J. DIFFERENTIAL GEOMETRY
99 (2015) 497-538

ISOMETRIC EMBEDDINGS VIA HEAT KERNEL

Xiaowei Wang & Ke Zhu

Abstract

For any n-dimensional compact Riemannian manifold M with smooth metric g, we construct a canonical t-family of isometric embeddings $I_t : M \to \mathbb{R}^{q(t)}$, with $t > 0$ sufficiently small and $q(t) >> t^{-\frac{n}{2}}$. This is done by intrinsically perturbing the heat kernel embedding introduced in [**BBG**]. As $t \to 0_+$, asymptotic geometry of the embedded images is discussed.

1. Introduction

Given an n-dimensional Riemannian manifold (M, g), one seeks for the embeddings $u : M \to \mathbb{R}^q$ for some q such that the induced metric is g, i.e., $u^* g_{can} = g$, where g_{can} is the standard Euclidean metric in $\mathbb{R}^q$. This is called the *isometric embedding problem* and has long history, with contributions from many people (see, e.g., [**G2**], [**HH**] for survey). In a celebrated paper [**N2**] in 1956, Nash proved the existence of global isometric embeddings of class C^s for $g \in C^s$, with $s \geq 3$ or $s = \infty$, and dimension $q_c = \frac{3}{2} n(n+1) + 4n$ in the compact case, $q = (n+1) q_c$ in the noncompact case.

Nash's proof used the so-called *hard implicit function theorem*, or Nash–Moser technique, which involves smoothing operators in the Newton iteration to preserve the differentiability of approximate solutions of the isometric problem. Günther (1989, [**G1**]) significantly simplified Nash's proof by inventing a *new iteration scheme* for the isometric embedding problem, such that there is no loss of differentiability in the iteration so the usual contracting mapping theorem is enough. A good exposition of his method can be found in [**G2**].

Nash and Günther's isometric embedding is very flexible. One can start with any *short embedding* as the approximate solution, i.e., any embedding $u : M \to \mathbb{R}^q$ such that the induced metric is less than or equal to g, to produce an isometric embedding (Nash's method was generalized in [**Gr1**]). On the other hand, such great flexibility of the initial embeddings usually makes the resulting isometric embeddings *noncanonical*. Their methods require one to perturb $f : M \to \mathbb{R}^q$ to a *free map*, i.e., the vectors $\{\partial_i f(x), \partial_j \partial_k f(x)\}_{1 \leq i,j,k \leq n,\, j \leq k}$ are linearly

Received 6/25/2013.

independent at every x on M (see Definition 13), and to estimate the right inverse of the matrix spanned by these vectors in order to apply the implicit function theorem.

Motivated by Kähler geometry and conformal geometry, one would like to find a *canonical* isometric embedding of a Riemannian manifold into S^{q-1}or $\mathbb{R}^q$ for $q \gg 1$, such that the corresponding geometry of the underlying Riemannian manifold is reflected via the symmetry groups on S^{q-1} or $\mathbb{R}^q$ (cf. [**Wa**]). To achieve that, one first has to construct such canonical embeddings. In 1994, Bérard, Besson, and Gallot [**BBG**] made progress by constructing an "asymptotically isometric embedding" of compact Riemannian manifolds M into ℓ^2, the space of real-valued and square summable series, using the *normalized heat kernel embedding* for $t > 0$:

$$\Psi_t : x \to \sqrt{2}\,(4\pi)^{n/4}\, t^{\frac{n+2}{4}} \cdot \left\{ e^{-\lambda_j t/2} \phi_j\,(x) \right\}_{j \geq 1}, \tag{1.1}$$

where λ_j is the jth eigenvalue of the Laplacian Δ_g of (M, g) and $\{\phi_j\}_{j\geq 0}$ is the L^2 orthonormal eigenbasis of Δ_g. The advantage is that the embeddings $\Psi_t : M \to \ell^2$ are *canonical,* in the sense that they are uniquely determined by the spectral geometry of (M, g). Moreover, when $t \to 0_+$ the embedding Ψ_t tends to an isometry in the following sense:

$$\Psi_t^* g_{can} = g + \frac{t}{3}\left(\frac{1}{2} S_g \cdot g - Ric_g\right) + O\left(t^2\right), \tag{1.2}$$

where g_{can} is the standard Euclidean metric in ℓ^2, S_g and Ric_g are scalar and Ricci curvatures of (M, g), respectively, and the convergence is in the C^r sense for any $r \geq 0$ (see [**BeGaM**, p. 213]). However, for any given $t > 0$, Ψ_t usually is only *asymptotically* isometric (with an error of order $O\,(t)$).

So we are in the following situation: Nash's embedding is isometric but *far from being canonical*, and the heat kernel embedding is canonical but only *asymptotically isometric*. In this paper, we are able to produce a *canonical* isometric embedding into $\mathbb{R}^q$ for $q \gg 1$ by modifying the almost isometric embedding Ψ_t in [**BBG**] to a better approximation with error bounded by $O\left(t^l\right)$ for any given $l \geq 2$, and then perturbing it to an isometry by Günther's theory ([**G2**]). Fixing two constants $\rho > 0$ and $0 < \alpha < 1$ throughout our paper, we have our main theorem, as follows.

Theorem 1. *Let (M, g) be a smooth n-dimensional compact Riemannian manifold without boundary. g is a smooth Riemannian metric on M. Then:*

1) *For any integer $l \geq 1$, there exist a canonical family of* almost *isometric embeddings $\tilde{\Psi}_t : M \to \ell^2$ such that*

$$\tilde{\Psi}_t^* g_{can} = g + O\left(t^l\right)$$

as $t \to 0_+$, where the above convergence is in C^r-norm for any $r \geq 0$.

2) *For any integer $k \geq 2$ satisfying $k + \alpha < l + \frac{1}{2}$, there exists a constant $t_0 > 0$ depending on k, α, l, ρ, and g, such that for $0 < t \leq t_0$, we can truncate $\tilde{\Psi}_t$ to $\mathbb{R}^{q(t)} \subset \ell^2$ and perturb it to a unique $C^{k,\alpha}$ isometric embedding*

$$I_t : M \to \mathbb{R}^{q(t)},$$

where dimension $q(t) \geq t^{-\frac{n}{2}-\rho}$ or $q(t) = \infty$, and $\left\| I_t - \tilde{\Psi}_t \right\|_{C^{k,\alpha}(M)} = O\left(t^{l+\frac{1}{2}-\frac{k+\alpha}{2}}\right)$.

The isometric embedding I_t is *canonical* in the sense that our $\tilde{\Psi}_t : M \to \ell^2$ is canonically constructed from Ψ_t in [**BBG**] (see Section 2) and our implicit function theorem only uses the *intrinsic* information of (M, g). More precisely, the smoothing operator needed in Günther's iteration scheme was constructed *directly* from g. The iteration attempts to adjust $\tilde{\Psi}_t$ to the nearest isometric embedding in each step with a *unique* minimal movement.

Our method has the following advantages. The heat kernel embedding $\Psi_t : M \to \mathbb{R}^{q(t)}$ is automatically a *free mapping* for small t. Furthermore, the row vectors $\{\partial_i \Psi_t(x)\}_{i=1}^n$ and $\{\partial_i \partial_j \Psi_t(x)\}_{1 \leq i \leq j \leq n}$ span a matrix $P(\Psi_t)$ with an *explicit right inverse bound* (Corollary 31) on the *whole* M. (These nice properties are inherited to $\tilde{\Psi}_t$ as well.) There is no need of sophisticated perturbation arguments in local charts to achieve the right inverse bound as was done in Nash and Günther's method.

We have good control of the second fundamental form and mean curvature of the embedded images $\Psi_t(M)$ and $I_t(M)$ in $\mathbb{R}^{q(t)}$. For simplicity, we only state the mean curvature part:

Proposition 2. Let M be a smooth n-dimensional compact Riemannian manifold with smooth metric. For any x on M, let $H(x,t)$ be the mean curvature vector at $\Psi_t(x)$ (or $I_t(x)$) in $\mathbb{R}^{q(t)}$. Then, as $t \to 0_+$,

$$\sqrt{t}\,|H(x,t)| \to \sqrt{\frac{n+2}{2n}}.$$

The second fundamental form also has certain normal form as $t \to 0_+$. (See Corollary 38, Remark 39 and its following paragraph).

Finally, we make a few remarks about our approach. First, in [**BBG**] the authors also constructed the heat kernel embedding into the infinite dimensional *unit sphere* $S^\infty \subset \ell^2$, by

$$K_t : x \to \left\{ e^{-\lambda_j t/2} \phi_j(x) \right\}_{j \geq 1} \Big/ \left(\Sigma_{j \geq 1} e^{-\lambda_j t} \phi_j^2(x) \right)^{1/2},$$

with the asymptotic behavior $K_t^* g_{can} = \frac{1}{2t}\left(g - \frac{t}{3}Ric_g + O\left(t^2\right)\right)$ as $t \to 0_+$, where g_{can} is the standard metric in ℓ^2. There is a parallel version of Theorem 1 for $S^{q(t)}$ provable by our method; i.e., we can truncate and perturb K_t to $\Upsilon_t : (M,g) \to S^{q(t)} \subset \mathbb{R}^{q(t)+1}$ such that $\Upsilon_t^* g_{can} = \frac{1}{2t}g$. But for the sake of simplicity, we write only the one for $\mathbb{R}^{q(t)}$. In the sequel to this paper, we will extend our method to construct a canonical *conformal embedding* $\Theta_t : (M,g) \to S^q \subset \mathbb{R}^{q+1}$ that keeps more information of K_t. It will be interesting to see the relation to the conformal volume defined in [**LY2**]. Second, we want to point out that it is not our main emphasis to optimize the dimension $q(t) \geq t^{-\frac{n}{2}-\rho}$, since the lower the dimension of embedding is, the less canonical the map is. Third, it will not make too much difference if one uses the heat kernel for a perturbation of the Laplacian operator (cf. [**Gil**]), but we choose to use the canonical one. Last, we only deal with compact Riemannian manifolds, while it is likely that our method can be extended to Riemannian manifolds with boundary or even complete Riemannian manifolds (with suitable condition at infinity). We leave these to future investigation. We notice there are several related works on embeddings of compact Riemannian manifolds by eigenfunctions (with various weights) and heat kernels recently, e.g., [**Ni**], [**Wu**], [**P**], and [**Po**].

The organization of the paper is follows: In Section 2 we review the heat kernel embedding $\Psi_t : M \to \ell^2$ in [**BBG**]. Then we modify Ψ_t to get improved error to isometry, and truncate the embedding to $\mathbb{R}^q \subset \ell^2$ and estimate the remainder. In Section 3 we recall the matrix $E(u)$ that appeared in the linearization of the isometric embedding problem, and we review Günther's iteration scheme and the implicit function theorem. In Section 4 we give higher derivative estimates of Ψ_t using the off-diagonal heat kernel expansion method, and establish the crucial uniform linear independence property of the matrix $E(\Psi_t)$. Then we give the operator norm estimate of $E(\Psi_t)$. In Section 5 we establish the uniform quadratic estimate of the nonlinear operator $Q(u)$ in the isometric embedding problem for all $\mathbb{R}^q$. In Section 6 we apply Günther's implicit function theorem to the modified Ψ_t to obtain isometric embeddings $I_t : (M,g) \to \mathbb{R}^{q(t)}$. The geometry of $I_t(M)$ is close to that of $\Psi_t(M)$ by our error estimate. In Section 7 we derive the asymptotic formulae of the second fundamental form and mean curvature of the embedded images $\Psi_t(M)$ as $t \to 0_+$. In Section 8 we illustrate our method by explicit calculations on $M = S^1$. In the appendix we make the constant in Günther's implicit function theorem explicit and discuss the minimal embedding dimension of our method.

Convention: In this paper, unless otherwise remarked, the constant C only depends on (M,g), its dimension n, and k, α in the $C^{k,\alpha}$-Hölder norm, but not on t and q of $\mathbb{R}^q$. In a sequence of inequalities, the constant C in successive appearances can be assumed to increase. The two constants $\rho > 0$ and $0 < \alpha < 1$ are fixed throughout the paper. The constant k in the $C^{k,\alpha}$-norm and the constant l in the error term $O\left(t^l\right)$ should not be confused with the indices k, l $(1 \leq k, l \leq n)$ in partial derivatives like ∂_k and $\partial_k \partial_l$.

Acknowledgments. We thank Gang Liu for an inspiring discussion on the off-diagonal heat kernel expansion, Deane Yang for helpful email correspondence, and Jiaping Wang for helpful discussions. Both authors would like to thank Professor S.-T. Yau for his influence on canonical embeddings. The second named author would like to thank Professor Clifford Taubes for his interest and support. He also thanks Peng Gao, Baosen Wu, and Qingchun Ji for discussions on the embedding problem in wider contexts. The work of K. Zhu is partially supported by a grant of Professor Clifford Taubes from the National Science Foundation. The work of X. Wang is partially supported by a Research Council Grant from Rutgers University and a Collaboration Grant for Mathematicians from the Simons Foundation.

2. The heat kernel embedding into ℓ^2 and modifications

Let ℓ^2 be the Hilbert space of real series $\{a_i\}_{i\geq 1}$ such that $\sum_{i=1}^{\infty} a_i^2 < \infty$, and g_{can} be the standard metric in ℓ^2. Let (M, g) be an n-dimensional compact Riemannian manifold with smooth metric g, and let $\{\phi_j(x)\}_{j\geq 0} \subset C^{\infty}(M)$ be a L^2-orthonormal basis of real eigenfunctions of the Laplacian of M, i.e., for eigenvalues $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$, $\Delta_g \phi_j = \lambda_j \phi_j$, and $\int_M \phi_i \phi_j dvol_g = \delta_{ij}$ for $\forall i, j$. The heat kernel of (M, g) is

$$H(t, x, y) = \Sigma_{s=1}^{\infty} e^{-\lambda_s t} \phi_s(x) \phi_s(y)$$

for $x, y \in M$ and $t > 0$.

Definition 3. We call the family of maps

$$\Phi_t : \begin{array}{ccc} M & \longrightarrow & \ell^2 \\ x & \longmapsto & \left\{e^{-\lambda_j t/2} \phi_j(x)\right\}_{j\geq 1} \end{array} \quad \text{for } t > 0$$

the *heat kernel embeddings*, and call $\Psi_t = \sqrt{2}(4\pi)^{n/4} t^{\frac{n+2}{4}} \cdot \Phi_t$ the *normalized heat kernel embeddings.*

From the definition, we clearly have $H(t, x, y) = \langle \Phi_t(x), \Phi_t(y) \rangle$, where $\langle , \rangle$ is the standard inner product in ℓ^2.

In [**BBG**], Bérard, Besson, and Gallots introduced the above maps and proved the following.

Theorem 4 ([**BBG**] Theorem 5)**.** *As $t \to 0_+$, there is an expansion*

$$\Psi_t^* g_{can} = g + \sum_{i=1}^{l} t^i A_i(g) + O\left(t^{l+1}\right), \tag{2.1}$$

with

$$A_1 = \frac{1}{3}\left(\frac{1}{2}S_g \cdot g - Ric_g\right),$$

where the A_i's are universal polynomials of the covariant differentiations of the metric g and its curvature tensors up to order $2i$, and the convergence in (2.1) *is in the C^r sense for any $r \geq 0$.*

As a direct consequence, we have the following singular perturbation result parallel to Theorem 26 in [**Don**]. Denote the space of symmetric 2-tensors on M by $\Gamma\left(\mathrm{Sym}^{\otimes 2}\left(T^*M\right)\right)$.

Proposition 5. For any $l \geq 1$, there are $h_i \in \Gamma\left(\mathrm{Sym}^{\otimes 2}\left(T^*M\right)\right)$ $(1 \leq i \leq l-1)$ such that for the family of metrics

$$g(s) = g + \sum_{i=1}^{l-1} s^i h_i,$$

the induced metric from the heat kernel embeddings $\Psi^*_{t,g(s)} : (M, g(s)) \to \ell^2$ satisfies the estimate

$$\left\| \Psi^*_{t,g(t)} g_{can} - g \right\|_{C^r(g)} \leq C(g, l, r)\, t^l, \tag{2.2}$$

for any $r \geq 0$, where the constant $C(g,l,r)$ depends only on l, r and the geometry of (M, g).

