M')) = 0$.

Let $a_0 = \{X_n\}_{n \in \mathbb{N}}$ and $a' = \{X'_n\}_{n \in \mathbb{N}}$. From the definition of Hausdorff distance and the compactness of $V_t^{a_0}(M)$ and $V_t^{a'}(M')$, we have

$$\begin{aligned} & \text{for all } x \in M, \text{ there exists } y'_t \in M' \text{ s.t. for all } n, m \geq 1 \quad (51) \\ & \text{Vol}(M)e^{-(\lambda_n + \lambda_m)t/2} \langle X_n(x), X_m(x) \rangle = \text{Vol}(M')e^{-(\lambda'_n + \lambda'_m)t/2} \langle X'_n(y'_t), X'_m(y'_t) \rangle \\ & \text{for all } y' \in M', \text{ there exists } x_t \in M \text{ s.t. for all } n, m \geq 1 \\ & \text{Vol}(M)e^{-(\lambda_n + \lambda_m)t/2} \langle X_n(x_t), X_m(x_t) \rangle = \text{Vol}(M')e^{-(\lambda'_n + \lambda'_m)t/2} \langle X'_n(y'), X'_m(y') \rangle \end{aligned}$$

Because $\{\langle X_n(x), X_m(x) \rangle\}_{n, m=1}^\infty$ (resp. $\{\langle X_n(y'), X_m(y') \rangle\}_{n, m=1}^\infty$) separate the points² in the manifold by [Theorem 3.1](#), the point y'_t (resp. x_t) is uniquely defined and hence the corresponding map $f_t : x \rightarrow y'_t$ (resp. $h_t : y' \rightarrow x_t$) is well-defined, and it is clear that f_t and h_t are inverse to each other. It is also clear that f_t and h_t are continuous. Indeed, by [Lemma 3.2](#), the geodesic distances between x, \bar{x} and $y'_t = f_t(x), \bar{y}'_t = f_t(\bar{x})$ are related by:

$$\begin{aligned} d_g(x, \bar{x}) &= \frac{(4\pi)^{d/2}}{d^{1/2} \text{Vol}(M)} t^{\frac{d+1}{2}} d_{\text{VDM}, t}(x, \bar{x})(1 + O(t)) \\ &= \frac{(4\pi)^{d/2}}{d^{1/2} \text{Vol}(M')} t^{\frac{d+1}{2}} d_{\text{VDM}, t}(y'_t, \bar{y}'_t)(1 + O(t)) = d_{g'}(y'_t, \bar{y}'_t)(1 + O(t)), \quad (52) \end{aligned}$$

which implies the continuity of f_t . Similarly we get the continuity of h_t . Then, we show that f_t and h_t are C^∞ diffeomorphism. Define a map $F : M \times M' \rightarrow \mathbb{R}^d$ by

² A set $\{\langle X_n(x), X_m(x) \rangle\}_{n, m=1}^\infty$ separates points provided that for $x \neq y$ there exists $f \in \{\langle X_n(x), X_m(x) \rangle\}_{n, m=1}^\infty$ such that $f(x) \neq f(y)$.

$$F(x, y') = (\langle X_{n_i}(x), X_{m_i}(x) \rangle - c_{n_i, m_i}(t) \langle X'_{n_i}(y'), X'_{m_i}(y') \rangle)_{i=1}^d, \quad (53)$$

where

$$c_{n_i, m_i}(t) = \frac{\text{Vol}(M')}{\text{Vol}(M)} e^{(\lambda_{n_i} + \lambda_{m_i} - \lambda'_{n_i} - \lambda'_{m_i})t/2}. \quad (54)$$

Note that $F(h_t(y'), y') = 0$. Let $y'_0 = f_t(x_0)$. From [Lemma 4.5](#), it follows that the partial differentiation of F with related to the first variable at (x_0, y_0) is an isomorphism and hence h_t is locally smooth at y'_0 by the implicit function theorem. It follows that h_t is smooth. The same proof shows that f_t is smooth, too.

Next, we show that $\text{Vol}(M) = \text{Vol}(M')$. Denote the induced volume form $(f_t)_* dV_M$ by $a_t dV_{M'}$, where a_t is smooth. Integrating the relation [\(51\)](#), we obtain for $n \neq m$, $n, m \in \mathbb{N}$:

$$\begin{aligned} 0 &= \text{Vol}(M) e^{-(\lambda_n + \lambda_m)t/2} \int_M \langle X_n(x), X_m(x) \rangle dV_M(x) \\ &= \text{Vol}(M') e^{-(\lambda'_n + \lambda'_m)t/2} \int_M \langle X'_n(f_t(x)), X'_m(f_t(x)) \rangle dV_M(x) \\ &= \text{Vol}(M') e^{-(\lambda'_n + \lambda'_m)t/2} \int_{M'} \langle X'_n(y), X'_m(y) \rangle a_t(y) dV_{M'}(y), \end{aligned} \quad (55)$$

and similarly for $n = m$, $n \in \mathbb{N}$:

$$\begin{aligned} 1 &= \text{Vol}(M) e^{-\lambda_n t} \int_M \langle X_n(x), X_n(x) \rangle dV_M(x) \\ &= \text{Vol}(M') e^{-\lambda'_n t} \int_M \langle X'_n(f_t(x)), X'_n(f_t(x)) \rangle dV_M(x) \\ &= \text{Vol}(M') e^{-\lambda'_n t} \int_{M'} \langle X'_n(y), X'_n(y) \rangle a_t(y) dV_{M'}(y). \end{aligned} \quad (56)$$

In other words, we have when $n \neq m$,

$$\int_{M'} \langle X'_n(y), X'_m(y) \rangle a_t(y) dV_{M'}(y) = 0, \quad (57)$$

and when $n = m$,

$$\int_{M'} \langle X'_n(y), X'_n(y) \rangle a_t(y) dV_{M'}(y) = 1. \quad (58)$$

We claim that from (57) and (58), $a_t = 1$. Indeed, since a_t is smooth and $\{X_k\}_{k=1}^\infty$ form an orthonormal basis of $L^2(TM)$, by Lemma 4.4 and (57) we conclude that $a_t(y)$ is constant from (57). From (58), we know $a_t(y) = 1$. Hence,

$$\text{Vol}(M) = \int_M dV_M(x) = \int_{M'} (f_t)_* dV_{M'}(y) = \text{Vol}(M'). \quad (59)$$

Plug $\text{Vol}(M) = \text{Vol}(M')$ into (51). Integrating (51) gives $e^{-(\lambda_n + \lambda_m)t/2} = e^{-(\lambda'_n + \lambda'_m)t/2}$, which implies $e^{-\lambda_n t} = e^{-\lambda'_n t}$ for all $n \geq 1$ and hence $\lambda_n = \lambda'_n$ for all $n \geq 1$.

So far, (51) becomes: for all $n, m = 1, 2, \dots$, $\lambda_n = \lambda'_n$ and

$$\langle X_n(x), X_m(x) \rangle = \langle X'_n(f_t(x)), X'_m(f_t(x)) \rangle. \quad (60)$$

We now show that f_t is an isometry. Fix $p \in M$. Note that since f_t is a diffeomorphism, we can find an orthonormal frame $\{E_i\}_{i=1}^d$ around p and $\{E'_i\}_{i=1}^d$ around $f_t(p)$ so that $df_t|_p E_i = a_i(p) E'_i(f_t(p))$, where $a_i(p) > 0$. Indeed, we can find normal coordinates around p and $f(p)$ so that we have orthonormal bases $\{e_i\}_{i=1}^d$ of $T_p M$ and $\{e'_i\}_{i=1}^d$ of $T_{f_t(p)} M'$. Rewrite the linear transformation $f_{t*}|_p$ in the matrix form, denoted as $[f_{t*}|_p]$, with related to these bases. By the singular value decomposition we have $[f_{t*}|_p] = U \Lambda V^T$, where $V, U \in O(d)$. Clearly, $\{V e_i\}$ and $\{U e'_i\}$ are orthonormal bases of $T_p M$ and $T_{f_t(p)} M'$ respectively. The orthonormal frames E_i and E'_i are thus defined so that $E_i(p) = V e_i$ and $E'_i(f_t(p)) = U e'_i$. Denote $a_i(p) = \Lambda(i, i)$. To finish the proof, we have to show that $a_i(p) = 1$ for all $i = 1, \dots, d$. Choose a vector field $Z \in C^\infty(TM)$ so that

$$\begin{cases} Z(p) = 0, \quad \nabla Z(p) = 0, \\ \nabla_{E_1, E_1}^2 Z(p) \neq 0, \\ \nabla_{E_1, E_1}^2 Z(p) + \nabla_{E_2, E_2}^2 Z(p) = 0, \\ \nabla_{E_l, E_l}^2 Z(p) = 0 \text{ for all } l = 3, \dots, d. \end{cases} \quad (61)$$

Since $Z \in C^\infty(TM)$, we have $Z = \sum_{n=1}^\infty \alpha_n X_n$. Also, by the construction, we have $\Delta^{\text{TM}} Z(p) = 0$. To show the existence of this vector field, we could consider, for example, $Z = (x_1^2 - x_2^2)(\partial/\partial x_1) + (x_1^2 - x_2^2)(\partial/\partial x_2)$, where $\{x_i\}_{i=1}^d$ is the normal coordinate around p so that $(\partial/\partial x_i)(p) = E_i(p)$ for $i = 1, \dots, d$. Note that we have $x_i(p) = 0$, locally near p we have $\partial/\partial x_j(x) = E_j(x) - \frac{1}{6} \sum_i x_k x_l R_{kilj}(p) E_i(x) + \dots$, where R_{kilj} is the Riemannian curvature, and $\nabla_{E_i}(\partial/\partial x_j)(p) = 0$.