Proof. Let us assume that

$$g(s) = g + \sum_{i=1}^{l-1} s^i h_i \text{ with } h_i \in \Gamma\left(\mathrm{Sym}^{\otimes 2}\left(T^*M\right)\right), \tag{2.3}$$

where the h_i's are to be determined. Then by (2.1), for metric $g(s)$ we have

$$G(s,t) := \Psi^*_{t,g(s)} g_{can} = g(s) + tA_1(g(s)) + t^2 A_2(g(s)) + \cdots \tag{2.4}$$

with A_i's universal polynomials of the covariant differentiations of any metric and its curvature tensors up to order $2i$. Using the Taylor expansion of $A_i(g(s))$ at $s = 0$, we have

$$A_i(g(s)) = A_i(g) + \sum_{j=1}^{\infty} A_{i,j}(h_1, \cdots h_j)\, s^j,$$

where each

$$A_{i,j}(h_1, \cdots h_j) := \left.\frac{\partial^j}{\partial s^j}\right|_{s=0} \frac{1}{j!} A_i(g(s))$$

is a universal polynomial of the covariant differentiations of the metric g and its curvature tensors and is multi-linear in $h_1, \cdots h_j$ by the chain rule. Putting this into (2.1), we have in the C^r norm convergence

$$\begin{aligned} G(t,t) &= \left(g + th_1 + t^2 h_2 + \cdots\right) \\ &\quad + t\left(A_1(g) + A_{1,1}(h_1)\,t + A_{1,2}(h_1,h_2)\,t^2 + \cdots\right) \\ &\quad + t^2\left(A_2(g) + A_{2,1}(h_1)\,t + A_{2,2}(h_1,h_2)\,t^2 + \cdots\right) + \cdots \end{aligned}$$

$$+t^l\left(A_{l-1}(g) + A_{l-1,1}(h_1)\,t + A_{l-1,2}(h_1,h_2)\,t^2 + \cdots\right) + O\left(t^l\right). \tag{2.5}$$

Now we let $h_1 = -A_1(g)$, $h_2 = -A_{1,1}(h_1) - A_2(g)$, and in general let

$$\begin{aligned} h_j :=& -A_{1,j-1}(h_1,\cdots,h_{j-1}) - A_{2,j-2}(h_1,\cdots,h_{j-2}) \\ & - \cdots A_{j-1,1}(h_1) - A_j(g) \end{aligned}$$

inductively for $1 \leq j \leq l-1$. Then we are able to construct a curve $g(s) \subset \Gamma\left(\mathrm{Sym}^{\otimes 2}(T^*M)\right)$ by (2.3) such that

$$\Psi^*_{t,g(t)} g_{can} = G(t,t) = g + O\left(t^l\right)$$

in the C^r sense for any $r \geq 0$, as we claimed. q.e.d.

Definition 6 (Modified heat kernel embedding)**.** We call the $\Psi_{t,g(t)} : M \to \ell^2$ constructed above the *modified heat kernel embedding*, and denote

$$\tilde{\Psi}_t := \Psi_{t,g(t)}. \tag{2.6}$$

To get the embedding into $\mathbb{R}^q$, let

$$\Pi_q : \ell^2 \longrightarrow \mathbb{R}^q \tag{2.7}$$

be the projection of ℓ^2 to the first q components. To get a finite-dimensional isometric embedding, we introduce the truncated embedding

$$\Psi_t^{q(t)} := \Pi_q \circ \Psi_t : (M,g) \longrightarrow \ell^2 \xrightarrow{\Pi_q} \mathbb{R}^{q(t)}.$$

Remark 7. Since the metrics $g(s)$ constructed in (2.3) depend on s analytically, given any $\mu_0 > 0$ not in the spectrum of Δ_g, there exists $\delta_0 > 0$, such that for Δ_{g_s} with $0 \leq s < \delta_0$, for their eigenvalues $0 = \lambda_0 \leq \lambda_1(s) \leq \cdots \leq \lambda_{j_0}(s) < \mu_0$, the total multiplicity j_0 is independent on s and each $\lambda_j(s)$ $(0 \leq j \leq j_0)$ depends on s analytically. Furthermore, we can choose the eigenfunctions $\phi_j(s,x)$ of Δ_{g_s} associated with these $\lambda_j(s)$ such that they are orthonormal in $L^2(M,g_s)$ and depend on s analytically (see [**A**, Lemma 2.1] and earlier [**R**]). Therefore, for $0 \leq t < \delta_0$, the truncated heat kernel mapping $\Psi_t^{j_0} : M \to \mathbb{R}^{j_0}$ can be made depending on t analytically.

In order to estimate the truncated tail, we recall the following well-known derivative estimates of eigenfunctions ϕ_j and extend it to the Hölder derivative setting.

Lemma 8. *For any integer $k \geq 0$ and $0 < \alpha < 1$, we have*

$$\left\|\nabla^{(k)}\phi_j\right\|_{C^0(M)} \leq C(k,g)\,\lambda_j^{\frac{n+2k}{4}}, \tag{2.8}$$

$$\|\phi_j\|_{C^{k,\alpha}(M)} \leq C(k,\alpha,g)\,\lambda_j^{\frac{n+2k+2\alpha}{4}} \tag{2.9}$$

for some positive constants $C(k,g)$ and $C(k,\alpha,g)$.

Proof. The estimate (2.8) is in Theorem 17.5.3 of [**H2**] and Theorem 1 of [**X**] (when $k=0$ in earlier [**H1**] and [**S**], and $k=1$ in [**LY1**]). We only prove (2.9). Since $\lambda_j \to +\infty$ as $j \to \infty$, starting from some j we must have $\lambda_j^{-\frac{1}{2}}$ less than the injective radius of (M,g). Without loss of generality we assume this holds from $j=1$. We consider two cases:

For $x, y \in M$ with $d(x,y) \leq \lambda_j^{-\frac{1}{2}}$, we have

$$\begin{aligned}
\left|\frac{\nabla^{(k)}\phi_j(x) - \nabla^{(k)}\phi_j(y)}{(d(x,y))^{\alpha}}\right| &= \left|\frac{\nabla^{(k)}\phi_j(x) - \nabla^{(k)}\phi_j(y)}{d(x,y)}\right|(d(x,y))^{1-\alpha} \\
&\leq C(g)\left|\nabla^{(k+1)}\phi_j\right|_{C^0(M)}\lambda_j^{-\frac{1-\alpha}{2}} \\
(\text{by } (2.8)) \quad &\leq C(g)\,C(k+1,g)\,\lambda_j^{\frac{n+2k+2\alpha}{4}},
\end{aligned}$$

where the constant $C(g)$ only depends on g.

For $x, y \in M$ with $d(x,y) \geq \lambda_j^{-\frac{1}{2}}$, we have

$$\begin{aligned}
\left|\frac{\nabla^{(k)}\phi_j(x) - \nabla^{(k)}\phi_j(y)}{(d(x,y))^{\alpha}}\right| &\leq \frac{2\left|\nabla^{(k)}\phi_j\right|_{C^0(M)}}{\lambda_j^{-\frac{\alpha}{2}}}(\text{by } (2.8)) \\
&\leq 2C(k,g)\,\lambda_j^{\frac{n+2k+2\alpha}{4}}.
\end{aligned}$$

Combining the two cases and letting

$$C(k,\alpha,g) := \max\{C(g)\,C(k+1,g), 2C(k,g)\},$$

we obtain (2.9). q.e.d.

Proposition 9. Let $\{g_s\}_{s\in K}$ be a compact family of smooth metrics on a compact n-dimensional Riemannian manifold M, where g_s depends on s smoothly. Given $x \in M$, let $\left\{x^k\right\}_{1\leq k\leq n}$ be the normal coordinates in its neighborhood. Then for any multiple-indices $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$, and $q(t) \geq t^{-\left(\frac{n}{2}+\rho\right)}$,

$$\Sigma_{j\geq q(t)+1}e^{-\lambda_j t}D^{\overrightarrow{\alpha}}\phi_j(x)\,D^{\overrightarrow{\beta}}\phi_j(x) \leq C\exp\left(t^{-\frac{\rho}{n}}\right) \tag{2.10}$$

for any $l \geq 1$. The convergence is uniform for $x \in M$ and $s \in K$ in the C^r-norm for any $r \geq 0$.

Proof. We first prove the proposition when K consists of a *single* metric g. By Lemma 8 we have for any multi-index $\overrightarrow{\alpha}$,

$$\left\| D^{\overrightarrow{\alpha}}\phi_j \right\|_{C^0(M)} \leq C\left(\overrightarrow{\alpha}, g\right)\lambda_j^{\frac{n+2|\overrightarrow{\alpha}|}{4}} \tag{2.11}$$

for some constant $C\left(\overrightarrow{\alpha}, g\right)$, with $|\overrightarrow{\alpha}|$ being the degree of $\overrightarrow{\alpha}$. Combining *Weyl's asymptotic formula* ([**Ch**, p. 9]) for eigenvalues on compact manifolds (M, g) that

$$\lambda_j \sim \frac{4\pi^2}{(\omega_n Vol(M))^{\frac{2}{n}}} j^{\frac{2}{n}} \geq A\left(g\right) j^{\frac{2}{n}} \tag{2.12}$$

for some constant $A := A\left(g\right)$ as $j \to \infty$ (where ω_n is the volume of the unit ball in $\mathbb{R}^n$), we have

$$\begin{aligned}
& \left|\Sigma_{j\geq q(t)} e^{-\lambda_j t} D^{\overrightarrow{\alpha}}\phi_j\left(x\right) D^{\overrightarrow{\beta}}\phi_j\left(x\right)\right| \\
\leq\ & C\Sigma_{j\geq q(t)}\left(j^{\frac{2}{n}}\right)^{\frac{n+|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}} e^{-Aj^{\frac{2}{n}}t} \leq C\int_{q(t)}^{\infty} j^{\frac{n+|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{n}} e^{-Aj^{\frac{2}{n}}t} dj \\
\leq\ & Ct^{-\left(\frac{2n+|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|}{2}\right)} \int_{A(q(t))^{\frac{2}{n}}t}^{\infty} \mu^{\frac{2n+|\overrightarrow{\alpha}|+\left|\overrightarrow{\beta}\right|-2}{2}} e^{-\mu} d\mu \ \left(\mu = Aj^{\frac{2}{n}}t\right) \\
\leq\ & C\exp\left(t^{-\frac{\rho}{n}}\right),
\end{aligned}$$

where we have used $q\left(t\right) \geq t^{-\left(\frac{n}{2}+\rho\right)}$ and $\mu^{\sigma} = o\left(e^{\frac{\mu}{2}}\right)$ as $\mu \to \infty$ for any fixed $\sigma > 0$. Therefore, we have proved (2.10) in the C^0-convergence. The C^r-convergence of (2.10) follows by the Leibniz rule, adding the indices $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$ by $\overrightarrow{\gamma}$ with $|\overrightarrow{\gamma}| \leq r$ in the above argument.

For a compact family of metrics $\{g_s\}_{s\in K}$ smoothly depending on s, notice the constant $C\left(\overrightarrow{\alpha}, g_s\right)$ in (2.11) has a uniform upper bound and the constant $A\left(g_s\right)$ in (2.12) has a uniform positive lower bound for all $s \in K$, because they are determined by the following geometric quantities continuously depending on s: the dimension, the curvature bound, the diameter, and volume of (M, g_s) (see Remark 11). So the truncation estimate (2.10) can be made uniform for all $s \in K$. q.e.d.

Corollary 10. *Given any $l \geq 1$, for $q = q\left(t\right) \geq Ct^{-\left(\frac{n}{2}+\rho\right)}$, the truncated modified heat kernel embedding $\tilde{\Psi}_t^{q(t)} : (M, g) \to \mathbb{R}^{q(t)}$ still satisfies the asymptotic formula*

$$\left(\tilde{\Psi}_t^{q(t)}\right)^* g_{can} = g + O\left(t^l\right)$$

in the C^r-sense for any $r \geq 0$.

Proof. From Proposition 5 we have in the C^r-sense

$$\left(\tilde{\Psi}_t\right)^* g_{can} = g + O\left(t^l\right).$$

To see it still holds after truncating $\tilde{\Psi}_t$ to $\tilde{\Psi}_t^{q(t)}$, let $D^{\overrightarrow{\alpha}} = \nabla_i$, $D^{\overrightarrow{\beta}} = \nabla_j$ for local normal coordinates, let $\{g_s\}_{s\in[0,t_0]}$ be the compact family of metrics defined in (2.3), and then apply the above proposition. q.e.d.

Remark 11. To get an effective truncation of ℓ^2 to $\mathbb{R}^q$, it is useful to have an estimate of the jth eigenvalue λ_j of M in terms of geometric quantities of (M,g), since the Weyl asymptotic formula (2.12) does not tell how fast the λ_j converges to its limit in (2.12). Given a real number Λ, for all n-dimensional compact Riemannian manifolds (M,g) satisfying Ricci curvature $Ric_g \geq \Lambda g$ and diameter bounded by D, there exists a constant $A(n,\Lambda,D) > 0$ such that $\lambda_j \geq A(n,\Lambda,D)\, j^{\frac{2}{n}}$ ([**SY**, **Gr2**, **BBG**]). (Similar lower bound of λ_j was established earlier in [**LY1**] under stronger assumptions). The upper bound $\lambda_j \leq B(n,\Lambda,D)\, j^{\frac{2}{n}}$ was established in [**LY1**].

The estimate of $\|\phi_j\|_{C^{k,\alpha}(M)}$ for $k \geq 2$ can be reduced to $\|\phi_j\|_{C^1(M)}$ by inductively using the elliptic estimate

$$\|\phi_j\|_{C^{k,\alpha}(M)} \leq C_e \left(\|\Delta_g \phi_j\|_{C^{k-1,\alpha}(M)} + \|\phi_j\|_{C^0(M)}\right),$$

where the constant C_e depends on $n, D, Vol(M)$, and the sectional curvature bound K. In $\|\phi_j\|_{C^1(M)} \leq C(1,g)\,\lambda_j^{\frac{n+1}{2}}$, the constant $C(1,g)$ depends on the these quantities too (e.g. [**WaZh**]). Hence n, K, D, and $Vol(M)$ determine $C(k,\alpha,g)$.

Remark 12. Recently, [**Po**] studied a similar almost isometric embedding of compact Riemannian manifolds into Euclidean spaces via heat kernel plus certain recording points on M, with a weaker regularity assumption on g. The embedding dimension is controlled by similar geometric quantities in Remark 11.

3. Günther's iteration for isometric embedding

3.1. The perturbation problem and free mappings. To solve the isometric embedding problem $du\cdot du = g$, Nash studied the perturbation problem $d(u+v)\cdot d(u+v) = du\cdot du + f$ for small symmetric 2-tensors f. In local coordinates $\{x_i\}_{1=1}^n$, the perturbation $v : M \to \mathbb{R}^q$ should satisfy $\partial_i u \cdot \partial_j v + \partial_j u \cdot \partial_i v + \partial_i v \cdot \partial_j v = f_{ij}$. Imposing the condition $\partial_i u \cdot v = 0$, the equation becomes the system

$$\partial_i u \cdot v = 0, \quad \partial_j \partial_i u \cdot v = -\frac{1}{2} f_{ij} - \frac{1}{2}\partial_i v \cdot \partial_j v. \tag{3.1}$$

The *linear part* of the system is determined by a matrix whose row vectors are the $\frac{n(n+3)}{2}$ vectors $\partial_i u(x)$ and $\partial_j \partial_k u(x)$ in $\mathbb{R}^q$. This motivates the following definition.

Definition 13 (Free mapping)**.** A C^2 map $u : M \to \mathbb{R}^q$ (including ℓ^2, if $q = \infty$) is called a *free mapping* if the $\frac{n(n+3)}{2}$ vectors $\{\partial_i u(x),$ $\partial_j \partial_k u(x)\}_{1\le i,j,k\le n,\ j\le k}$ in $\mathbb{R}^q$ are linearly independent at any $x \in M$, where ∂_i is the derivative with respect to a coordinate $\{x_i\}_{i=1}^n$ of M near x. (Note this property is *independent* of choice of coordinates).

3.2. $C^{k,\alpha}$ norms for $\mathbb{R}^q$-valued functions. We first define the $C^{k,\alpha}$ norms for $\mathbb{R}^q$*-valued functions* for any integer $k \ge 0$ and $\alpha \in (0,1)$. Since our $q = q(t) \ge Ct^{-\frac{n}{2}-\rho} \to \infty$ as $t \to 0_+$, several equivalent $C^{k,\alpha}$ norms for any fixed q will diverge from each other as $q \to \infty$. To get the uniform quadratic estimate for *all* q, we will carefully choose the definition of the $C^{k,\alpha}$ norm.

Definition 14. Let $f : M \to \mathbb{R}^q$ (or ℓ^2, if $q = \infty$) be a $\mathbb{R}^q$-valued function $f = (f_1, \cdots, f_q)$, where each $f_j : M \to \mathbb{R}$. We let $|\cdot|$ be the standard Euclidean norm in $\mathbb{R}^q$, ∇ be the covariant derivative of (M, g), $\beta \ge 0$ be an integer, and let

$$\begin{aligned} \left\|\nabla^\beta f\right\|_{C^0(M,\mathbb{R}^q)} &= \sup_{x\in M}\left(\Sigma_{j=1}^q \left|\nabla^\beta f_j(x)\right|^2\right)^{1/2}, \\ \|f\|_{C^k(M,\mathbb{R}^q)} &= \Sigma_{0\le\beta\le k}\left\|\nabla^\beta f\right\|_{C^0(M,\mathbb{R}^q)}, \\ [f]_{\alpha,M;\mathbb{R}^q} &= \sup_{x\neq y\in M}\frac{|f(x)-f(y)|}{\operatorname{dist}(x,y)^\alpha}, \\ \|f\|_{C^{k,\alpha}(M,\mathbb{R}^q)} &= \|f\|_{C^k(M,\mathbb{R}^q)} + \left[\nabla^k f\right]_{\alpha,M;\mathbb{R}^q}. \end{aligned} \tag{3.2}$$

Then we have the following.

Lemma 15. *Let $\cdot$ be the standard inner product in $\mathbb{R}^q$ (including ℓ^2). For f and g in $C^{k,\alpha}(M,\mathbb{R}^q)$, we have*

$$\|f\cdot g\|_{C^{k,\alpha}(M)} \le C(k,\alpha,M)\,\|f\|_{C^{k,\alpha}(M,\mathbb{R}^q)}\,\|g\|_{C^{k,\alpha}(M,\mathbb{R}^q)}, \tag{3.3}$$

where the constant $C(k,\alpha,M) = n^k$ is uniform for all q.