Denote the Levi-Civita connection of M' as ∇' and the associated connection Laplacian as Δ'^{TM} . Construct $Z' \in C^\infty(TM')$ by $Z' = \sum_{n=1}^\infty \alpha_n X'_n$, where $\{X'_n\}$ is the eigen-vector fields of Δ'^{TM} . We claim that

$$\begin{cases} Z'(f_t(p)) = 0, & \nabla' Z'(f_t(p)) = 0, \\ \nabla'^2_{E'_1, E'_1} Z'(f_t(p)) \neq 0, \\ \nabla'^2_{E'_1, E'_1} Z'(f_t(p)) + \nabla'^2_{E'_2, E'_2} Z'(f_t(p)) = 0, \\ \nabla'^2_{E'_l, E'_l} Z'(f_t(p)) = 0 \text{ for all } l = 3, \dots, d. \end{cases} \quad (62)$$

Note that this claim implies $\Delta'^{\text{TM}} Z'(f_t(p)) = 0$. By (60), we know

$$\begin{aligned} \langle Z'(f_t(p)), Z'(f_t(p)) \rangle &= \sum_{k,l=1}^{\infty} \alpha_k \alpha_l \langle X'_k(f_t(p)), X'_l(f_t(p)) \rangle \\ &= \sum_{k,l=1}^{\infty} \alpha_k \alpha_l \langle X_k(p), X_l(p) \rangle = \langle Z(p), Z(p) \rangle = 0 \end{aligned} \quad (63)$$

and

$$\begin{aligned} \langle \Delta'^{\text{TM}} Z'(f_t(p)), \Delta'^{\text{TM}} Z'(f_t(p)) \rangle &= \sum_{k,l=1}^{\infty} \alpha_k \lambda'_k \alpha_l \lambda'_l \langle X'_k(f_t(p)), X'_l(f_t(p)) \rangle \\ &= \sum_{k,l=1}^{\infty} \alpha_k \lambda_k \alpha_l \lambda_l \langle X_k(p), X_l(p) \rangle = \langle \Delta^{\text{TM}} Z(p), \Delta^{\text{TM}} Z(p) \rangle = 0, \end{aligned} \quad (64)$$

which implies $Z'(f_t(p)) = 0$ and $\Delta'^{\text{TM}} Z'(f_t(p)) = 0$. Take $v \in T_p M$ and a curve $\gamma : [0, \epsilon) \rightarrow M$ so that $\gamma(0) = p$ and $\gamma'(0) = v$. By extending v to $V \in \Gamma(TM)$ so that $V(\gamma(t)) = \gamma'(t)$, we have

$$\begin{aligned} \frac{d^2}{dt^2} \Big|_{t=0} \langle Z, Z \rangle(\gamma(t)) &= 2 \frac{d}{dt} \Big|_{t=0} \langle \nabla_V Z, Z \rangle(\gamma(t)) \\ &= 2 \langle \nabla_V \nabla_V Z, Z \rangle(p) + \langle \nabla_V Z, \nabla_V Z \rangle(p) = 2 \langle \nabla_V Z, \nabla_V Z \rangle(p), \end{aligned} \quad (65)$$

where the last equality holds since $Z(p) = 0$. On the other hand, by (60) we have

$$\begin{aligned} &\frac{d^2}{dt^2} \Big|_{t=0} \langle Z, Z \rangle(\gamma(t)) \\ &= \frac{d^2}{dt^2} \Big|_{t=0} \langle Z', Z' \rangle(f_t \circ \gamma(t)) \\ &= 2 \frac{d}{dt} \Big|_{t=0} \langle \nabla'_{f_{t*} V} Z', Z' \rangle(f_t \circ \gamma(t)) \\ &= 2 \langle \nabla'_{f_{t*} V} \nabla'_{f_{t*} V} Z', Z' \rangle(f_t(p)) + \langle \nabla'_{f_{t*} V} Z', \nabla'_{f_{t*} V} Z' \rangle(f_t(p)) \\ &= 2 \langle \nabla'_{f_{t*} V} Z', \nabla'_{f_{t*} V} Z' \rangle(f_t(p)), \end{aligned}$$

where the last equality comes from the fact that $Z'(f_t(p)) = 0$. Thus, we have

$$\langle \nabla'_{f_{t*} V} Z', \nabla'_{f_{t*} V} Z' \rangle(f_t(p)) = \langle \nabla_V Z, \nabla_V Z \rangle(p) = 0, \quad (66)$$

which implies $\nabla'_{f_{t*}v}Z'(f_t(p)) = 0$. Since f_t is diffeomorphic and v is arbitrary, we conclude that $\nabla'Z'(f_t(p)) = 0$.

Next choose the same curve γ and take the fourth order derivative:

$$\begin{aligned} & \frac{d^4}{dt^4}|_{t=0}\langle Z, Z \rangle(\gamma(t)) \\ &= 2\langle \nabla_V \nabla_V \nabla_V \nabla_V Z, Z \rangle(p) \\ & \quad + 8\langle \nabla_V \nabla_V \nabla_V Z, \nabla_V Z \rangle(p) \\ & \quad + 6\langle \nabla_V \nabla_V Z, \nabla_V \nabla_V Z \rangle(p) \\ &= 6\langle \nabla_V \nabla_V Z, \nabla_V \nabla_V Z \rangle(p) \\ &= 6\langle \nabla_{V,V}^2 Z + \nabla_{\nabla_V V} Z, \nabla_{V,V}^2 Z + \nabla_{\nabla_V V} Z \rangle(p) \\ &= 6\langle \nabla_{V,V}^2 Z, \nabla_{V,V}^2 Z \rangle(p) \end{aligned}$$

where the second and the fourth equalities come from the fact that $Z(p) = 0$ and $\nabla Z(p) = 0$. Similarly, by the fact that $Z'(f_t(p)) = 0$ and $\nabla'Z'(f_t(p)) = 0$ we have

$$\begin{aligned} & \frac{d^4}{dt^4}|_{t=0}\langle Z, Z \rangle(\gamma(t)) \\ &= \frac{d^4}{dt^4}|_{t=0}\langle Z', Z' \rangle(f_t \circ \gamma(t)) \\ &= 2\langle \nabla'_{f_{t*}V} \nabla'_{f_{t*}V} \nabla'_{f_{t*}V} \nabla'_{f_{t*}V} Z', Z' \rangle(f_t(p)) \\ & \quad + 8\langle \nabla'_{f_{t*}V} \nabla'_{f_{t*}V} \nabla'_{f_{t*}V} Z', \nabla'_{f_{t*}V} Z' \rangle(f_t(p)) \\ & \quad + 6\langle \nabla'_{f_{t*}V} \nabla'_{f_{t*}V} Z', \nabla'_{f_{t*}V} \nabla'_{f_{t*}V} Z' \rangle(f_t(p)) \\ &= 6\langle \nabla'_{f_{t*}V} \nabla'_{f_{t*}V} Z', \nabla'_{f_{t*}V} \nabla'_{f_{t*}V} Z' \rangle(f_t(p)) \\ &= 6\langle \nabla'^2_{f_{t*}V, f_{t*}V} Z' + \nabla'_{\nabla'_{f_{t*}V} f_{t*}V} Z', \nabla'^2_{f_{t*}V, f_{t*}V} Z' + \nabla'_{\nabla'_{f_{t*}V} f_{t*}V} Z' \rangle(f_t(p)) \\ &= 6\langle \nabla'^2_{f_{t*}V, f_{t*}V} Z', \nabla'^2_{f_{t*}V, f_{t*}V} Z' \rangle(f_t(p)). \end{aligned}$$

Since v is arbitrary, we have

$$\langle \nabla^2_{E_i, E_i} Z, \nabla^2_{E_i, E_i} Z \rangle(p) = \langle \nabla'^2_{f_{t*}E_i, f_{t*}E_i} Z', \nabla'^2_{f_{t*}E_i, f_{t*}E_i} Z' \rangle(f_t(p)). \quad (67)$$

Thus, we have shown the claim (62).

Next we claim that $a_1(p) = a_2(p)$. Take another smooth vector field $Y \in C^\infty(TM)$ so that

$$Y = \sum_{n=1}^{\infty} \gamma_n X_n \quad (68)$$

and $Y(p) \neq 0$ and construct $Y' \in C^\infty(TM)$ so that

$$Y' = \sum_{n=1}^{\infty} \gamma_n X'_n. \quad (69)$$

Then by taking the curve γ so that $\gamma'(0) = E_i$, where $i = 1, \dots, d$, we have

$$\begin{aligned} & \frac{d^2}{dt^2} \Big|_{t=0} \langle Z, Y \rangle (\gamma(t)) \\ &= \langle \nabla_{E_i} \nabla_{E_i} Z, Y \rangle (p) + 2 \langle \nabla_{E_i} Z, \nabla_{E_i} Y \rangle (p) + \langle Z, \nabla_{E_i} \nabla_{E_i} Y \rangle (p) \\ &= \langle \nabla_{E_i, E_i}^2 Z + \nabla_{\nabla_{E_i} E_i} Z, Y \rangle (p) \\ &= \langle \nabla_{E_i, E_i}^2 Z, Y \rangle (p) \end{aligned} \quad (70)$$

since $Z(p) = 0$ and $\nabla Z(p) = 0$. On the other hand, by the same arguments as those for (70), the construction of Z' and E'_i , we have

$$\begin{aligned} & \frac{d^2}{dt^2} \Big|_{t=0} \langle Z, Y \rangle (\gamma(t)) \\ &= \frac{d^2}{dt^2} \Big|_{t=0} \langle Z', Y' \rangle (f \circ \gamma(t)) \\ &= \langle \nabla_{f_{t*} E_i, f_{t*} E_i}^2 Z', Y' \rangle (f_t(p)) \\ &= a_i(p)^2 \langle \nabla_{E'_i, E'_i}^2 Z', Y' \rangle (f_t(p)) \end{aligned} \quad (71)$$

From (70) and (71), we have

$$\langle \nabla_{E_i, E_i}^2 Z, Y \rangle (p) = a_i(p)^2 \langle \nabla_{E'_i, E'_i}^2 Z', Y' \rangle (f_t(p)). \quad (72)$$

Hence, by the assumption of Z we have

$$\begin{aligned} 0 &= \langle \Delta^{\text{TM}} Z, Y \rangle (p) = \langle \nabla_{E_1, E_1}^2 Z + \nabla_{E_2, E_2}^2 Z, Y \rangle (p) \\ &= a_1(p)^2 \langle \nabla_{E'_1, E'_1}^2 Z', Y' \rangle (f_t(p)) + a_2(p)^2 \langle \nabla_{E'_2, E'_2}^2 Z', Y' \rangle (f_t(p)) \end{aligned} \quad (73)$$

Since $Y'(f_t(p))$ is arbitrary, we know

$$a_1(p)^2 \nabla_{E'_1, E'_1}^2 Z'(f_t(p)) + a_2(p)^2 \nabla_{E'_2, E'_2}^2 Z'(f_t(p)) = 0. \quad (74)$$

Combining with the fact that

$$\nabla_{E'_1, E'_1}^2 Z'(f_t(p)) + \nabla_{E'_2, E'_2}^2 Z'(f_t(p)) = 0 \quad (75)$$

in (62), we know $a_1(p) = a_2(p)$.