Proof. The inequality already appeared in [**G1**]. We adapt it to $\mathbb{R}^q$-valued functions and further observe that the constant $C(k,\alpha,M)$ is *uniform* for any dimensional $\mathbb{R}^q$. This is because $D^{\overrightarrow{\gamma}}(f\cdot g)$ with $|\overrightarrow{\gamma}| \le k$ produces at most n^k inner product terms, and the *Cauchy–Schwartz inequality* $|a\cdot b| \le |a|_{\mathbb{R}^q}\,|b|_{\mathbb{R}^q}$ is valid for all $\mathbb{R}^q$ with coefficient 1 on the right-hand side. q.e.d.

3.3. Günther's implicit function theorem. Given a free mapping $u : M \to \mathbb{R}^q$ (or ℓ^2, if $q = \infty$) and $(h, f) \in C^{s,\alpha}(M, T^*M \oplus \mathrm{Sym}^{\otimes 2}(T^*M))$ with $s \geq 2$ and $0 < \alpha < 1$, let $v(x) \in C^{s,\alpha}(M, \mathbb{R}^q)$ be the unique solution to the system

$$P(u) \cdot v := \begin{bmatrix} \nabla u \\ \nabla^2 u \end{bmatrix} v = \begin{bmatrix} h \\ f \end{bmatrix}, \text{ and } v(x) \perp \ker P(u)(x), \tag{3.4}$$

where

$$\begin{aligned} \nabla u &= (\nabla u_1, \cdots, \nabla u_q) \in C^{s,\alpha}(M, T^*M \otimes \mathbb{R}^q), \\ \nabla^2 u &= \left(\nabla^2 u_1, \cdots, \nabla^2 u_q\right) \in C^{s,\alpha}\left(M, \mathrm{Sym}^{\otimes 2}(T^*M) \otimes \mathbb{R}^q\right), \\ P(u) &: C^{s,\alpha}(M, \mathbb{R}^q) \to C^{s,\alpha}\left(M, T^*M \oplus \mathrm{Sym}^{\otimes 2}(T^*M)\right), \end{aligned}$$

and ∇u_k and $\nabla^2 u_k$ are the gradient and Hessian of u_k, respectively, for $1 \leq k \leq q$. (Here we identify $T^*M \simeq TM$ by the metric g, so ∇u_k can be regarded in T^*M). If u is a free mapping, then (3.4) is always solvable for any (h, f). We define the right inverse of $P(u)$ as

$$E(u) : \begin{array}{ccc} C^{s,\alpha}\left(M, T^*M \oplus \mathrm{Sym}^{\otimes 2}(T^*M)\right) & \longrightarrow & C^{s,\alpha}(M, \mathbb{R}^q) \\ (h, f) & \longmapsto & v \end{array}. \tag{3.5}$$

By viewing $E(u)$ as a section of $C^{s,\alpha}(M, \mathrm{Hom}(T^*M \oplus \mathrm{Sym}^{\otimes 2}(T^*M), \mathbb{R}^q))$, the $C^{s,\alpha}(M)$-norm of $E(u)$ is induced from the Riemannian and Euclidean metrics g and g_{can}. By linear algebra, there is an explicit expression

$$E(u)(x) = P^T(u)\left[P(u)P^T(u)\right]^{-1}(x), \tag{3.6}$$

where "T" is the transpose. For an orthonormal frame field $\{V_i\}_{1 \leq i \leq n}$ near x on M, we have

$$P(u)(x) \simeq \begin{bmatrix} \{\nabla u(V_i)(x)\}^T \\ \left\{\nabla^2 u(V_j, V_k)(x)\right\}^T \end{bmatrix}_{1 \leq i,j,k \leq n, j \leq k}, \tag{3.7}$$

and $E(u)(x)$ is the unique *right inverse* of $P(u)(x)$ with its column vectors *orthogonal* to $\ker P(u)(x)$. The $C^{s,\alpha}(M)$-norm of $E(u)$ is the maximum of the $C^{s,\alpha}(M)$-norms of its column vectors, each viewed as an $\mathbb{R}^q$-valued function, defined in finitely many charts covering M by (3.2).

Theorem 16 (**[G2]**)**.** *Let $u : (M^n, g) \to \mathbb{R}^q$ be a C^∞*-free embedding *(cf. Definition* 13*). For $f \in C^{s,\alpha}\left(M, \mathrm{Sym}^{\otimes 2}(T^*M)\right)$ with $s \geq 2$ or $s = \infty$, and $0 < \alpha < 1$, there is a positive number θ (independent of u, s, and f) with the following property: If*

$$\|E(u)\|_{C^{2,\alpha}(M)} \|E(u)(0, f)\|_{C^{2,\alpha}(M)} \leq \theta, \tag{3.8}$$

then there exists a $v = v(u, f) \in C^{s,\alpha}(M, \mathbb{R}^q)$ solving $d(u+v) \cdot d(u+v) = du \cdot du + f$.

Remark 17. As will be seen in Section 5 and Section 9, Günther's implicit function theorem has a *uniform quadratic estimate* in $\mathbb{R}^q$ for all q, so the constant θ is independent on q, including $q=\infty$.

To obtain the $v \in C^{s,\alpha}(M,\mathbb{R}^q)$ for $2 \le s \le \infty$, Günther's theorem only requires to verify the $C^{2,\alpha}$-condition (3.8). This makes it easier to apply his theorem, but there appears no explicit control of the $C^{s,\alpha}$-norm of v, especially when $s>2$. For our purpose, it is useful to know how close the isometric embedding map I_t is to the "canonical" heat kernel embedding map Ψ_t in the $C^{s,\alpha}$-norm for any given $s \ge 2$, so we will prove a stronger inequality,

$$\|E(u)\|_{C^{s,\alpha}(M)} \|E(u)(0,f)\|_{C^{s,\alpha}(M)} \le \theta$$

in our case and give the estimate of $\|v\|_{C^{s,\alpha}(M)}$.

3.4. Günther's iteration scheme. The difficulty of applying the usual Banach space fixed point theorem to system (3.1) lies in the quadratic terms $\partial_i v \cdot \partial_j v$, which lose one order of differentiability after each iteration of v. Günther ([**G1**]) invented a new iteration scheme with no loss of differentiability, which we will recall in the following.

Let $\Delta_{(1)}$ and $\Delta_{(2)}$ be the *Lichnerowicz Laplacian* for vector fields and symmetric covariant 2-tensors on (M,g), respectively, i.e., in local coordinates

$$\begin{aligned} \Delta_{(1)} t_i \ &: \ = \Delta t_i - R^l_{i.} t_l, \ \Delta := \nabla^l \nabla_l, \\ \Delta_{(2)} t_{ij} \ &: \ = \Delta t_{ij} - R^{kl}_{i.j.} t_{kl} - R^l_{i.} t_{lj} - R^l_{j.} t_{il}. \end{aligned}$$

Fix a constant $\Lambda_0 \notin \operatorname{Spec}\left(\Delta_{(r)}\right)$, the spectrum of $\Delta_{(r)}$ on (M,g) for $r=1,2$. We introduce the smoothing operator $\left(\Delta_{(r)} - \Lambda_0\right)^{-1}$,

$$\begin{aligned} &\left(\Delta_{(1)} - \Lambda_0\right)^{-1} : C^{s-2,\alpha}(M, T^*M) \to C^{s,\alpha}(M, T^*M), \\ &\left(\Delta_{(2)} - \Lambda_0\right)^{-1} : C^{s-2,\alpha}\left(M, \operatorname{Sym}^{\otimes 2}(T^*M)\right) \to C^{s,\alpha}\left(M, \operatorname{Sym}^{\otimes 2}(T^*M)\right), \end{aligned}$$

with the *operator norm* denoted by $\left\|\left(\Delta_{(r)} - \Lambda_0\right)^{-1}\right\|_{\text{op}}$ and let

$$\sigma(\Lambda_0, \alpha, M) := \max_{r=1,2} \left\|\left(\Delta_{(r)} - \Lambda_0\right)^{-1}\right\|_{\text{op}}. \tag{3.9}$$

Given a free mapping $u \in C^{s,\alpha}(M,\mathbb{R}^q)$ and $f \in C^{s,\alpha}\left(M, \operatorname{Sym}^{\otimes 2}(T^*M)\right)$, let the vector field $N(v)$ and symmetric 2-tensor fields $L(v)$ and $M(v)$ be

$$\begin{aligned}
N_i(v) &= -\Delta v\cdot\nabla_i v,\\
L_{ij}(v) &= 2\nabla^l\nabla_i v\cdot\nabla_l\nabla_j v-2\Delta v\cdot\nabla_i\nabla_j v\\
&\quad -2R^{kl}_{i.j.}\nabla_k v\cdot\nabla_l v-\Lambda_0\nabla_i v\cdot\nabla_j v,\\
M_{ij}(v) &= \frac{1}{2}L_{ij}(v)\\
&\quad+\left(\nabla_i R^l_{j.}+\nabla_j R^l_{i.}-\nabla^l R_{ij}\right)\left(\left(\Delta_{(1)}-\Lambda_0\right)^{-1}N(v)\right)_l
\end{aligned} \tag{3.10}$$

in local coordinates, where the Einstein summation convention is used. Günther defined the iteration $\Upsilon_u : C^{s,\alpha}(M,\mathbb{R}^q)\to C^{s,\alpha}(M,\mathbb{R}^q)$ by

$$\Upsilon_u(v) := E(u)\left(0,-\frac{1}{2}f\right)+Q(u)(v,v), \tag{3.11}$$

where

$$\begin{aligned}
Q(u)(v,v) :=& E(u)\left(\left(\Delta_{(1)}-\Lambda_0\right)^{-1}N(v),\left(\Delta_{(2)}-\Lambda_0\right)^{-1}M(v)\right)\\
&\in C^{s,\alpha}(M,\mathbb{R}^q).
\end{aligned} \tag{3.12}$$

For later reference, we denote the components

$$\begin{aligned}
Q_i(v) &:= \left(\left(\Delta_{(1)}-\Lambda_0\right)^{-1}N(v)\right)_i,\\
Q_{jk}(v) &:= \left(\left(\Delta_{(2)}-\Lambda_0\right)^{-1}M(v)\right)_{jk}
\end{aligned}$$

for $1\le i,j,k\le n$.

4. Uniform linear independence property of $P(\Psi_t)$

Recall that the (un-normalized) heat kernel embedding $\Phi_t : (M,g)\to \ell^2$ in [**BBG**] is

$$\Phi_t : x\in M\to\left(e^{-\frac{\lambda_1}{2}t}\phi_1(x),e^{-\frac{\lambda_2}{2}t}\phi_2(x),\cdots,e^{-\frac{\lambda_q}{2}t}\phi_q(x),\cdots\right)\in\ell^2.$$

For any $x\in M$ we take an orthonormal frame field $\{V_i\}_{1\le i\le n}$ in its neighborhood. Following our notation $P(u)$ for a smooth map $u : M\to\mathbb{R}^q$, we consider the following $\frac{n(n+3)}{2}\times\infty$ matrix $P(\Phi_t)$ (with $q=\infty$) consisting of the n first derivatives of Φ_t and the $n(n+1)/2$ second covariant derivatives (i.e., Hessian) of Φ_t with respect to $\{V_i\}_{1\le i\le n}$:

$$P(\Phi_t)=\begin{bmatrix}
\cdots & \cdots & \cdots & \cdots & \cdots\\
e^{-\frac{\lambda_1}{2}t}\nabla_i\phi_1 & e^{-\frac{\lambda_2}{2}t}\nabla_i\phi_2 & \cdots & e^{-\frac{\lambda_q}{2}t}\nabla_i\phi_q & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots\\
e^{-\frac{\lambda_1}{2}t}\nabla_j\nabla_k\phi_1 & e^{-\frac{\lambda_2}{2}t}\nabla_j\nabla_k\phi_2 & \cdots & e^{-\frac{\lambda_q}{2}t}\nabla_j\nabla_k\phi_q & \cdots\\
\cdots & \cdots & \cdots & \cdots & \cdots
\end{bmatrix}.$$

We will prove that as $t \to 0_+$, $P(\Phi_t)(x)$ has full rank for any x on M, so Φ_t is a free mapping. For this we will compute the inner products of the row vectors of $P(\Phi_t)$. The following *uniform freeness* result of Φ_t will be proved in this section.

Theorem 18 (Uniform linear independence)**.** *As $t \to 0_+$, we have the following asymptotic formulae:*

$$\frac{\langle \nabla_i \Phi_t, \nabla_j \Phi_t \rangle}{|\nabla_i \Phi_t| \, |\nabla_j \Phi_t|} = \delta_{ij} + O(t), \quad \frac{\langle \nabla_i \nabla_j \Phi_t, \nabla_k \Phi_t \rangle}{|\nabla_i \nabla_j \Phi_t| \, |\nabla_k \Phi_t|} = O(t),$$

and

$$\frac{\langle \nabla_i \nabla_j \Phi_t, \nabla_k \nabla_l \Phi_t \rangle}{|\nabla_i \nabla_j \Phi_t| \, |\nabla_k \nabla_l \Phi_t|} = O(t) + \begin{cases} 1, & \{i,j\} = \{k,l\} \text{ as sets,} \\ 1/3, & i = j \text{ and } k = l, \text{ but } i \neq k \\ 0, & \text{otherwise,} \end{cases},$$

where $\langle , \rangle$ is the standard inner product in ℓ^2. The above convergence is uniform for all x on M in the C^r-norm for any $r \geq 0$. Moreover, if we truncate $\Phi_t : M \to \mathbb{R}^q \subset \ell^2$ for $q = q(t) \geq t^{-\frac{n}{2}-\rho}$ with sufficiently small $t > 0$, the above results still hold.

4.1. The Minakshisundaram–Pleijel expansion. As in [**BBG**], we have the *Minakshisundaram–Pleijel expansion* of the heat kernel

$$H(t,x,y) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{r^2}{4t}} U(t,x,y), \tag{4.1}$$

where $r = r(x,y)$ is the distance function for points x and y on M and

$$U(t,x,y) = u_0(x,y) + t u_1(x,y) + \cdots + t^p u_p(x,y) + O\left(t^{p+1}\right), \tag{4.2}$$

in the C^r sense for any $r \geq 0$ (see [**BeGaM**, p. 213]).

It is known that $u_0(x,y) = [\theta(x,y)]^{-1/2}$ (for x and y close enough) and

$$\theta(x,y) = \frac{\text{volume density at } y \text{ read in the normal coordinate around } x}{r^{n-1}} \tag{4.3}$$

with $r = r(x,y)$ ([**BeGaM**, p. 208]), in particular

$$\theta(x,x) = 1 = u_0(x,x). \tag{4.4}$$

From this we immediately see that

$$\langle \Phi_t, \Phi_t \rangle (x) = H(t,x,x) = \frac{1}{(4\pi t)^{n/2}} (1 + O(t)). \tag{4.5}$$

We also have

$$\partial_i U|_{x=y} = \partial_i u_0(x,y)|_{x=y} + t \partial_i u_1(x,y)|_{x=y} + O\left(t^2\right) = O(t), \tag{4.6}$$

where $\partial_i u_0(x,y)|_{x=y} = 0$ in (4.6) is due to

$$\begin{aligned} u_0(x,0) &= [\theta(x(s),0)]^{-\frac{1}{2}} = \left[1 - \mathrm{Ric}(\dot{x}(s),\dot{x}(s))\frac{s^2}{3!} + O\left(|s|^3\right)\right]^{-\frac{1}{2}} \\ &= 1 + \frac{1}{12}\mathrm{Ric}(\dot{x}(s),\dot{x}(s))s^2 + O\left(|s|^3\right). \end{aligned}$$

We recall the following useful lemma.

Lemma 19. *Let $r = r(x,y)$ be the shortest distance between x and y on M. For x and y that are close enough to each other, $r : M \times M \to \mathbb{R}$ is smooth. Using the normal coordinates $\left\{x^i\right\}_{1\le i\le n}$ near x, we have*

$$\partial_i^x r^2(x,y)|_{x=y} = \partial_i^y r^2(x,y)|_{x=y} = 0, \tag{4.7}$$

$$\partial_i^x\partial_j^x r^2(x,y)|_{x=y} = -\partial_i^x\partial_j^y r^2(x,y)|_{x=y} = 2\delta_{ij}, \tag{4.8}$$

$$\partial_k^x\partial_i^x\partial_j^x r^2(x,y)|_{x=y} = \partial_k^y\partial_i^x\partial_j^x r^2(x,y)|_{x=y} = 0, \tag{4.9}$$

where the notation ∂_i^x (resp. ∂_i^y) means the derivative is taken with respect to the variable in the first (resp. second) component of $M \times M$.

The first two identities are proved in [**BBG**]. The third identity can be computed in the normal coordinates near x (see e.g., [**De**, Chapter 16, p. 282]), using the Taylor expansion of the metric g near x (e.g., [**T**, Proposition 3.1, p. 41])

$$\begin{aligned} g_{ij}(x) &= \delta_{ij} + \frac{1}{3}R_{iklj}x^k x^l + \frac{1}{6}R_{iklj,s}x^k x^l x^s \\ &\quad + \left(\frac{1}{20}R_{iklj,st} + \frac{2}{45}\Sigma_m R_{iklm}R_{jstm}\right)x^k x^l x^s x^t + O\left(r^5\right), \end{aligned}$$

where r is the distance to the base point x_0.