By repeating the arguments from (61) to (74), we conclude that

$$a_1(p) = a_2(p) = \dots = a_d(p) =: a(p). \quad (76)$$

To finish the proof, we choose $W \in C^\infty(TM)$ so that $W = \sum_{l=1}^\infty \beta_l X_l$, $W(p) = 0$, $\nabla W(p) = 0$ and $\Delta^{\text{TM}} W(p) \neq 0$. Construct $W' = \sum_{l=1}^\infty \beta_l X'_l$. By the same argument, we know $W'(f_t(p)) = 0$ and $\nabla' W'(f_t(p)) = 0$. Take arbitrary vector fields Y on M and Y' on M' constructed by the same way as those in (68) and (69). The same arguments for (70) and (71) hold for W ; that is, when $\gamma(t)$ is a curve on M so that $\gamma(0) = p$ and $\gamma'(0) = E_i(p)$ we have

$$\frac{d^2}{dt^2} \Big|_{t=0} \langle W, Y \rangle (\gamma(t)) = \langle \nabla_{E_i, E_i}^2 W, Y \rangle (p) \quad (77)$$

and

$$\begin{aligned} \frac{d^2}{dt^2} \Big|_{t=0} \langle W, Y \rangle (\gamma(t)) &= \frac{d^2}{dt^2} \Big|_{t=0} \langle W', Y' \rangle (f \circ \gamma(t)) \\ &= \langle \nabla_{f_{t*} E_i, f_{t*} E_i}^2 W', Y' \rangle (f_t(p)) = a(p)^2 \langle \nabla_{E'_i, E'_i}^2 W', Y' \rangle (f_t(p)). \end{aligned} \quad (78)$$

Thus we have

$$\begin{aligned} a(p)^2 \langle \Delta'^{\text{TM}} W', Y' \rangle &= a(p)^2 \left\langle \sum_{i=1}^d \nabla_{E'_i, E'_i}^2 W', Y' \right\rangle \\ &= \left\langle \sum_{i=1}^d \nabla_{E_i, E_i}^2 W, Y \right\rangle = \langle \Delta^{\text{TM}} W, Y \rangle. \end{aligned} \quad (79)$$

On the other hand, the following equality holds due to the definition of W' and Y' :

$$\begin{aligned} \langle \Delta'^{\text{TM}} W', Y' \rangle &= \left\langle \sum_{l=1}^\infty \beta_l \lambda'_l X'_l, \sum_{k=1}^\infty \gamma_k X'_k \right\rangle = \sum_{l,k=1}^\infty \beta_l \lambda'_l \gamma_k \langle X'_l, X'_k \rangle \\ &= \sum_{l,k=1}^\infty \beta_l \lambda_l \gamma_k \langle X_l, X_k \rangle = \langle \Delta^{\text{TM}} W, Y \rangle, \end{aligned} \quad (80)$$

which gives us $a(p) = 1$ and hence f_t is an isometry. We have thus finished the proof. \square

5. Precompactness of $\mathcal{M}_{d,k,D}$

By Theorem 4.6, we have a distance, referred to as the VSD, in the space of the isometry classes in $\mathcal{M}_{d,k,D}$. We finally can state the pre-compactness theorem. We need Lemma 1.1 to finish the proof.

Theorem 5.1. *For any $t > 0$, the space of the isometry classes in $\mathcal{M}_{d,k,D}$ is d_t -precompact.*

Proof. Denote h^1 to be the Hilbert space consisting of infinite sequences $\{x_i\}_{i=1,2,\dots}$ so that $\|\{x_i\}_{i=1,2,\dots}\|_{h^1} := \sum_{i=1}^{\infty} (1+i^{2/d})|x_i|^2 < \infty$. Consider a one-to-one map $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ which maps (i, j) to $\frac{(i+j-1)(i+j-2)}{2} + j$. Thus we could view the double indexed sequence $\{e^{-(\lambda_i+\lambda_j)t} \langle X_i(x), X_j(x) \rangle\}_{i,j=1}^{\infty}$ as an infinite sequence. Fix $t > 0$. For any $M \in \mathcal{M}_{d,k,D}$, $a \in \mathcal{B}(M, g)$ and $x \in M$, we have

$$\begin{aligned} \|V_t^a(x)\|_{h^1}^2 &= \text{Vol}(M)^2 \sum_{i,j \geq 1} \left[1 + \left(\frac{(i+j-1)(i+j-2)}{2} + j \right)^{2/d} \right] \\ &\quad \times e^{-(\lambda_i+\lambda_j)t} \langle X_i(x), X_j(x) \rangle^2 \\ &\leq \text{Vol}(M)^2 \sum_{i,j \geq 1} (1 + (i+j)^{4/d}) e^{-(\lambda_i+\lambda_j)t} \langle X_i(x), X_j(x) \rangle^2 \\ &\leq \frac{\text{Vol}(M)^2}{A(d, k, D)^2} \sum_{i,j \geq 1} (A(d, k, D)^2 + c(d)(\lambda_i + \lambda_j)^2) e^{-(\lambda_i+\lambda_j)t} \langle X_i(x), X_j(x) \rangle^2 \\ &\leq E(d, k, D) F(0, d) t^{-d} + \frac{c(d)E(d, k, D)}{A(d, k, D)^2} F(2, d) t^{-2-d}, \end{aligned}$$