4.2. Derivative estimates of the heat kernel embedding Ψ_t. Using Lemma 19 and the Minakshisundaram–Pleijel expansion (4.1), we will derive-higher derivative estimates of $D^{\overrightarrow{\alpha}}\Phi_t(x)$. Let $\overrightarrow{\gamma} = \left(\overrightarrow{\alpha},\overrightarrow{\beta}\right)$ be a multi-index, with $\overrightarrow{\alpha}$ and $\overrightarrow{\beta}$ being the multi-indices in x and y variables of $M \times M$, respectively. Let $D^{\overrightarrow{\gamma}}$ be the corresponding multi-derivative operator. From the heat kernel expression $H(t,x,y) = \Sigma_{s=1}^{\infty} e^{-\lambda_s t}\phi_s(x)\phi_s(y)$, it is easy to check that

$$\left\langle D^{\overrightarrow{\alpha}}\Phi_t(x), D^{\overrightarrow{\beta}}\Phi_t(x)\right\rangle = D^{\overrightarrow{\gamma}}H(t,x,y)|_{x=y}. \tag{4.10}$$

Proposition 20. As $t \to 0_+$, there exists a constant $C > 0$ such that

$$\begin{aligned} \left|D^{\overrightarrow{\gamma}}H(t,x,y)|_{x=y}\right| &\le Ct^{-\frac{n}{2}-\left[\frac{|\overrightarrow{\gamma}|}{2}\right]}, \\ \left|D^{\overrightarrow{\alpha}}\Phi_t(x)\right|^2 &\le Ct^{-\frac{n}{2}-\left|\overrightarrow{\alpha}\right|}, \end{aligned} \tag{4.11}$$

where $[b]$ is the largest integer less than or equal to a given real number b.

Proof. We write $H(t,x,y) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{r^2}{4t}} U(t,x,y)$, and

$$D^{\overrightarrow{\gamma}} H(t,x,y) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{r^2}{4t}} P_{\overrightarrow{\gamma}}(t,x,y), \tag{4.12}$$

where $P_{\overrightarrow{\gamma}}(t,x,y)$ is a *polynomial* in $D^{\overrightarrow{\mu_j}}\left(\frac{r^2(x,y)}{t}\right)$ and $D^{\overrightarrow{\eta_k}} U(t,x,y)$ for multi-indices $\overrightarrow{\mu_j}$ and $\overrightarrow{\eta_k}$ by the Leibniz rule. For example, when $\overrightarrow{\gamma} = \partial_{x_i}$,

$$P_{\overrightarrow{i}}(t,x,y) = -\frac{1}{4}\partial_i \left(\frac{r^2(x,y)}{t}\right) U(t,x,y) + \partial_i U(t,x,y).$$

We have the following

1. Each summand of $P_{\overrightarrow{\gamma}}(x,t)$ is of the form

$$\left(\Pi_{j=1}^{s} D^{\overrightarrow{\mu_j}}\left(\frac{r^2(x,y)}{t}\right)\right)\Big|_{x=y} \cdot D^{\overrightarrow{\eta}} U(t,x,y)\Big|_{x=y} \tag{4.13}$$

with

$$\Sigma_{j=1}^{s} |\overrightarrow{\mu_j}| + |\overrightarrow{\eta}| = |\overrightarrow{\gamma}|. \tag{4.14}$$

There are only finitely many terms in $P_{\overrightarrow{\gamma}}(x,t)$; we denote the total number by $n(\overrightarrow{\gamma})$.

2. As $t \to 0_+$, the terms involving the highest possible power of $\frac{1}{t}$ must have $\overrightarrow{\mu_j}$'s with $|\overrightarrow{\mu_j}| = 2$ as many as possible. This is because of the total degree condition (4.14) and Lemma 19. So if there are more than one $\overrightarrow{\mu_j}$ with $|\overrightarrow{\mu_j}| \geq 3$, the summand (4.13) loses the potential to have the maximal number of factors $\frac{1}{t}$, which is $\left[\frac{|\overrightarrow{\gamma}|}{2}\right]$ by (4.14).

Therefore, from (4.12) we have

$$\begin{aligned}
\left|D^{\overrightarrow{\gamma}} H(t,x,y)|_{x=y}\right| &\leq C\frac{1}{(4\pi t)^{n/2}} \cdot t^{-\left[\frac{|\overrightarrow{\gamma}|}{2}\right]}, \\
\left|D^{\overrightarrow{\alpha}} \Psi_t(x)\right|^2 &= \left|D_x^{\overrightarrow{\alpha}} D_y^{\overrightarrow{\alpha}} H(t,x,y)|_{x=y}\right| \leq C\frac{1}{(4\pi t)^{n/2}} \cdot t^{-\left[\frac{2|\overrightarrow{\alpha}|}{2}\right]} \\
&= C\frac{1}{(4\pi t)^{n/2}} \cdot t^{-|\overrightarrow{\alpha}|}.
\end{aligned}$$

The constant C can be taken as $C = n(\overrightarrow{\alpha}) \sup \|U(t,x,y)\|_{C^{|\overrightarrow{\alpha}|}(M\times M)}$, where the sup is taken for $t > 0$ in the range of the Minakshisundaram–Pleijel expansion (4.1) and $(x,y) \in M \times M$ with $\mathrm{dist}(x,y)$ less than the injective radius of M. q.e.d.

We can get precise asymptotic formulae if we are more careful in item 2 of the above argument, as in the following.

Proposition 21. For any x on M, let $\{x^i\}_{i=1}^n$ be the normal coordinates near x. Then as $t \to 0_+$, we have

$$\begin{aligned}\langle \partial_i \Phi_t, \partial_j \Phi_t \rangle (x) &= (2t)^{-1} (4\pi t)^{-\frac{n}{2}} (\delta_{ij} + O(t)), \\ \langle \partial_i \partial_j \Phi_t, \partial_k \Phi_t \rangle (x) &= t^{-\frac{n}{2}} \cdot O(1),\end{aligned}$$

$$\begin{aligned}&\langle \partial_j \partial_i \Phi_t, \partial_m \partial_k \Phi_t \rangle (x) \\ &= (2t)^{-2} (4\pi t)^{-\frac{n}{2}} (\delta_{ij}\delta_{km} + \delta_{im}\delta_{jk} + \delta_{ik}\delta_{jm} + O(t))\end{aligned} \tag{4.15}$$

for $1 \le i, j, k, m \le n$. The convergence is uniform for all x on M in the C^r-norm for any $r \ge 0$. The coefficients of $O(t)$ and $O(1)$ are determined by the covariant differentiations of the metric g and its curvature tensors.

Proof. The first asymptotic formula is in [**BBG**]. We will prove the remaining two. For simplicity, we let $\partial_i^x = \partial_i$ and $\partial_k^y = \partial_{\bar{k}}$ from now on (the bar notation in "$\overline{k}$" means that the derivative is taken with respect to the y variable for $(x, y) \in M \times M$). We will compute $D^{\overrightarrow{\gamma}} H(t, x, y)|_{x=y}$ for multi-indices $\overrightarrow{\gamma}$ in the following two cases and observe the summands (4.13) in $P_{\overrightarrow{\gamma}}(t, x, y)$ (4.12) of leading order:

1) $\overrightarrow{\gamma} = (\partial_j \partial_i, \partial_{\bar{k}})$: By (4.14), $\Sigma_{j=1}^s |\overrightarrow{\mu_j}| + |\overrightarrow{\eta}| = 3$. So the highest possible power of $\left(\frac{1}{t}\right)$ in $P_{\overrightarrow{\gamma}}(t, x, y)$ is $\left[\frac{3}{2}\right] = 1$, with $s = 1$, $|\overrightarrow{\mu_1}| = 2$, $|\overrightarrow{\eta}| = 1$. The term in $P_{\overrightarrow{\gamma}}(t, x, y)$ containing $\frac{1}{t}$ is

$$\begin{aligned}&\partial_{\overline{k}} \partial_j \left(-\frac{r^2}{4t}\right) \partial_i U + \partial_j \partial_i \left(-\frac{r^2}{4t}\right) \partial_{\overline{k}} U + \partial_{\overline{k}} \partial_i \left(-\frac{r^2}{4t}\right) \partial_j U \Big|_{x=y} \\ = \quad & -\frac{1}{2t} \left(-\delta_{kj} \partial_i U + \delta_{ij} \partial_{\overline{k}} U - \delta_{ik} \partial_j U\right) = O(1),\end{aligned}$$

we have used (4.6). Since we have used $\nabla U|_{x=y} = O(t)$ to decrease the $\left(\frac{1}{t}\right)$ power by 1, we must also consider the term in $P_{\overrightarrow{\gamma}}(t, x, y)$ with $s = 0$ and $|\overrightarrow{\eta}| = 3$, but clearly $\partial_j \partial_i \partial_{\bar{k}} U|_{x=y} = O(1)$. Hence, $|\langle \partial_i \partial_j \Phi_t, \partial_k \Phi_t \rangle (x)| = \partial_{\bar{k}} \partial_j \partial_i H(t, x, y)|_{x=y} = t^{-n/2} \cdot O(1)$.

2) $\overrightarrow{\gamma} = (\partial_j \partial_i, \partial_{\bar{m}} \partial_{\bar{k}})$: By (4.14), $\Sigma_{j=1}^s |\overrightarrow{\mu_j}| + |\overrightarrow{\eta}| = 4$. So the highest possible power of $\left(\frac{1}{t}\right)$ in $P_{\overrightarrow{\gamma}}(t, x, y)$ is $\left[\frac{4}{2}\right] = 2$, with $s = 2$, $|\overrightarrow{\mu_1}| = |\overrightarrow{\mu_2}| = 2$, $|\overrightarrow{\eta}| = 0$. The term in $P_{\overrightarrow{\gamma}}(t, x, y)$ containing $\left(\frac{1}{t}\right)^2$ is

$$\begin{aligned}&\left(-\frac{1}{4t}\right)^2 \left[\partial_{\bar{m}} \partial_{\bar{k}} \left(r^2\right) \partial_j \partial_i \left(r^2\right) + \partial_{\overline{k}} \partial_j \left(r^2\right) \partial_{\bar{m}} \partial_i \left(r^2\right)\right. \\ &\left.+ \partial_{\bar{m}} \partial_j \left(r^2\right) \partial_{\overline{k}} \partial_i \left(r^2\right)\right] U\Big|_{x=y} \\ = \quad & \left(\frac{1}{2t}\right)^2 (\delta_{ij}\delta_{km} + \delta_{im}\delta_{jk} + \delta_{ik}\delta_{jm} + O(t)).\end{aligned}$$

Hence,

$$\begin{aligned}\langle \partial_j\partial_i\Phi_t, \partial_m\partial_k\Phi_t\rangle (x) &= \partial_{\bar{m}}\partial_{\bar{k}}\partial_j\partial_i H(t,x,y)\,|_{x=y} \\ = \quad & \left(\frac{1}{2t}\right)^2 \frac{1}{(4\pi t)^{n/2}} \left(\delta_{ij}\delta_{km} + \delta_{im}\delta_{jk} + \delta_{ik}\delta_{jm} + O(t)\right).\end{aligned}$$

The coefficients of $O(t)$ and $O(1)$ in the above argument are determined by the covariant differentiations of the metric g and its curvature tensors, because of (4.10) and the heat kernel expansion of $D^{\overrightarrow{\gamma}}H$ $(t,x,y)\,|_{x=y}$, similar to the expansion of $H(t,x,x)$ (e.g., [**Gil**]). q.e.d.

Remark 22. In the expansion of $\partial_{\bar{m}}\partial_{\bar{k}}\partial_j\partial_i H(t,x,y)$, the term $\frac{\partial_{\bar{m}}\partial_{\bar{k}}\partial_j\partial_i(r^2)}{4t}U$ involves the *curvature tensors* on M (see, e.g., [**De**, Chapter 16, p. 282],), but it is a lower-order term (order t^{-1} v.s. leading order t^{-2}) and so does not affect the asymptotic behavior.

Remark 23. Propositions 20 and 21 with precise coefficients C (such that the inequalities become equalities as $t \to 0_+$) were also obtained in [**Ni**] in the context of random function theory by different argument. For the purpose of our paper we do not need that general result, prefering to give a self-contained derivation. [**Ni**] also proved the almost isometric embeddings by eigenfunctions and a wide class of weights. Potentially, some of them may be perturbed to isometric embeddings by our method.

For later applications, we need to estimate the Hölder derivatives of $D^{\overrightarrow{\alpha}}\Psi_t(x)$. This can be done by interpolating between estimates of the integral exponent obtained in Proposition 20. Then we have the following.

Proposition 24. As $t \to 0_+$, the Hölder derivatives satisfy

$$\begin{aligned}\left[D^{\overrightarrow{\alpha}}\Phi_t(x)\right]_{\alpha;M} &\leq Ct^{-\frac{n}{4}-\frac{|\overrightarrow{\alpha}|+\alpha}{2}}, \\ \|\Phi_t(x)\|_{C^{k,\alpha}(M)} &\leq Ct^{-\frac{n}{4}-\frac{k+\alpha}{2}}\end{aligned}$$

for some constant $C > 0$.

4.3. Uniform linear independence property of $P(\Psi_t)$.

Proof of Theorem 18. For any $x \in M$, we choose the normal coordinates $\{x^i\}_{1\leq i\leq n}$ near x such that $\left\{\frac{\partial}{\partial x^i}\right\}_{1\leq i\leq n}$ agree with the frame field $\{V_i\}_{1\leq i\leq n}$ at x. Then $\nabla_i\Phi_t(x) = \partial_i\Phi_t(x)$, $\nabla_j\nabla_k\Phi_t(x) = \partial_j\partial_k\Phi_t(x)$.

From Proposition 21 we have as $t \to 0_+$,

$$\begin{aligned} |\nabla_i \Phi_t(x)|^2 &= (2t)^{-1}(4\pi t)^{-n/2} \cdot (1 + O(t)), \\ \frac{\langle \nabla_i \Phi_t, \nabla_j \Phi_t \rangle}{|\nabla_i \Phi_t| \, |\nabla_j \Phi_t|}(x) &= \delta_{ij} + O(t). \end{aligned} \tag{4.16}$$

From (4.15) we have

$$\begin{aligned} \langle \nabla_j \nabla_i \Phi_t, \nabla_m \nabla_k \Phi_t \rangle (x) = (2t)^{-2}(4\pi t)^{-n/2} \cdot (&\delta_{ij}\delta_{km} \\ &+ \delta_{im}\delta_{jk} + \delta_{ik}\delta_{jm} + O(t)). \end{aligned}$$

In particular, for $i \neq j$ and $\{k, m\} \neq \{i, j\}$ as sets, we have

$$\begin{aligned} \langle \nabla_i \nabla_i \Phi_t, \nabla_i \nabla_i \Phi_t \rangle (x) &= (2t)^{-2}(4\pi t)^{-n/2} \cdot (3 + O(t)), \\ \langle \nabla_j \nabla_i \Phi_t, \nabla_j \nabla_i \Phi_t \rangle (x) &= (2t)^{-2}(4\pi t)^{-n/2} \cdot (1 + O(t)), \\ \langle \nabla_i \nabla_i \Phi_t, \nabla_j \nabla_j \Phi_t \rangle (x) &= (2t)^{-2}(4\pi t)^{-n/2} \cdot (1 + O(t)), \\ \langle \nabla_j \nabla_i \Phi_t, \nabla_m \nabla_k \Phi_t \rangle (x) &= (2t)^{-2}(4\pi t)^{-n/2} \cdot (0 + O(t)), \end{aligned} \tag{4.17}$$

So we conclude that

$$\begin{aligned} &\frac{\langle \nabla_i \nabla_j \Phi_t, \nabla_k \nabla_l \Phi_t \rangle}{|\nabla_i \nabla_j \Phi_t| \, |\nabla_k \nabla_l \Phi_t|}(x) \\ &= \begin{cases} 0 + O(t), & \text{if } \{i,j\} \neq \{k,l\} \text{ and } \{i,k\} \neq \{j,l\} \text{ as sets,} \\ 1/3 + O(t), & \text{if } i = j \text{ and } k = l \text{ , but } i \neq k, \\ 1 + O(t), & \text{if } \{i,j\} = \{k,l\} \text{ as sets.} \end{cases} \end{aligned} \tag{4.18}$$

By (4.17) $|\nabla_j \nabla_i \Phi_t|^2 \to C t^{-\frac{n}{2}-2}$, combining Proposition 21 we have

$$\begin{aligned} &\frac{|\langle \nabla_i \nabla_j \Phi_t, \nabla_k \Phi_t \rangle|}{|\nabla_i \nabla_j \Phi_t| \, |\nabla_k \Phi_t|}(x) \\ &= \frac{t^{-n/2} \cdot O(1)}{\left[(2t)^{-2}(4\pi t)^{-n/2} \cdot (1 + O(t)) \cdot (2t)^{-1}(4\pi t)^{-n/2} \cdot (1 + O(t))\right]^{1/2}} \\ &= O\left(t^{3/2}\right). \end{aligned} \tag{4.19}$$

By Proposition 9, the above inner product results pass to the truncated map $\Phi_t : (M, g) \to \mathbb{R}^{q(t)}$ as well.

The linear independence of the row vectors follows from (4.16), (4.18), (4.19), and taking

$$\alpha_i = \nabla_i \nabla_i \Phi_t(x) \text{ for } i = 1, \cdots, n$$

in the following linear algebra lemma. q.e.d.

Lemma 25. *For n vectors $\alpha_1, \cdots, \alpha_n$ in a real linear space V equipped with an inner product $\langle , \rangle$, if there is a constant $\sigma \in \left(-\frac{1}{n-1}, 1\right)$ such that*

$$\frac{\langle \alpha_i, \alpha_j \rangle}{|\alpha_i|\,|\alpha_j|} = \sigma, \text{ for all } i \neq j,$$

then $\{\alpha_i\}_{1 \leq i \leq n}$ are linearly independent.

Proof. We will find an explicit invertible linear transform $L_0 : V \to V$ to change the angle $\cos^{-1} \sigma$ between α_i and α_j to $\frac{\pi}{2}$. Without loss of generality, we can assume all α_i are unit vectors. Let

$$\alpha_0 = \frac{\alpha_1 + \cdots + \alpha_n}{n}, \; c = c(\sigma, n) = \sqrt{\frac{1 + (n-1)\sigma}{1 - \sigma}} \; (\text{note } \sigma > -\frac{1}{n-1}),$$
$$\widetilde{\alpha_i} = \alpha_0 + c(\alpha_i - \alpha_0) := L_0 \alpha_i, \text{ for } i = 1, \cdots, n;$$

then $\widetilde{\alpha_i}$ is nonzero by checking the orthogonal relation $\langle \alpha_0, \alpha_i - \alpha_0 \rangle = 0$. We also have for $i \neq j$,

$$\begin{aligned} \langle \widetilde{\alpha_i}, \widetilde{\alpha_j} \rangle &= |\alpha_0|^2 + c^2 (\alpha_i - \alpha_0)(\alpha_j - \alpha_0) \\ &= \frac{(n-1)\sigma + 1}{n} + \frac{1 + (n-1)\sigma}{1 - \sigma} \left(\sigma - \frac{(n-1)\sigma + 1}{n} \right) = 0, \end{aligned}$$

and so $\{\widetilde{\alpha_i}\}_{1 \leq i \leq n}$ is an orthogonal set. Since $\{\widetilde{\alpha_i}\}_{1 \leq i \leq n}$ is obtained from linear combinations of $\{\alpha_i\}_{1 \leq i \leq n}$, $\{\alpha_i\}_{1 \leq i \leq n}$ must be linearly independent. q.e.d.

Corollary 26. *Let $\sigma \in \left(-\frac{1}{n-1}, 1\right)$. Then the $n \times n$ matrix*

$$\Xi_n(\sigma) := [\theta_{ij}]_{1 \leq i, j \leq n} \tag{4.20}$$

with $\theta_{ii} = 1$ and $\theta_{ij} = \sigma$ $(i \neq j)$ is invertible.

Proof. Let P be the matrix whose row vectors are the above unit vectors α_i $(1 \leq i \leq n)$. Then P is a matrix of full rank, and $PP^T = \Xi_n(\sigma)$. q.e.d.

4.4. Operator norm estimate of $E(\Psi_t)$. We start with the following elementary linear algebra lemmas.

Lemma 27. *Let A be an $m \times n$ matrix. Regarding A as a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, the operator norm $\|A\|$ of A, defined as*

$$\|A\| = \sup_{v \in \mathbb{R}^n, \; |v| = 1} \frac{|Av|}{|v|},$$

is less than or equal to $\sqrt{n}$ times the length of its longest column vector. If the column vectors are orthogonal to each other, then $\|A\|$ is equal to the length of the longest column vector.

Lemma 28. *Let $A_i(t)$ be $m_i \times m_i$ symmetric, invertible matrices with operator norm $\left\|(A_i(t))^{-1}\right\| \le \rho_0$ for $i = 1, 2$ and $t \in (0, t_0]$, and let $b(t)$ be an $m_2 \times m_1$ matrix with $\|b(t)\| \to 0$ as $t \to 0_+$. Then for sufficiently small $t > 0$, the inverse matrix for $\begin{bmatrix} A_1(t) & b^T(t) \\ b(t) & A_2(t) \end{bmatrix}$ is*

$$\begin{bmatrix} A_1^{-1}(t) & c^T(t) \\ c(t) & A_2^{-1}(t) \end{bmatrix} \begin{bmatrix} \left(I_{m_1} + b^T(t)\,c(t)\right)^{-1} & 0 \\ 0 & \left(I_{m_2} + b(t)\,c^T(t)\right)^{-1} \end{bmatrix},$$

where $c(t)$ is the $m_2 \times m_1$ matrix given by $c(t) = A_2^{-1}(t)\,b(t)\,A_1^{-1}(t)$. In particular,

$$\|c(t)\| \le \left\|(A_2(t))^{-1}\right\| \|b(t)\| \left\|(A_1(t))^{-1}\right\|.$$

From now on we consider the *normalized heat kernel embedding* $\Psi_t = \sqrt{2}\,(4\pi)^{n/4}\, t^{\frac{n+2}{4}}\,\Phi_t$. Theorem 18 still holds if we replace Φ_t by Ψ_t, for they only differ by a scaling factor $\sqrt{2}\,(4\pi)^{n/4}\, t^{\frac{n+2}{4}}$.

Corollary 29. *The matrix $P(\Psi_t)(x)$ has a right inverse $E(\Psi_t)(x)$ with uniform operator norm bound C for all $q \ge t^{-\frac{n}{2}-\rho}$ and all $x \in M$ as $t \to 0_+$.*

Proof. Since $\Psi_t = \sqrt{2}\,(4\pi)^{n/4}\, t^{\frac{n+2}{4}}\,\Phi_t$, by Proposition 21, as $t \to 0_+$ we have in the C^r-sense (for any $r \ge 0$)

$$\begin{aligned} & P(\Psi_t)\,P^T(\Psi_t)(x) \\ & = \begin{bmatrix} I_n + O(t) & O(t) \\ O(t) & \frac{1}{2t} \cdot \left(\begin{bmatrix} I_{\frac{n(n-1)}{2}} & 0 \\ 0 & \Xi_n\left(\frac{1}{3}\right) \end{bmatrix} + O(t) \right) \end{bmatrix}, \end{aligned} \tag{4.21}$$

where I_n corresponds to $\langle \nabla_i \Psi_t, \nabla_j \Psi_t \rangle$, $I_{\frac{n(n-1)}{2}}$ corresponds to $\langle \nabla_i \nabla_j \Psi_t, \nabla_k \nabla_l \Psi_t \rangle$ for $i \ne j$ and $k \ne l$, and $\Xi_n\left(\frac{1}{3}\right)$ (defined in Corollary 26) corresponds to $\langle \nabla_i \nabla_i \Psi_t, \nabla_k \nabla_k \Psi_t \rangle$ for $1 \le i, j, k, l \le n$. By Proposition 9, (4.21) still holds when we truncate Ψ_t from ℓ^2 to $\mathbb{R}^{q(t)}$ with $q(t) \ge t^{-\frac{n}{2}-\rho}$. Therefore as $t \to 0_+$, by Lemma 28 we have

$$\begin{aligned} & \left[P(\Psi_t)\,P^T(\Psi_t)\right]^{-1}(x) \\ & = \begin{bmatrix} I_n + O(t) & O(t) \\ O(t) & 2t \cdot \left(\begin{bmatrix} I_{\frac{n(n-1)}{2}} & 0 \\ 0 & \left(\Xi_n\left(\frac{1}{3}\right)\right)^{-1} \end{bmatrix} + O(t) \right) \end{bmatrix} \end{aligned} \tag{4.22}$$

in the C^r-sense. By Lemma 27, for the right inverse

$$E(\Psi_t) = P^T(\Psi_t)\left[P(\Psi_t)\,P^T(\Psi_t)\right]^{-1}, \tag{4.23}$$

its operator norm is controlled by the length of its longest column vector. From Proposition 21, $|\nabla_i \Psi_t| \to 1$ and $|\nabla_j \nabla_k \Psi_t| \to \frac{1}{\sqrt{2t}}$ in the C^r-sense as $t \to 0_+$, so plugging this and (4.22) into (4.23), we then have

$$\|E(\Psi_t)(x)\| \leq C \sup_{1\leq i,j,k\leq n} \left(\|\nabla_i \Psi_t(x)\| + t\,\|\nabla_j \nabla_k \Psi_t(x)\|\right) \leq C.$$

q.e.d.

Proposition 30. For any multi-index $\overrightarrow{\alpha}$ with $|\overrightarrow{\alpha}| = k$ and $0 < \alpha < 1$, for any $x \in M$, the operator norms of the linear maps $D^{\overrightarrow{\alpha}} E(\Psi_t)(x)$ and $\left[D^{\overrightarrow{\alpha}} E(\Psi_t)\right]_{\alpha,M}(x) : \mathbb{R}^{\frac{n(n+3)}{2}} \to \mathbb{R}^{q(t)}$ satisfy

$$\begin{aligned} \left\|D^{\overrightarrow{\alpha}} E(\Psi_t)(x)\right\| &\leq Ct^{-\frac{k}{2}}, \\ \left\|\left[D^{\overrightarrow{\alpha}} E(\Psi_t)\right]_{\alpha,M}(x)\right\| &\leq Ct^{-\frac{k+\alpha}{2}}, \end{aligned} \tag{4.24}$$

respectively, for all $q(t) \geq t^{-\frac{n}{2}-\rho}$ as $t \to 0_+$.

Proof. In each chart U of M where we use the orthonormal frame field $\{V_i\}_{1\leq i\leq n}$ to trivialize $P(\Psi_t)$, the $O(t)$'s in (4.22) are smooth functions on U with the C^r-norm of order $O(t)$. So for any multi-index $\overrightarrow{\gamma}$ and real number $\gamma \in [0,1)$ with $|\overrightarrow{\gamma}| + \gamma > 0$, by (4.22) we have

$$\left[D^{\overrightarrow{\gamma}} \left[P(\Psi_t) P^T(\Psi_t)\right]^{-1}\right]_{\gamma}(x) = O(t) \tag{4.25}$$

in the C^r-sense for any $r \geq 0$. Therefore, for any multi-index $\overrightarrow{\alpha}$ with $|\overrightarrow{\alpha}| = k$, applying $D^{\overrightarrow{\alpha}}$ to (4.23), using the Leibniz rule, and noticing Lemma 27 and (4.25), we have

$$\begin{aligned} &\left\|D^{\overrightarrow{\alpha}} E(\Psi_t)\ \ (x)\right\| \\ \leq\ & C\Sigma_{\overrightarrow{\beta}\cup\overrightarrow{\gamma}=\overrightarrow{\alpha}} \left\|D^{\overrightarrow{\beta}} P^T(\Psi_t)(x) \cdot D^{\overrightarrow{\gamma}} \left[P(\Psi_t) P^T(\Psi_t)\right]^{-1}(x)\right\| \\ \leq\ & C\left(\left|\nabla^{k+1}\Psi_t(x)\right| + \left|\nabla^{k+2}\Psi_t(x)\right| \cdot O(t)\right) \leq Ct^{-\frac{k}{2}}, \\ &\left\|\left[D^{\overrightarrow{\alpha}} E(\Psi_t)\right]_{\alpha,M}(x)\right\| \\ \leq\ & C\left(\left\|\left[D^{\overrightarrow{\alpha}} P^T(\Psi_t)(x)\right]_{\alpha} \cdot \left[P(\Psi_t) P^T(\Psi_t)\right]^{-1}(x)\right\|\right. \\ &\left. + \left\|D^{\overrightarrow{\alpha}} P^T(\Psi_t)(x) \cdot \left[\left[P(\Psi_t) P^T(\Psi_t)\right]^{-1}(x)\right]_{\alpha}\right\|\right) \\ \leq\ & C\left(t^{-\frac{k}{2}-\frac{\alpha}{2}} + t^{-\frac{k}{2}} \cdot O(t)\right) \leq Ct^{-\frac{k+\alpha}{2}}, \end{aligned}$$

where in both inequalities we have used Proposition 20 and Proposition 24 for the derivative estimates of $\Psi_t(x)$. q.e.d.

Corollary 31. *For* $q \geq Ct^{-\frac{n}{2}-\rho}$, $\|E(\Psi_t)\|_{C^{k,\alpha}(M)}$ *is of order* $t^{-\frac{k+\alpha}{2}}$, *and so is the* operator norm $\|E(\Psi_t)\|_{op}$ *of*

$$E(\Psi_t) : C^{k,\alpha}(M, T^*M) \times C^{k,\alpha}\left(M, \mathrm{Sym}^{\otimes 2}(T^*M)\right) \to C^{k,\alpha}(M, \mathbb{R}^q),$$

i.e.,

$$\|E(\Psi_t)\|_{C^{k,\alpha}(M)}, \|E(\Psi_t)\|_{\mathrm{op}} \leq Ct^{-\frac{k+\alpha}{2}} \tag{4.26}$$

for a constant $C > 0$.

Proof. Taking the supremum for $x \in M$ in the inequalities in (4.24), we obtain

$$\|E(\Psi_t)\|_{C^{k,\alpha}(M)} \leq Ct^{-\frac{k+\alpha}{2}}. \tag{4.27}$$

Now we estimate the operator norm $\|E(\Psi_t)\|_{\mathrm{op}}$. For any section $\varphi \in C^{k,\alpha}(M, T^*M) \times C^{k,\alpha}\left(M, \mathrm{Sym}^{\otimes 2}(T^*M)\right)$, using Proposition 30, by the Leibniz rule we have

$$\left|\left[D^{\overrightarrow{\alpha}}\left(E(\Psi_t)(x)\varphi(x)\right)\right]_{\alpha,M}\right| \leq Ct^{-\frac{k+\alpha}{2}}\|\varphi\|_{C^{k,\alpha}(M)}$$

for any multi-index $\overrightarrow{\alpha}$ with $|\overrightarrow{\alpha}| = k$. (In a local trivialization of TM, $\varphi : M \to \mathbb{R}^{\frac{n(n+3)}{2}}$. The $C^{k,\alpha}$-norm for vector-valued functions is given in Section 3.2). Taking the supremum for $x \in M$ in the above inequalities, we have

$$\|E(u)\varphi\|_{C^{k,\alpha}(M,\mathbb{R}^q)} \leq Ct^{-\frac{k+\alpha}{2}}\|\varphi\|_{C^{k,\alpha}(M,\mathbb{R}^q)},$$

so the *operator norm* $\|E(\Psi_t)\|_{\mathrm{op}}$ is of order $Ct^{-\frac{k+\alpha}{2}}$. Note this operator norm agrees with the $C^{k,\alpha}$-Hölder norm of $E(u)$ by (4.27). q.e.d.

Definition 32 (The constant C_E)**.** Due to the importance of the operator norm of $E(\Psi_t)$, we denote the maximum of the constants C appeared in the coefficients of the above estimates of $\|E(\Psi_t)(x)\|$, $\|E(\Psi_t)\|$, $\|E(\Psi_t)\|_{C^{k,\alpha}(M,\mathbb{R}^q)}$, $\left\|D^{\overrightarrow{\alpha}}\Psi_t\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)}$ in Proposition 24, $O(t)$ in (4.22), and $2\left\|\left(\Xi\left(\frac{1}{3}\right)\right)^{-1}\right\|$ by C_E, where "E" indicates $E(\Psi_t)$.

5. Uniform quadratic estimate of $Q(u)$

For any given map $u \in C^{k,\alpha}(M, \mathbb{R}^q)$, the quadratic estimate of $Q(u)$ was established in [**G1**, Lemma 4]. In this section we show the constant in the quadratic estimate is *uniform* for all $\mathbb{R}^q$. This is essentially due to Lemma 15, where the constant $C(k,\alpha,M)$ is uniform for all q.

Proposition 33. For any $v \in C^{k,\alpha}(M,\mathbb{R}^q)$, we have

$$\begin{aligned}
\|Q_i(v,v)\|_{C^{k,\alpha}(M,\mathbb{R}^q)} &\le \Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)\|v\|^2_{C^{k,\alpha}(M,\mathbb{R}^q)}, \\
\|Q_{ij}(v,v)\|_{C^{k,\alpha}(M,\mathbb{R}^q)} &\le \Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)\|v\|^2_{C^{k,\alpha}(M,\mathbb{R}^q)}, \\
\|Q(\Psi_t)(v,v)\|_{C^{k,\alpha}(M,\mathbb{R}^q)} &\le C_E\Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)t^{-\frac{k+\alpha}{2}}\|v\|^2_{C^{k,\alpha}(M,\mathbb{R}^q)},
\end{aligned}$$

where the constant $\Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)$ is uniform for all q, where $\|R\|_{C^1}$ is the C^1-norm of the Riemannian curvature tensor R on M. (The constants $\sigma(\Lambda_0,\alpha,M)$, $C(k,\alpha,M)$, and C_E are in (3.9), Lemma 15 and Definition 32 respectively).