where $c(d)$ is a universal constant depending on d , the first inequality holds by a simple bound $\frac{(i+j-1)(i+j-2)}{2} + j \leq (i+j)^2$, the second inequality follows from Lemma 4.2(a) and the third inequality follows from Lemma 4.2(c). Since $c(d)$, $A(d, k, D)$, $E(d, k, D)$, $F(0, d)$ and $F(2, d)$ are universal constants, we know that the set

$$K_0 := \{V_t^a(x)\}_{x \in M, M \in \mathcal{M}_{d,k,D}, a \in \mathcal{B}(M, g)} \subset h^1 \quad (81)$$

is bounded in $h^1 \subset \ell^2$, which is hence relative compact inside ℓ^2 by Rellich's Theorem [7]. Denote the closure of K_0 in ℓ^2 by K . Denote the set of all non-empty closed subsets of K by $\mathcal{F}(K)$, equipped with the Hausdorff distance HD associated with the canonical metric on ℓ^2 . By Lemma 1.1, the metric space $(\mathcal{F}(K), \text{HD})$ is precompact.

By Theorem 3.1, since $M \in \mathcal{M}_{d,k,D}$ is compact, $V_t^a(M)$ is compact inside ℓ^2 for any $a \in \mathcal{B}(M, g)$. Hence $V_t^a(M) \in \mathcal{F}(K)$. Consider a subset E of $\mathcal{F}(K)$ consisting of $V_t^a(M)$, where $M \in \mathcal{M}_{d,k,D}$, $a \in \mathcal{B}(M, g)$, that is,

$$E = \{V_t^a(M)\}_{M \in \mathcal{M}_{d,k,D}, a \in \mathcal{B}(M, g)} \subset \mathcal{F}(K). \quad (82)$$

Note that E is precompact with related to the distance HD since a subset of a compact set is precompact. Given $M \in \mathcal{M}_{d,k,D}$, define a subset $V_t(M)$ of E consisting of $V_t^a(M)$ for all $a \in \mathcal{B}(M, g)$, that is,

$$V_t(M) := \{V_t^a(M)\}_{a \in \mathcal{B}(M, g)} \subset E. \quad (83)$$

Here we view $V_t^a(M)$ as a point in the set E . By Lemma 4.3, $V_t(M)$ is a closed subset of E with related to the Hausdorff distance HD. Indeed, by the closeness of $\mathcal{B}(M, g)$ and Lemma 4.3 we have

$$\text{HD}(V_t^a(M), V_t^b(M)) \leq 2\text{Vol}(M)^2 d_{\mathcal{B}(M,g)}(a, b) \sup_{x \in M} |K_{TM}^{(N)}(t, x, x)|, \quad (84)$$

which implies the closeness of $V_t(M)$. Then, consider $\mathcal{F}(E)$ the set of non-empty closed subsets of E , equipped with the Hausdorff distance h_{HD} associated with the distance HD . By Lemma 1.1 again, we conclude that $\mathcal{F}(E)$ is precompact with related to the distance h_{HD} . Finally, the set $\{V_t(M)\}_{M \in \mathcal{M}_{d,k,D}}$, which is a subset of $\mathcal{F}(E)$, is precompact with related to the Hausdorff distance h_{HD} .

Notice that by the definition of the Hausdorff distance in (42) and Theorem 4.6, (43) is nothing but the Hausdorff distance h_{HD} , that is,

$$d_t(M, M') = d_{\text{HD}}(V_t(M), V_t(M')), \quad (85)$$

and we conclude the proof. \square

6. Manifold learning

Over the last couple of decades, explosive technological advances lead to exponential growth of massive datasets in almost all fields. In addition to the big size, the data is often high dimensional and difficult to quantify its hidden structure. In order to deal with this kind of data, several new techniques and algorithms were proposed in the applied mathematics and statistics literature. The geometric object, in particular the differentiable manifold, is commonly considered as a candidate model to capture or approximate the nonlinear structure inside the massive dataset, and several techniques based on spectral graph theory [9] are extensively studied in order to capture this structure [3,10]. Here we summarize the algorithm and the relevant theory.