Proof. For brevity, we write $C^{k,\alpha}(M,\mathbb{R}^q)$ as $C^{k,\alpha}(M)$. Let "$\cdot$" be the standard inner product in $\mathbb{R}^q$. Recall that $Q_i(v,v) = \left(\left(\Delta_{(1)}-\Lambda_0\right)^{-1}(N(v))\right)_i$ and $Q_{ij}(v,v) = \left(\left(\Delta_{(2)}-\Lambda_0\right)^{-1}(M(v))\right)_{ij}$ in Section 3.4, where

$$\begin{aligned}
N_i(v) &= -\Delta v\cdot\nabla_i v, \\
L_{ij}(v) &= 2\nabla^l\nabla_i v\cdot\nabla_l\nabla_j v - 2\Delta v\cdot\nabla_i\nabla_j v - 2R^{kl}_{i.j.}\nabla_k v\cdot\nabla_l v \\
&\quad - \Lambda_0\nabla_i v\cdot\nabla_j v, \\
M_{ij}(v) &= \frac{1}{2}L_{ij}(v) + \left(\nabla_i R^l_{j.} + \nabla_j R^l_{i.} - \nabla^l R_{ij}\right)\left(\left(\Delta_{(1)}-\Lambda_0\right)^{-1}N(v)\right)_l.
\end{aligned}$$

By the definition of the operator norm $\left\|\left(\Delta_{(r)}-\Lambda_0\right)^{-1}\right\|_{\mathrm{op}}$ and (3.9), we have

$$\begin{aligned}
&\|Q_{ij}(v,v)\|_{C^{k,\alpha}(M)} \\
\le\ &\sigma(\Lambda_0,\alpha,M)\left(\frac{1}{2}\|L_{ij}(v)\|_{C^{k-2,\alpha}(M)}\right. \\
&\left.+\|\nabla R\|_{C^0(M)}\,\sigma(\Lambda_0,\alpha,M)\,\|N(v)\|_{C^{k-2,\alpha}(M)}\right).
\end{aligned} \tag{5.1}$$

For $L_{ij}(v)$, by (3.3) we have

$$\begin{aligned}
&\left\|L_{ij}(v)\right\|_{C^{k,\alpha}(M)}\\
\leq\ & 2C(k,\alpha,M)\Bigg(\left\|\nabla^l\nabla_i v\right\|_{C^{k-2,\alpha}(M)}\left\|\nabla_l\nabla_j v\right\|_{C^{k-2,\alpha}(M)}\\
&+\left\|\Delta v\right\|_{C^{k-2,\alpha}(M)}\left\|\nabla_i\nabla_j v\right\|_{C^{k-2,\alpha}(M)}\\
&+\left\|R\right\|_{C^0(M)}\left\|\nabla_k v\right\|_{C^{k-2,\alpha}(M)}\left\|\nabla_l v\right\|_{C^{k-2,\alpha}(M)}\\
&+\left|\frac{\Lambda_0}{2}\right|\left\|\nabla_i v\right\|_{C^{k-2,\alpha}(M)}\left\|\nabla_j v\right\|_{C^{k-2,\alpha}(M)}\Bigg)\\
\leq\ & 2C(k,\alpha,M)\Big(\|v\|^2_{C^{k,\alpha}(M)}+\|v\|^2_{C^{k,\alpha}(M)}+\\
&\|R\|_{C^0(M)}\|v\|^2_{C^{k-1,\alpha}(M)}+\left|\frac{\Lambda_0}{2}\right|\|\nabla_i v\|^2_{C^{k-1,\alpha}(M)}\Big)\\
\leq\ & 2C(k,\alpha,M)\left(2+\|R\|_{C^0(M)}+\left|\frac{\Lambda_0}{2}\right|\right)\|v\|^2_{C^{k,\alpha}(M)}.
\end{aligned}$$

Similarly,

$$\left\|N_i(v)\right\|_{C^{k-2,\alpha}(M)}\leq\|\Delta v\|_{C^{k-2,\alpha}(M)}\|\nabla_i v\|_{C^{k-2,\alpha}(M)}\leq\|v\|^2_{C^{k,\alpha}(M)}.$$

Putting these into (5.1), we have

$$\begin{aligned}
(5.2)\qquad &\left\|Q_{ij}(v,v)\right\|_{C^{k,\alpha}(M)}\\
\leq\ & \sigma(\Lambda_0,\alpha,M)\,C(k,\alpha,M)\left(2+\|R\|_{C^0(M)}+\left|\frac{\Lambda_0}{2}\right|\right)\|v\|^2_{C^{k,\alpha}(M)}\\
&+\sigma^2(\Lambda_0,\alpha,M)\,\|\nabla R\|_{C^0(M)}\|v\|^2_{C^{k,\alpha}(M)}\\
=\ & \Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)\|v\|^2_{C^{k,\alpha}(M)},
\end{aligned}$$

where the constant

$$\begin{aligned}
(5.3)\qquad &\Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)\\
:\ &=\sigma(\Lambda_0,\alpha,M)\,C(k,\alpha,M)\left(2+\|R\|_{C^0(M)}+\left|\frac{\Lambda_0}{2}\right|\right)\\
&+\sigma^2(\Lambda_0,\alpha,M)\,\|\nabla R\|_{C^0(M)}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\left\|Q_i(v,v)\right\|_{C^{k,\alpha}(M)} &\leq \sigma(\Lambda_0,\alpha,M)\,C(k,\alpha,M)\,\|v\|^2_{C^{k,\alpha}(M)}\\
&\leq \Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)\|v\|^2_{C^{k,\alpha}(M)}.
\end{aligned}$$

Finally, since

$$Q(u)(v,v)=E(u)\left(\left[Q_i(u)(v,v)\right],\left[Q_{ij}(u)(v,v)\right]\right)$$

and the *operator norms* of

$$E(\Psi_t):C^{k,\alpha}(M,T^*M)\times C^{k,\alpha}\left(M,\mathrm{Sym}^{\otimes 2}(T^*M)\right)\to C^{k,\alpha}(M,\mathbb{R}^q)$$

is of order $C_E t^{-\frac{k}{2}-\frac{\alpha}{2}}$ by Corollary 31, we have

$$\begin{aligned} (5.4) \qquad & \left\|Q\left(\Psi_t\right)(v,v)\right\|_{C^{k,\alpha}(M)} \\ \leq \quad & \Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)\left\|E\left(\Psi_t\right)\right\|_{C^{k,\alpha}(M)}\|v\|^2_{C^{k,\alpha}(M)} \\ \leq \quad & C_E\Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)t^{-\frac{k+\alpha}{2}}\|v\|^2_{C^{k,\alpha}(M)}, \end{aligned}$$

where the constant $C_E\Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)$ is uniform for all q. q.e.d.

6. The implicit function theorem: isometric embedding

In previous sections we have considered the $\frac{n(n+3)}{2}\times\infty$ matrix $P\left(\Psi_t\right)$ and its right inverse $E\left(\Psi_t\right)$. If we truncate ℓ^2 to $\mathbb{R}^{q(t)}$ with $q\left(t\right)\geq Ct^{-\frac{n}{2}-\rho}$ and consider the *modified heat kernel embedding map* $\tilde{\Psi}_t : M\to\mathbb{R}^{q(t)}$, then $E\left(\tilde{\Psi}_t\right)$ is a $q\left(t\right)\times\frac{n(n+3)}{2}$ matrix. For each fixed t, the modified $\tilde{\Psi}_t=\Psi_{t,g(t)}$ is the heat kernel embedding map in [**BBG**] for the *modified metric* g_t. The modified metrics $\{g_s\}_{0\leq s\leq t_0}$ are a compact family and depend on s smoothly. By Proposition 9 and Proposition 24, $E\left(\tilde{\Psi}_t\right)$ still has the operator bounds as in Proposition 30 and Corollary 31 for $E\left(\Psi_t\right)$. This is because from our construction of $E\left(\Psi_{t,g(s)}\right)$ (4.23), $\left\|E\left(\Psi_{t,g(s)}\right)\right\|_{C^{k,\alpha}(M)}$ is determined by the (derivatives of) inner products of the row vectors $\partial_i\Psi_{t,g(s)}$ and $\partial_i\partial_j\Psi_{t,g(s)}$ $(1\leq i\leq j\leq n)$ for the parameter s, which is in a *compact* interval $[0,t_0]$.

Now we are ready to give the proof of Theorem 1. We divide the proof into two propositions: isometric immersion and one-to-one map.

Proposition 34 (Isometric immersion))**.** Under the conditions of Theorem 1, there exists $t_0>0$ depending on (g,ρ,α), such that for the integer $q=q\left(t\right)\geq t^{-\frac{n}{2}-\rho}$ and $0<t\leq t_0$, the modified heat kernel embedding $\tilde{\Psi}_t$ can be truncated to

$$\tilde{\Psi}_t : M\to\mathbb{R}^q\subset\ell^2$$

and can be perturbed to an isometric embedding $I_t : M\to\mathbb{R}^q$, with the perturbation of $\tilde{\Psi}_t$ of order $O\left(t^{\frac{k+1}{2}-\frac{\alpha}{2}}\right)$ in the $C^{k,\alpha}$-norm.

Proof. Given the truncated heat kernel embedding $u=\tilde{\Psi}_t : M\to\mathbb{R}^{q(t)}$ with $q=q\left(t\right)\geq t^{\frac{n}{2}-\rho}$ and the error $f:=\left(\tilde{\Psi}_t\right)^*g_{can}-g$ to the isometric embedding, we consider the nonlinear functional

$$(6.1) \qquad F : C^{k,\alpha}\left(M,\mathbb{R}^q\right)\to C^{k,\alpha}\left(M,\mathbb{R}^q\right),$$

$$F\left(v\right)=v-E\left(\tilde{\Psi}_t\right)(0,f)+E\left(\tilde{\Psi}_t\right)\left(\left[Q_i\left(\tilde{\Psi}_t\right)(v,v)\right],\left[Q_{jk}\left(\tilde{\Psi}_t\right)(v,v)\right]\right).$$

We stress that this iteration is *coordinate free* and is defined on the *whole* M, as it is the coordinate expression of the iteration of tensors (see equations (12)–(21) in [**G1**]). We want to find the zeros of F. By the general implicit function theorem (e.g., Proposition A.3.4. in [**MS**]), the operator norm estimate in Corollary 31, and the uniform quadratic estimates in Proposition 33, it is enough to verify that

$$\left\|E\left(\tilde{\Psi}_t\right)\right\|_{C^{k,\alpha}(M)}\left\|E\left(\tilde{\Psi}_t\right)(0,f)\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)}\to 0$$

as $t\to 0_+$. By Corollary 31 we have

$$\left\|E\left(\tilde{\Psi}_t\right)\right\|_{C^{k,\alpha}(M)}\le C_E t^{-\frac{k+\alpha}{2}}.$$

By Theorem 10 we have $f=\left(\tilde{\Psi}_t\right)^* g_{can}-g=O\left(t^l\right)$ in the C^{k+1} norm, so for small t,

$$\|f\|_{C^{k,\alpha}\left(M,\mathrm{Sym}^{\otimes 2}(T^*M)\right)}\le Gt^l \tag{6.2}$$

for the constant

$$G:=C\left(g,l,k+1\right) \tag{6.3}$$

in Proposition 5 (when $k=l=2$, we have partial estimate of G by curvature terms in Section 9.2). By our construction in (4.23) and (4.22), we have

$$E\left(\tilde{\Psi}_t\right)(0,f)=P^T\left(\tilde{\Psi}_t\right)\begin{bmatrix} O\left(t\right)\cdot f \\ 2t\cdot\left(\begin{bmatrix} I_{\frac{n(n-1)}{2}} & 0 \\ 0 & \left(\Xi_n\left(\frac{1}{3}\right)\right)^{-1}\end{bmatrix}+O\left(t\right)\right)\cdot f\end{bmatrix}.$$

When t is small, $\left|\nabla_i\tilde{\Psi}_t\right|<<\left|\nabla_j\nabla_k\tilde{\Psi}_t\right|$, so we have

$$\begin{aligned}\left\|E\left(\tilde{\Psi}_t\right)(0,f)\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)} &= \left\|\left[\left(\nabla_j\nabla_k\tilde{\Psi}_t\right)^T\right]_{1\le j\le k\le n}\cdot O\left(t\right)\cdot f\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)} \\ &\le C_E\left(t^{-\frac{k+1}{2}-\frac{\alpha}{2}}\right)\cdot C_E t\cdot Gt^l \\ &= C_E^2 G t^{l+\frac{1}{2}-\frac{k+\alpha}{2}}, \end{aligned}\tag{6.4}$$

where $[\cdot]_{1\le j\le k\le n}$ is the notation for a matrix and we have used Proposition 24 to estimate $\left\|\nabla_j\nabla_k\tilde{\Psi}_t\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)}$. Hence,

$$\begin{aligned}&\left\|E\left(\tilde{\Psi}_t\right)\right\|_{C^{k,\alpha}(M)}\left\|E\left(\tilde{\Psi}_t\right)(0,f)\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)} \\ \le\ & C_E^3 G t^{-\frac{k+\alpha}{2}}\cdot t^{l+\frac{1}{2}-\frac{k+\alpha}{2}} \\ =\ & C_E^3 G t^{l+\frac{1}{2}-k-\alpha}\to 0\end{aligned}\tag{6.5}$$

as $t\to 0_+$, for $l+\frac{1}{2}>k+\alpha$ by our assumption.

The same quadratic estimate still holds for $\tilde{\Psi}_t$ for $0<t\le t_0$ and is *uniform* for all $q\left(t\right)\ge t^{\frac{n}{2}-\rho}$, by Corollary 31, Proposition 33, and

our remark on $\left\| E\left(\tilde{\Psi}_t\right)\right\|_{C^{k,\alpha}(M)}$ in the beginning of this subsection, as follows:

$$\begin{aligned}
& \left\| Q\left(\tilde{\Psi}_t\right)(v,v)\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)} \\
= & \left\| E\left(\tilde{\Psi}_t\right)\left(\left[Q_i\left(\tilde{\Psi}_t\right)(v,v)\right],\left[Q_{ij}\left(\tilde{\Psi}_t\right)(v,v)\right]\right)\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)} \\
\leq & \left\| E\left(\tilde{\Psi}_t\right)\right\|_{C^{k,\alpha}(M)} \cdot \Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)\|v\|^2_{C^{k,\alpha}(M,\mathbb{R}^q)} \\
\leq & C_E\Gamma\left(\Lambda_0,k,\alpha,\|R\|_{C^1}\right)t^{-\frac{k}{2}-\frac{\alpha}{2}}\|v\|^2_{C^{k,\alpha}(M,\mathbb{R}^q)}.
\end{aligned}$$

By Günther's implicit function theorem we obtain a smooth map $I_t : M \to \mathbb{R}^q$ such that $I_t^* g_{can} = g$. From this we immediately see I_t is an isometric immersion. From the implicit function theorem we also see the needed perturbation from $\tilde{\Psi}_t$ to I_t is of order $O\left(t^{l+\frac{1}{2}-\frac{k+\alpha}{2}}\right)$ in the $C^{k,\alpha}$-norm (For readers interested in more details about this, see the appendix). q.e.d.

Remark 35. The condition that $\|f\|_{C^{k,\alpha}\left(M,\mathrm{Sym}^{\otimes 2}(T^*M)\right)}$ is of order $O\left(t^l\right)$ with $l+\frac{1}{2} > k+\alpha$ is used in (6.5). Since $k \geq 2$ we must have $l \geq 2$. This is the reason that we need to modify the Ψ_t in [**BBG**] for higher-order (at least $O\left(t^2\right)$) approximation to isometry. If we can make the remainder terms in (6.2) explicit, then we can give the estimate of the smallness of t in the above implicit function theorem. See Section 9.2 for partial results in this direction.

To show the map $I_t : M \to \mathbb{R}^{q(t)}$ is one-to-one for small enough $t > 0$, we prove the following proposition.

Proposition 36. Let (M,g) be a compact Riemannian manifold with smooth metric g. Then there exists $\delta_0 > 0$, such that for $0 < t \leq \delta_0$ and $q(t) \geq Ct^{-\frac{n}{2}-\rho}$, the truncated heat kernel mapping $\Psi_t^{q(t)} : M \to \mathbb{R}^{q(t)}$ can distinguish any two points on the manifold, i.e., for any $x \neq y$ on M, $\Psi_t^{q(t)}(x) \neq \Psi_t^{q(t)}(y)$. The same is true for the isometric immersion $I_t : M \to \mathbb{R}^{q(t)}$.

Proof. The proof is adapted from Section 4 of [**SZ**]. If there is no such δ_0, then there is a sequence of $t_k \to 0$, and $x_k \neq y_k$ on M, such that

(6.6)
$$\Psi_{t_k}^{q(t_k)}(x_k) = \Psi_{t_k}^{q(t_k)}(y_k), \text{ i.e., } \phi_j(x_k) = \phi_j(y_k) \text{ for } 1 \leq j \leq q(t_k).$$

Therefore,

$$\Sigma_{j=1}^{q(t_k)} e^{-\lambda_j t_k}\phi_j(x_k)\phi_j(y_k) = \Sigma_{j=1}^{q(t_k)} e^{-\lambda_j t_k}\phi_j^2(x_k).$$

By Proposition 9 and (4.5), letting $k \to \infty$, we have

$$\lim_{k\to\infty} (4\pi t_k)^{n/2} H(t_k, x_k, y_k) = \lim_{k\to\infty} (4\pi t_k)^{n/2} H(t_k, x_k, x_k) \to 1. \tag{6.7}$$

Let $r_k = dist(x_k, y_k)$. From (6.7) we see $\lim_{k\to\infty} r_k = 0$. Otherwise, $r_k \geq c_0 > 0$, by the compactness of M we can assume x_k and y_k converge to different limits $x_\infty \neq y_\infty$ on M. So in the left-side of (6.7) we have $\lim_{k\to\infty} H(t_k, x_k, y_k) = \lim_{k\to\infty} H(t_k, x_\infty, y_\infty) = 0$, a contradiction. We further claim that when k is large,

$$r_k := \operatorname{dist}(x_k, y_k) \leq A t_k^{\frac{1}{2}} \tag{6.8}$$

for some constant $A > 0$. Otherwise, $r_k \to 0$ and $\frac{r_k}{\sqrt{t_k}} \to +\infty$. By the Minakshisundaram–Pleijel expansion,

$$\begin{aligned}\lim_{k\to\infty} (4\pi t_k)^{n/2} |H(t_k, x_k, y_k)| &= \lim_{k\to\infty} e^{-\frac{r_k^2}{4t_k}} |U(t_k, x_k, y_k)| \\ &\leq e^{\lim_{k\to\infty}\left(-\frac{r_k^2}{4t_k}\right)} \cdot 2 = 0,\end{aligned}$$

contradicting (6.7).