Suppose we have a dataset $\mathcal{X} = \{x_i\}_{i=1}^n$ independently and uniformly sampled from a random variable. The first challenge to data scientists is finding a suitable metric to quantify the relationship among data points. Consider the special case that \mathcal{X} is located on a d -dim closed and smooth manifold M embedded in \mathbb{R}^p . Clearly we may use the canonical distance of \mathbb{R}^p to study the relationship. Once we have the pairwise relationship among data points, we are able to build up an affinity graph and study the dataset by studying its associated graph Laplacian. Precisely, set the vertices $V = \mathcal{X}$, the edges $E = \{(x_i, x_j)\}$ and the affinity function $w : E \rightarrow \mathbb{R}^+$ by $w(x_i, x_j) := e^{-\|x_i - x_j\|_{\mathbb{R}^p}^2 / \epsilon}$, where $\epsilon > 0$ is chosen by the user. Here w could be defined with other suitable kernels, but we present the algorithm with the Gaussian kernel for the simplicity. Second, build up a $n \times n$ matrix W and a $n \times n$ diagonal matrix D by

$$W_{ij} = w(x_i, x_j), \quad D_{ii} = \sum_{j=1}^n W_{ij}, \quad (86)$$

and then construct the graph Laplacian by $L = I_n - D^{-1}W$, where I_n is a $n \times n$ identity matrix. Clearly the constructed graph is undirected and W is symmetric. It has been well

studied that as $n \rightarrow \infty$ and $\epsilon \rightarrow 0$ depending on n , the graph Laplacian converges to the Laplace–Beltrami operator in the spectral sense in probability [4,24]. If we have further information about the parallel transport among two close points, then we can further build up the *connection graph*, which consists of the affinity graph discussed above and a connection function $\mathbf{g} : E \rightarrow O(d)$, where

$$\mathbf{g}(x_i, x_j) := b_i^T P_{x_i, x_j} b_j, \quad (87)$$

P_{x_i, x_j} is the parallel transport from x_j to x_i , and $b_i : \mathbb{R}^d \rightarrow T_{x_i}M$ can be viewed as a chosen basis of the tangent plane $T_{x_i}M$. In this case, we can establish the GCL in the following way. Recall that GCL stands for graph connection Laplacian. First, build up a $n \times n$ block matrix \mathbf{S} and a $n \times n$ block diagonal matrix \mathbf{D} with $d \times d$ blocks:

$$\mathbf{S}_{ij} = \mathbf{w}(x_i, x_j) \mathbf{g}(x_i, x_j) \in \mathbb{R}^{d \times d}, \quad \mathbf{D}_{ii} = \sum_{j=1}^n \mathbf{w}(x_i, x_j) I_d \in \mathbb{R}^{d \times d}, \quad (88)$$

and hence the GCL by $\mathbf{C} = I_{nd} - \mathbf{D}^{-1} \mathbf{S}$. By the construction of \mathbf{g} , we know \mathbf{S} is symmetric. As is shown in [24], under this setup, as $n \rightarrow \infty$ and $\epsilon \rightarrow 0$ depending on n , \mathbf{C} converges to the connection Laplacian in the spectral sense in probability. Furthermore, in [23, Theorem B1 and Theorem B2], it has been shown that via the techniques including local principal component analysis and rotational alignment, we are able to estimate the parallel transport between two close points directly from the point clouds. In other words, even if the parallel transport information among two close points is missing, we are able to extract it directly from the point clouds. We refer the reader to [10,23,24,13, 14] for more mathematical details as well as their practical algorithms. Note that all the above discussions are limited to one manifold.

Now one consider a different scenario – what happens if the dataset \mathcal{X} is more abstract than a subset of vectors? Take a set of 2-dim shapes embedded in \mathbb{R}^3 for example. In this setup, there has been several metrics proposed in the literature to compare two different shapes, for example [8,26,20] and the abundant literature summarized there. In general, the idea is that we view a shape as a manifold and the discretization of the shape as a set of point cloud sampled from the shape, and study its graph Laplacian or GCL. As is proved in [6] and this paper, if all the shapes in \mathcal{X} are of fixed dimension, bounded diameter and Ricci curvature, that is, $\mathcal{X} \subset \mathcal{M}_{d,k,D}$, then the SD and VSD among these shapes are indeed metrics distinguishing isometric classes inside \mathcal{X} . We can then take these metrics to define the affinity function \mathbf{w} in the affinity graph with $V = \mathcal{X}$ and $E = \{(x_i, x_j)\}$, and study the dataset by the graph Laplacian or other relevant techniques.

One immediate question we might ask when we take a new metric into account in the analysis – what is the property of \mathcal{X} under the SD or VSD? The pre-compactness theorem proved in [6] and this paper provide a partial answer. Indeed, we know that even these shapes might have quite diverse geometric and topological profiles, under

the SD or VSD the relationship among these manifolds is fairly restricted. Furthermore, since $\mathcal{M}_{d,k,D}$ is pre-compact under these metrics, we have control over several graph quantities. For example, if the constructed affinity graph is connected, we know the graph diameter is finite, and hence the first non-trivial eigenvalue is also non-trivial [9, Lemma 1.9]. A further study about the pre-compactness property and its application to the data analysis will be reported in the future work. Also, the main difference between the VSD and SD deserves a further study. In particular, since the connection Laplacian is intimately related to the Hodge Laplacian, the topological properties might be reflected in the VSD.

Acknowledgments

The author acknowledges the support partially by FHWA grant DTFH61-08-C-00028, partially by Award Number FA9550-09-1-0551 from AFOSR and partially by Connaught New Researcher grant 498992. He acknowledges Professor Charlie Fefferman and Professor Amit Singer for their time and inspiring and helpful discussions; in particular Professor Amit Singer, who introduced him the massive data analysis field. He also acknowledges the valuable discussion with Professor Gérard Besson, Professor Richard Bamler and Dr. Chen-Yun Lin. The author also thanks the anonymous reviewer's constructive and helpful comments.

References

- [1] B. Alexeev, A.S. Bandeira, M. Fickus, D.G. Mixon, Phase retrieval with polarization, *SIAM J. Imaging Sci.* 7 (1) (2013) 35–66.
- [2] J. Bates, The embedding dimension of Laplacian eigenfunction maps, *Appl. Comput. Harmon. Anal.* 37 (3) (2014) 516–530.
- [3] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, *Neural Comput.* 15 (6) (June 2003) 1373–1396.
- [4] M. Belkin, P. Niyogi, Convergence of Laplacian eigenmaps, in: *Proceedings of the 2006 Conference*, in: *Adv. Neur. In.*, vol. 19, The MIT Press, 2007, p. 129.
- [5] P. Bérard, *Spectral Geometry: Direct and Inverse Problems*, Springer, 1986.
- [6] P. Bérard, G. Besson, S. Gallot, Embedding Riemannian manifolds by their heat kernel, *Geom. Funct. Anal.* 4 (4) (1994) 373.
- [7] J. Bisgard, A compact embedding for sequence spaces, *Missouri J. Math. Sci.* 24 (2) (2012) 182–189.
- [8] F. Chazal, D. Cohen-Steiner, L.J. Guibas, F. Mémoi, S.Y. Oudot, Gromov–Hausdorff stable signatures for shapes using persistence, in: *Proceedings of the Symposium on Geometry Processing*, SGP'09, Eurographics Association, 2009, pp. 1393–1403.
- [9] F. Chung, *Spectral Graph Theory*, American Mathematical Society, 1996.
- [10] R.R. Coifman, S. Lafon, Diffusion maps, *Appl. Comput. Harmon. Anal.* 21 (1) (2006) 5–30.
- [11] M. Cucuringu, Y. Lipman, A. Singer, Sensor network localization by eigenvector synchronization over the Euclidean group, *ACM Trans. Sens. Netw.* 8 (3) (2012) 19:1–19:42.
- [12] M. Cucuringu, A. Singer, D. Cowburn, Eigenvector synchronization, graph rigidity and the molecule problem, *Inf. Inference J. IMA* 1 (2012) 21–67.
- [13] N. El Karoui, On information plus noise kernel random matrices, *Ann. Statist.* 38 (5) (2010) 3191–3216.
- [14] N. El Karoui, H.-T. Wu, Connection graph Laplacian methods can be made robust to noise, *Ann. Statist.* 44 (1) (2015) 346–372.
- [15] P. Gilkey, *The Index Theorem and the Heat Equation*, Princeton, 1974.

- [16] R. Hadani, A. Singer, Representation theoretic patterns in three dimensional cryo-electron microscopy II – the class averaging problem, *Found. Comput. Math.* 11 (5) (2011) 589–616.
- [17] P. Jones, M. Maggioni, R. Schul, Manifold parametrizations by eigenfunctions of the Laplacian and heat kernels, *Proc. Natl. Acad. Sci. USA* 105 (6) (2008) 1803–1808.
- [18] P. Jones, M. Maggioni, R. Schul, Manifold parametrizations by eigenfunctions of the Laplacian and heat kernels, *Ann. Acad. Sci. Fenn., Math.* 35 (2010) 131–174.
- [19] S. Marchesini, Y.-C. Tu, H.-T. Wu, Alternating projection, ptychographic imaging and phase synchronization, *Appl. Comput. Harmon. Anal.* (2015), <http://dx.doi.org/10.1016/j.acha.2015.06.005>.
- [20] F. Mémoli, A spectral notion of Gromov–Wasserstein distance and related methods, *Appl. Comput. Harmon. Anal.* 30 (3) (2011) 363–401.
- [21] J.W. Portegies, Embeddings of Riemannian manifolds with heat kernels and eigenfunctions, *Comm. Pure Appl. Math.* (2015), arXiv:1311.7568 [math.DG].
- [22] A. Singer, H.-T. Wu, Orientability and diffusion map, *Appl. Comput. Harmon. Anal.* 31 (1) (2011) 44–58.
- [23] A. Singer, H.-T. Wu, Vector diffusion maps and the connection Laplacian, *Comm. Pure Appl. Math.* 65 (8) (2012) 1067–1144.
- [24] A. Singer, H.-T. Wu, Spectral convergence of the connection Laplacian from random samples, *Inf. Inference J. IMA* (2016), in press, arXiv:1306.1587 [math.NA].
- [25] A. Singer, Z. Zhao, Y. Shkolnisky, R. Hadani, Viewing angle classification of cryo-electron microscopy images using eigenvectors, *SIAM J. Imaging Sci.* 4 (2) (2011) 543–572.
- [26] J. Sun, M. Ovsjanikov, L. Guibas, A concise and provably informative multi-scale signature based on heat diffusion, in: *Proceedings of the Symposium on Geometry Processing, SGP’09*, Eurographics Association, 2009, pp. 1383–1392.
- [27] Z. Zhao, A. Singer, Rotationally invariant image representation for viewing direction classification in cryo-em, *J. Struct. Biol.* 186 (1) (2014) 153–166.