By (6.8), for large k, we can write

$$y_k = \exp_{x_k}\left(2\sqrt{t_k} v_k\right), \text{ for } 0 \neq v_k \in T_{x_k}M, \ |v_k| = O(1),$$

and $y_k^s = \exp_{x_k}(s v_k)$ for $-1 \leq s \leq 2$. We consider the function

$$f_k(s) := \frac{(H(t_k, x_k, y_k^s))^2}{H(t_k, x_k, x_k) H\left(t_k, y_k^s, y_k^s\right)} = \frac{|\langle \Psi_{t_k}(x_k), \Psi_{t_k}(y_k^s)\rangle|^2}{|\Psi_{t_k}(x_k)|^2 \left|\Psi_{t_k}\left(y_k^s\right)\right|^2}, \quad -1 \leq s \leq 2.$$

By the Cauchy–Schwartz inequality, $0 \leq f_k(s) \leq 1$. By the definition of $f_k(s)$ and our assumption $x_k = y_k$, we have $f_k(0) = f_k(1) = 1$, achieving the maximum of f_k on $[-1, 2]$. So there exists some $s_k \in [0, 1]$, $f_k''(s_k) = 0$.

In the following we let $z = (t, x, x)$, $z_k = (t_k, x_k, x_k)$. By (4.1), we have

$$H(t_k, x_k, x_k) = (4\pi t_k)^{-n/2} U(z_k), \tag{6.9}$$

(6.10)

$$\begin{aligned}H(t_k, y_k^s, y_k^s) = (4\pi t_k)^{-n/2} \\ \times \left[U(z_k) + 2s\sqrt{t_k} A(z_k)(v_k) + s^2 t_k E(z_k)(v_k) + O\left(s^3 t_k^{\frac{3}{2}} |v_k|^3\right)\right]\end{aligned}$$

in the C^3-norm, where $A(t,x,x) = \partial_y U(t,x,y)|_{x=y}$ and $E(z)(v)$ is quadratic in v. We also have

$$\begin{aligned}
&(6.11)\ H(t_k, x_k, y_k^s) \\
&\quad = (4\pi t_k)^{-n/2} e^{-s^2|v_k|^2} U(t_k, x_k, y_k^s) \\
&\quad = (4\pi t_k)^{-n/2} \left[1 - s^2 |v_k|^2 + O\left(s^4\right)\right] \\
&\qquad \times \left[U(z_k) + s\sqrt{t_k} A(z_k)(v_k) + s^2 t_k B(z_k)(v_k) + O\left(s^3 t_k^{\frac{3}{2}} |v_k|^3\right)\right],
\end{aligned}$$

where $B(z)(v)$ is quadratic in v. Therefore, as $k \to \infty$, in the C^3-norm we have

$$\begin{aligned}
f_k(s) &= \left[1 - 2s^2 |v_k|^2 + O\left(s^4 |v_k|^4\right)\right] \\
&\quad \times \frac{\left[U(z_k) + s\sqrt{t_k} A(z_k)(v_k) + s^2 t_k B(z_k)(v_k) + O\left(s^3 t_k^{\frac{3}{2}} |v_k|^3\right)\right]^2}{U(z_k)\left[U(z_k) + 2s\sqrt{t_k} A(z_k)(v_k) + s^2 t_k E(z_k)(v_k) + O\left(s^3 t_k^{\frac{3}{2}} |v_k|^3\right)\right]} \\
&= \left[1 - 2s^2 |v_k|^2 + O\left(s^4 |v_k|^4\right)\right] \\
&\quad \times \left[\frac{1 + 2s\sqrt{t_k}\widetilde{A}(z_k)(v_k) + s^2 t_k \widetilde{B}(z_k)(v_k) + O\left(s^3 t_k^{\frac{3}{2}} |v_k|^3\right)}{1 + 2s\sqrt{t_k}\widetilde{A}(z_k)(v_k) + s^2 t_k \widetilde{E}(z_k)(v_k) + O\left(s^3 t_k^{\frac{3}{2}} |v_k|^3\right)}\right]
\end{aligned}$$

($\widetilde{A}$, $\widetilde{B}$, and $\widetilde{E}$ are functions A, B, and E divided by $U(z_k)$, respectively)

$$\begin{aligned}
&= \left[1 - 2s^2 |v_k|^2 + O\left(s^4 |v_k|^4\right)\right]\left[1 + s^2 t_k \widehat{C}(z_k)(v_k) + O\left(s^3 t_k^{\frac{3}{2}} |v_k|^3\right)\right] \\
&= 1 + s^2 |v_k|^2 \left(-2 + t_k \widehat{C}(z_k)\left(\frac{v_k}{|v_k|}\right)\right) + O\left(s^3 t_k^{\frac{3}{2}} |v_k|^3\right),
\end{aligned}$$

where the function $\widehat{C}(z)$ is constructed from $\widetilde{A}$, $\widetilde{B}$, and $\widetilde{E}$ and $\widehat{C}(z)(v)$ is quadratic in v. Hence, as $k \to \infty$,

$$0 = f_k''(s_k) = 2|v_k|^2 \left[-2 + O(t_k)\right] + O\left(s_k t_k^{\frac{3}{2}} |v_k|^3\right) = 2|v_k|^2 \left[-2 + O(1)\right].$$

But by our assumption $v_k \neq 0$, so the right-hand side will be nonzero for large k, a contradiction.

The proof of the one-to-one property of $I_t : M \to \mathbb{R}^{q(t)}$ is almost identical to that of Ψ_t. This is because we have $\|I_t - \Psi_t\|_{C^{k,\alpha}(M)} \leq Ct^{l+\frac{1}{2}-\frac{k+\alpha}{2}}$ for $k \geq 2$ from Proposition 34, so the function

$$f_k(s) := \frac{\left|\langle I_t(x_k), I_t(y_k^s)\rangle\right|^2}{|I_t(x_k)|^2 \left|I_t(y_k^s)\right|^2}$$

has the same properties as the above functions $f_k(s)$ for Ψ_t, up to the second-order derivatives. This is because if we replace $H(t,x,y)$ for Ψ_t

by $\langle I_t(x), I_t(y)\rangle$ for I_t in the key estimates (6.6), (6.7), (6.9), (6.10), and (6.11), the argument still holds. The proposition follows. q.e.d.

Corollary 37. *Let (M, g) be a compact Riemannian manifold with smooth metric g. Then there exists an integer $N_0 > 0$ depending on g, such that the first N_0 eigenfunctions $\{\phi_j\}_{j=1}^{N_0}$ can distinguish any two points on M, i.e., for any $x \neq y$ on M, there exists some $j_0 \in \{1, 2, \cdots N_0\}$, such that $\phi_{j_0}(x) \neq \phi_{j_0}(y)$.*

Proof. Take $N_0 = q(\delta_0)$ in the above proposition, and note that $\Psi_t^{N_0}(x) = \Psi_t^{N_0}(y) \Longleftrightarrow \phi_j(x) = \phi_j(y)$ for $1 \leq j \leq N_0$. q.e.d.

7. Geometry of the embedded images in ℓ^2 (and $\mathbb{R}^{q(t)}$)

In this section we study the geometry of the embedded images $\Psi_t(M)$ and $I_t(M)$ in ℓ^2. We first combine Theorem 18 and Proposition 21 to give the following consequence on the second fundamental form and mean curvature of the embedded image $\Psi_t(M) \subset \ell^2$.

Corollary 38. *For any $x \in M$, let $(x_1, \cdots, x_n)$ be the normal coordinates near x. The second fundamental form $A(x,t) = \Sigma_{1\leq i\leq j\leq n} h_{ij}(x,t)\, dx^i dx^j$ of the submanifold $\Psi_t(M) \subset \ell^2$ can be written as*

$$A(x,t) = \frac{1}{\sqrt{2t}}\left(\Sigma_{i=1}^n \sqrt{3} a_{ii}(x,t)\left(dx^i\right)^2 + \Sigma_{1\leq j<k\leq n} 2a_{jk}(x,t)\, dx^j dx^k\right),$$

where $a_{jk}(x,t)$ $(1 \leq j \leq k \leq n)$ are vectors in ℓ^2. Then as $t \to 0_+$, we have the following:

1) *For any two subsets $\{i,j\}$ and $\{k,l\} \subset \{1, 2, \cdots, n\}$,*

$$\langle a_{ij}, a_{ij}\rangle \to 1, \tag{7.1}$$

$$\langle a_{ij}, a_{kl}\rangle \to 0, \text{ if } \{i,j\} \neq \{k,l\} \text{ and } \{i,k\} \neq \{j,l\}, \tag{7.2}$$

$$\langle a_{ii}, a_{jj}\rangle \to \frac{1}{3}, \text{ if } i \neq j. \tag{7.3}$$

2) *The mean curvature vector $H(x,t) = \frac{1}{n}\Sigma_{i=1}^n h_{ii}(x,t)$, after scaled by a factor $\sqrt{t}$, converges to constant length:*

$$\sqrt{t}\,|H(x,t)| \to \sqrt{\frac{n+2}{2n}}. \tag{7.4}$$

The convergence is uniform for all x on M in the C^r-norm for any $r \geq 0$.

Proof. From Proposition 21 we have

$$|\langle \nabla_i \nabla_j \Phi_t, \nabla_k \Phi_t\rangle(x)| = O\left(t^{-n/2}\right). \tag{7.5}$$

Therefore, for the *normalized heat kernel embedding* $\Psi_t = \sqrt{2}\,(4\pi)^{n/4}$ $t^{\frac{n+2}{4}} \cdot \Phi_t$, its first derivative and second derivative vectors become orthogonal as $t \to 0_+$ by (7.5):

$$\left|\langle \nabla_i \nabla_j \Psi_t, \nabla_k \Psi_t \rangle (x)\right| \to 2\,(4\pi)^{n/2}\, t^{\frac{n+2}{2}} \cdot t^{-n/2} \cdot O(1) = Ct \cdot O(1) \to 0.$$

So as $t \to 0_+$, the *second fundamental form* at $\Psi_t(x)$ on $\Psi_t(M) \subset \ell^2$ is approximated by the second-order terms in the Taylor expansion of $\Psi_t : M \to \ell^2$ near x on M, i.e.,

$$\lim_{t\to 0_+} \left[A(x,t) - \left(\Sigma_{1\le i\le j\le n} \nabla_i \nabla_j \Psi_t(x,t)\, dx^i dx^j\right)\right] = 0.$$

From Proposition 21 we have

$$\begin{aligned}
(7.6) \qquad & \langle \nabla_j \nabla_i \Psi_t, \nabla_m \nabla_k \Psi_t \rangle (x) \\
\to \;& 2\,(4\pi)^{n/2}\, t^{\frac{n+2}{2}} \cdot \left(\frac{1}{2t}\right)^2 \frac{1}{(4\pi t)^{n/2}} \cdot (\delta_{ij}\delta_{km} + \delta_{im}\delta_{jk} + \delta_{ik}\delta_{jm}) \\
= \;& \frac{1}{2t}\,(\delta_{ij}\delta_{km} + \delta_{im}\delta_{jk} + \delta_{ik}\delta_{jm}).
\end{aligned}$$

In particular, as $t \to 0_+$,

$$\begin{aligned}
\left|\frac{1}{\sqrt{2t}} \cdot \sqrt{3} a_{ii}(x,t)\right| &= |\nabla_i \nabla_i \Psi_t| \to \frac{\sqrt{3}}{\sqrt{2t}}, \\
\left|\frac{1}{\sqrt{2t}} \cdot a_{jk}(x,t)\right| &= |\nabla_j \nabla_k \Psi_t| \to \frac{1}{\sqrt{2t}} \ (j \neq k),
\end{aligned}$$

and so (7.1) follows. Similarly, (7.2) and (7.3) follow from (7.6). For the mean curvature, we have

$$H(x,t) = \frac{1}{n}\Sigma_{i=1}^n h_{ii}(x,t) = \frac{1}{n}\frac{\sqrt{3}}{\sqrt{2t}}\left(\Sigma_{i=1}^n a_{ii}(x,t)\right).$$

Using $|a_{ii}| \to 1$ and $\langle a_{ii}, a_{jj}\rangle \to \frac{1}{3}$ as $t \to 0_+$, we have

$$|H(x,t)|^2 \to \frac{1}{n^2}\frac{3}{2t}\left(n \cdot 1 + n(n-1)\cdot \frac{1}{3}\right) = \frac{1}{2t}\frac{n+2}{n},$$

so (7.4) follows. q.e.d.

Remark 39. In Corollary 38, we have the following:

1) It is unknown if $\lim_{t\to 0_+} a_{jk}(x,t)$ exists, but $\lim_{t\to 0_+} \langle a_{ij}(x,t), a_{kl}(x,t)\rangle$ exists. There exists isometry $I(x,t) : \ell^2 \to \ell^2$, such that

$$a_{jk} := \lim_{t\to 0_+} I(x,t) \cdot a_{jk}(x,t) \quad (1 \le j \le k \le n)$$

exists, and $\{a_{jk}\}_{1\le j\le k\le n}$ is a fixed basis in $\mathbb{R}^{\frac{n(n+1)}{2}} \subset \ell^2$ satisfying the inner product relations in item 1 of Corollary 38.

2) The length of the mean curvature of $\Psi_t(M) \subset \ell^2$ converges to constant on M, but the constant is large (of order $t^{-\frac{1}{2}}$) as $t \to 0_+$ by (7.4). Intuitively, this is because the embedding Ψ_t uses more and more high-frequency eigenfunctions in its ℓ^2 norm as $t \to 0_+$, making the image $\Psi_t(M)$ evenly oscillating at all x on M.
3) By Proposition 9, the above results still hold if we replace ℓ^2 by $\mathbb{R}^{q(t)}$, for $q(t) \geq Ct^{-\frac{n}{2}-\rho}$ and sufficiently small $t > 0$.

As for our isometric embeddings $I_t : M \to \mathbb{R}^{q(t)}$ ($q(t) \geq Ct^{-\frac{n}{2}-\rho}$ or $q(t) = \infty$), they are obtained by $C^{k,\alpha}$-perturbation of $\tilde{\Psi}_t : M \to \mathbb{R}^{q(t)}$ of order $O\left(t^{l+\frac{1}{2}-\frac{k+\alpha}{2}}\right)$, with $k \geq 2$. Since the second fundamental form and mean curvature of any embedding $f : M \to \mathbb{R}^q$ are determined by up to the *second-order* derivatives of f, the statements in Corollary 38 also hold for the isometric embedding image $I_t(M) \subset \mathbb{R}^{q(t)}$, noticing that $\tilde{\Psi}_t := \Psi_{t,g(t)}$ is the heat kernel embedding map for the metric $g(t)$ on M. For the same reason, for any $0 \leq r \leq k$, the r-jet relations of I_t as $t \to 0_+$ are the same as those for Ψ_t in [**Z**]. This gives many constraints of the image $I_t(M) \subset \mathbb{R}^{q(t)}$.

8. Example $M = S^1$

As a concrete example, we write down $P(u)$ and $E(u)$ explicitly for the case $M = S^1 \simeq [0, 2\pi] / \sim$. Although the example is trivial, it exhibits almost all features of general cases. We use it to illustrate the proofs of our main results.

For the eigenvalue $\lambda_{2k-1} = \lambda_{2k} = k^2$, the L^2 orthonormal eigenfunctions are pairs $\phi_{2k-1}(x) = \frac{1}{\sqrt{\pi}} \cos kx$ and $\phi_{2k}(x) = \frac{1}{\sqrt{\pi}} \sin kx$. The heat kernel embedding $u : S^1 \to \mathbb{R}^{2q}$ is

$$u(x) = \sqrt{\frac{2}{\pi}} (4\pi)^{n/4} t^{\frac{n+2}{4}} \left\{ \left(e^{-\frac{k^2}{2}t} \cos kx, e^{-\frac{k^2}{2}t} \sin kx \right) \right\}_{1 \leq k \leq q},$$

so the system (3.4) becomes

$$R_1(x) \cdot v = h_1(x), \; R_2(x) \cdot v = f_{11}(x),$$

with the two row vectors R_1 and R_2 being

$$R_1(x) = \sqrt{\frac{2}{\pi}} (4\pi)^{n/4} t^{\frac{n+2}{4}} \left\{ \left(-e^{-\frac{k^2}{2}t} k \sin kx, e^{-\frac{k^2}{2}t} k \cos kx \right) \right\}_{1 \leq k \leq q},$$

$$R_2(x) = \sqrt{\frac{2}{\pi}} (4\pi)^{n/4} t^{\frac{n+2}{4}} \left\{ \left(-e^{-\frac{k^2}{2}t} k^2 \cos kx, -e^{-\frac{k^2}{2}t} k^2 \sin kx \right) \right\}_{1 \leq k \leq q}.$$

It is easy to check that $R_1(x)$ and $R_2(x)$ are *orthogonal*, so the solution v with minimal Euclidean norm is $v(x) = \frac{h_1(x)}{|R_1(x)|^2} R_1(x) + \frac{f_{11}(x)}{|R_2(x)|^2} R_2(x)$.

Hence,

$$E(u)=\begin{bmatrix} R_1(x)/|R_1(x)|^2 \\ R_2(x)/|R_2(x)|^2 \end{bmatrix}, \qquad E(u)(0,f_{11})=\frac{f_{11}(x)}{|R_2(x)|^2}R_2(x).$$

Recall the following well-known lemma.

Lemma 40. *$\lim_{t\to 0_+} t^{\frac{m+1}{2}}\Sigma_{k=1}^{\infty}k^m e^{-k^2t}=\int_0^\infty \mu^m e^{-\mu^2}d\mu$, and $\Sigma_{k=1}^{\infty} k^m e^{-k^2t}\le Kt^{-\frac{m+1}{2}}$, where the constant $K=\int_0^\infty \mu^m e^{-\mu^2}d\mu$.*

For q large, using $\int_0^\infty \mu^2 e^{-\mu^2}d\mu=\frac{\sqrt{\pi}}{4}$ and $\int_0^\infty \mu^4 e^{-\mu^2}d\mu=\frac{3\sqrt{\pi}}{8}$, we have

$$\begin{aligned}|R_1(x)|^2&=\frac{2}{\pi}(4\pi)^{n/2}t^{\frac{n+2}{2}}\Sigma_{k=1}^{q}k^2e^{-k^2t}\\&\to\frac{2}{\pi}(4\pi)^{1/2}t^{\frac{3}{2}}\cdot t^{-\frac{3}{2}}\int_0^\infty \mu^2e^{-\mu^2}d\mu=1,\\|R_2(x)|^2&=\frac{2}{\pi}(4\pi)^{n/2}t^{\frac{n+2}{2}}\Sigma_{k=1}^{q}k^4e^{-k^2t}\\&\to\frac{2}{\pi}(4\pi)^{1/2}t^{\frac{3}{2}}\cdot t^{-\frac{5}{2}}\int_0^\infty \mu^4e^{-\mu^2}d\mu=\frac{3}{2}t^{-1}.\end{aligned}$$

These agree with $\Psi_t^*g_{can}\to g$ in (1.2) and the mean curvature length $|H(x,t)|\to\sqrt{\frac{1+2}{2\cdot 1}}t^{-\frac{1}{2}}$ in Corollary 2, respectively. Thus, for $q=q(t)$ large, in C^3 convergence we have

$$E(u)\to\begin{bmatrix} R_1(x) \\ \frac{2}{3}tR_2(x)\end{bmatrix}, \qquad E(u)(0,f_{11})\to\frac{2t}{3}R_2(x)f_{11}(x). \tag{8.1}$$

We have the C^2 norm (according to Definition 14) for vector-valued functions

$$\begin{aligned}\|R_1(x)\|_{C^2(M)}&=\sqrt{\frac{2}{\pi}}(4\pi)^{n/4}t^{\frac{n+2}{4}}\left[\Sigma_{k=1}^{q}e^{-k^2t}k^6\right]^{\frac{1}{2}}\to Ct^{\frac{3}{4}}t^{-\frac{7}{4}}=Ct^{-1},\\\|R_2(x)\|_{C^2(M)}&=\sqrt{\frac{2}{\pi}}(4\pi)^{n/4}t^{\frac{n+2}{4}}\left[\Sigma_{k=1}^{q}e^{-k^2t}k^8\right]^{\frac{1}{2}}\to Ct^{\frac{3}{4}}t^{-\frac{9}{4}}=Ct^{-\frac{3}{2}},\\\|E(u)\|_{C^2(M)}&=\sqrt{\frac{2}{\pi}}(4\pi)^{n/4}t^{\frac{n+2}{4}}\cdot\left[\Sigma_{k=1}^{q}e^{-k^2t}k^6+t^2\cdot\Sigma_{k=1}^{q}e^{-k^2t}k^8\right]^{1/2}\\&\to Ct^{-1}.\end{aligned}$$

Notice that for S^1, the curvature tensor $R\equiv 0$, so $Ric_g-\frac{1}{2}g\cdot S_g\equiv 0$. By [**BBG**, Theorem 5] we have $f_{11}=\Psi_t^*g_{can}-g=O\left(t^2\right)$ as $t\to 0_+$ (it may be of higher vanishing order $O(t^p)$ for some $p>2$, but here we only use $O\left(t^2\right)$ to illustrate our method). So by (8.1) we have

$$\|E(u)(0,f_{11})\|_{C^2(M)}\to\left\|\frac{2t}{3}\cdot t^2R_2(x)\right\|_{C^2(M)}=Ct^3\cdot\|R_2(x)\|_{C^2(M)}=Ct^{\frac{3}{2}}.$$

Then

$$\|E(u)\|_{C^2(M)} \|E(u)(0,f_{11})\|_{C^2(M)} \to Ct^{-1} \cdot Ct^{\frac{3}{2}} = Ct^{\frac{1}{2}},$$

and similarly (using the interpolation technique in Lemma 8 to estimate the $C^{2,\alpha}$-norm from the C^2 and C^3-norms),

$$\|E(u)\|_{C^{2,\alpha}(M)} \|E(u)(0,f_{11})\|_{C^{2,\alpha}(M)} \to Ct^{-1-\frac{\alpha}{2}} \cdot Ct^{\frac{3}{2}-\frac{\alpha}{2}} = Ct^{\frac{1}{2}-\alpha} \to 0$$

for $0 < \alpha < \frac{1}{2}$. We see the estimates of the orders are exactly the same as obtained by the off-diagonal expansion of heat kernel method.

By Günther's implicit function theorem we obtain the isometric embeddings of S^1 into $\mathbb{R}^{q(t)}$.

9. Appendix

9.1. The implicit function theorem. For the sake of completeness, in this appendix we give a proof of Günther's implicit function theorem (Theorem 16) by applying Proposition A.3.4. in [**MS**] (which is an abstract implicit function theorem) to the nonlinear function $F : C^{k,\alpha}(M,\mathbb{R}^q) \to C^{k,\alpha}(M,\mathbb{R}^q)$ defined in Section 6. In particular, we obtained a little more: First, the constant θ in Günther's theorem is made explicit in (9.2). Second, the needed perturbation of Ψ_t is shown to be of order $O\left(t^{l+\frac{1}{2}-\frac{k+\alpha}{2}}\right)$ in the $C^{k,\alpha}$-norm.

Proposition 41 (Proposition A.3.4. of [**MS**])**.** Let X, Y be Banach spaces, and let U be an open set in X. The map $F : X \to Y$ is continuous differentiable. For $x_0 \in U$, $D := dF(x_0) : X \to Y$ is surjective and has a bounded linear right inverse $Q : Y \to X$, with $\|Q\| \le c$. Suppose that there exists $\delta > 0$ such that $x \in B_\delta(x_0) \subset U$

$$x \in B_\delta(x_0) \subset U \implies \|dF(x) - D\| \le \frac{1}{2c}.$$

Suppose $\|F(x_0)\| < \frac{\delta}{4c}$; then there exists a unique $x \in B_\delta(x_0)$ such that

$$F(x) = 0,\ x - x_0 \in \mathrm{Image}Q,\ \|x - x_0\| \le 2c\,\|F(x_0)\|.$$

Applying the above proposition to our case, we have (following their notations)

$$X = C^{k,\alpha}(M,\mathbb{R}^q),\ Y = C^{k,\alpha}(M,\mathbb{R}^q),\ F : X \to Y,$$

$$F(v) = v - E\left(\tilde{\Psi}_t\right)(0,f) + E\left(\tilde{\Psi}_t\right)\left(\left[Q_i\left(\tilde{\Psi}_t\right)(v,v)\right], \left[Q_{jk}\left(\tilde{\Psi}_t\right)(v,v)\right]\right),$$

$x_0 = 0$, x solution, $F(x) = 0$,

$$c = \left\|(dF(0))^{-1}\right\| = \left\|(id)^{-1}\right\| = 1,$$

$$\|F(0)\| = \left\|E\left(\tilde{\Psi}_t\right)(0,f)\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)} \le C_E^2 G t^{l+\frac{1}{2}-\frac{k+\alpha}{2}} \text{ (by (6.4))},$$

$$\|dF(v) - dF(0)\| \le \left\|E\left(\tilde{\Psi}_t\right)\right\|_{C^{k,\alpha}(M)} \cdot \Gamma\left(\Lambda_0, k, \alpha, \|R\|_{C^1}\right) \cdot \|v\|_{C^{k,\alpha}(M,\mathbb{R}^q)},$$

where the last inequality is from (5.2). Since the δ should satisfy

$$\|v-0\| \leq \delta \Longrightarrow \|dF(v) - dF(0)\| \leq \frac{1}{2c},$$

we can take

$$\begin{aligned} \delta &= \frac{1}{2\left\|E\left(\tilde{\Psi}_t\right)\right\|_{C^{k,\alpha}(M)} \cdot \Gamma\left(\Lambda_0, k, \alpha, \|R\|_{C^1}\right)} \\ &\geq \frac{t^{\frac{k+\alpha}{2}}}{2C_E \cdot \Gamma\left(\Lambda_0, k, \alpha, \|R\|_{C^1}\right)} (\text{by } (4.26)). \end{aligned}$$

The condition

$$\|F(0)\| \leq \frac{\delta}{4c} \tag{9.1}$$

is translated to

$$\left\|E\left(\tilde{\Psi}_t\right)(0,f)\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)} \leq \frac{1}{4} \cdot \frac{1}{2\left\|E\left(\tilde{\Psi}_t\right)\right\|_{C^{k,\alpha}(M)} \cdot \Gamma\left(\Lambda_0, k, \alpha, \|R\|_{C^1}\right)},$$

or

$$\begin{aligned} &\left\|E\left(\tilde{\Psi}_t\right)(0,f)\right\|_{C^{k,\alpha}(M,\mathbb{R}^q)} \left\|E\left(\tilde{\Psi}_t\right)\right\|_{C^{k,\alpha}(M)} \\ &\leq \left(8\Gamma\left(\Lambda_0, k, \alpha, \|R\|_{C^1}\right)\right)^{-1} := \theta. \end{aligned} \tag{9.2}$$

So the constant θ is essentially determined by $\Gamma\left(\Lambda_0, k, \alpha, \|R\|_{C^1}\right)$ in (5.3), which in turn depends on $\|R\|_{C^1}$ and $\sigma\left(\Lambda_0, \alpha, M\right) = \left\|\left(\Delta_{(r)} - \Lambda_0\right)^{-1}\right\|_{\text{op}}$ of the smoothing operators (note $C(k,\alpha,M) = n^k$ in Lemma 15 is independent on g). In terms of t the condition (9.1) is

$$C_E^2 G t^{l+\frac{1}{2}-\frac{k+\alpha}{2}} \leq \frac{t^{\frac{k+\alpha}{2}}}{8C_E \Gamma\left(\Lambda_0, k, \alpha, \|R\|_{C^1}\right)},$$

i.e.,

$$t \leq \left(8C_E^3 \cdot \Gamma\left(\Lambda_0, k, \alpha, \|R\|_{C^1}\right) \cdot G\right)^{-\frac{1}{(l+1/2)-(k+\alpha)}} := t_0. \tag{9.3}$$

So we see the smallness of t in our implicit function theorem depends on the following:

1) the constant C_E in (4.26) in the $C^{k,\alpha}$-norm estimate for $E\left(\tilde{\Psi}_t\right)$ (essentially the derivative estimates of $\tilde{\Psi}_t$);
2) the constant $\Gamma\left(\Lambda_0, k, \alpha, \|R\|_{C^1}\right)$ in (5.3) related to the smoothing operators $\left(\Delta_{(r)} - \Lambda_0\right)^{-1}$;
3) the constant G in (6.3) from the near diagonal expansion of the heat kernel of (M, g).

These constants are related to the dimension, diameter, volume, and curvature bounds of (M,g). Since the exponent $-\frac{1}{(l+1/2)-(k+\alpha)}$ in (9.3) is negative, we see the smaller the constants C_E, $\Gamma(\Lambda_0,k,\alpha,\|R\|_{C^1})$, and G are, the smaller the embedding dimension $q(t_0)$ is.

If we know t_0, we can obtain the estimate of the minimal embedding dimension $q(t_0)\geq t_0^{-\frac{n}{2}-\rho}$. From the above proposition the solution x satisfies $\|x\|\leq 2c\|F(0)\|$, i.e., the perturbation of $\tilde{\Psi}_t$ is of order

$$\|x\|_{C^{k,\alpha}(M,\mathbb{R}^q)}\leq 2C_E^2Gt^{l+\frac{1}{2}-\frac{k+\alpha}{2}}=O\left(t^{l+\frac{1}{2}-\frac{k+\alpha}{2}}\right).$$

9.2. The quadratic remainder. We give the estimate of the constant G in (6.3). To begin with, we state a refined version of (1.2) in [**BBG**]. The following lemma is well-known in physics literature.

Lemma 42. *Same notations as in Theorem* 4. *Then as* $t\to 0_+$, *we have*

$$\begin{aligned}&\left(\Psi_t^*g_{can}\right)(x)=g(x)+\frac{t}{3}\left(\frac{1}{2}S_g\cdot g-Ric_g\right)\\ &+t^2\left[u_2(x,x)+\Sigma_{i,j=1}^n 2\partial_{\bar{j}}\partial_i u_1(x,y)|_{x=y}\right]dx^idx^j+O\left(t^3\right).\end{aligned}\tag{9.4}$$

Proof. From (4.1) we have

$$\partial_iH(t,x,y)=\frac{1}{(4\pi t)^{n/2}}e^{-\frac{r^2}{4t}}\left[-\frac{\partial_i(r^2)}{4t}U+\partial_iU\right],\tag{9.5}$$

$$\begin{aligned}\partial_j\partial_iH(t,x,y)&=\frac{1}{(4\pi t)^{n/2}}e^{-\frac{r^2}{4t}}\left[-\frac{\partial_j(r^2)}{4t}\left(-\frac{\partial_i(r^2)}{4t}U+\partial_iU\right)\right.\\ &\left.+\left(-\frac{\partial_j\partial_i(r^2)}{4t}U-\frac{\partial_i(r^2)}{4t}\partial_jU+\partial_j\partial_iU\right)\right].\end{aligned}\tag{9.6}$$

From Lemma 19 and (4.2), letting $x=y$, we have

$$\begin{aligned}&\partial_{\bar{j}}\partial_iH(t,x,y)\big|_{x=y}\\ =&\frac{1}{(4\pi t)^{n/2}}\left[-\frac{\partial_{\bar{j}}\partial_i(r^2)}{4t}\left(u_0+tu_1+t^2u_2+O\left(t^2\right)\right)\right.\\ &\left.-\frac{\partial_i(r^2)}{4t}\partial_{\bar{j}}U+\partial_{\bar{j}}\partial_i\left(u_0+tu_1+O(t)\right)\right].\end{aligned}$$

Letting $i = j$ and using Lemma 19 in the above identity, we have for $V_i = \frac{\partial}{\partial x^i}$,

$$\begin{aligned}
&(\Psi_t^* g_{can})(V_i, V_i)(x) \\
&= 2(4\pi)^{\frac{n}{2}} t^{\frac{n+2}{2}} \cdot \frac{1}{(4\pi t)^{n/2}} \\
&\times \left[-\frac{\partial_{\bar{i}}\partial_i (r^2)}{4t} \left(u_0 + tu_1 + t^2 u_2 + O(t^2)\right) + \partial_{\bar{i}}\partial_i (u_0 + tu_1 + O(t)) \right] \Bigg|_{x=y} \\
&= \left[g - \frac{t}{3}\left(Ric_g - \frac{1}{2} S_g \cdot g \right) \right] (V_i, V_i) + t^2 [u_2(x,x) \\
&+ 2\partial_{\bar{i}}\partial_i u_1(x,y)|_{x=y}] + O(t^3).
\end{aligned}$$

Since $(\Psi_t^* g_{can})(V, W)$ is bilinear in V and W, the proposition follows. q.e.d.

Using the higher-order expansion of $H(t,x,y)$ in terms of curvature terms, it seems possible to make the quadratic terms in the above lemma explicit. On pp. 224–225 of [**BeGaM**] and Theorem 3.3.1 of [**Gil**] there is an explicit

$$u_2(x,x) = \frac{1}{180}|R_g(x)|^2 - \frac{1}{180}|Ric_g(x)|^2 + \frac{1}{72}|S_g(x)|^2 - \frac{1}{30}\Delta_g S_g(x),$$

where R_g is the Riemannian curvature tensor,

$$|R_g(x)|^2 = \Sigma_{1 \le i,j,k,l \le n} |R_g(x)(V_i, V_j, V_k, V_l)|^2$$

for the basis $V_i = \frac{\partial}{\partial x^i}$ of normal coordinates $\{x^i\}_{1 \le i \le n}$ near x, and similarly for $|Ric_g(x)|^2$. It remains to compute $\partial_{\bar{i}}\partial_i u_1(x,y)|_{x=y}$. Physics literature (e.g., [**DeFo**]) gives

(9.7)

$$\begin{aligned}
&\partial_{\bar{j}}\partial_i u_1(x,y)|_{x=y} \\
&= \Bigg[\frac{1}{20}\partial_j\partial_i S_g - \frac{1}{60}\Delta_g Ric_g(V_j, V_i) + \frac{1}{36} S_g Ric_g(V_j, V_i) \\
&\quad - \frac{1}{45}\Sigma_{k=1}^n Ric_g(V_j, V_k) Ric_g(V_i, V_k) \\
&\quad + \frac{1}{90}\Sigma_{k,l=1}^n \left(Ric_g(V_k, V_l) R(V_k, V_j, V_l, V_i) + |R(V_k, V_l, V_j, V_i)|^2 \right) \Bigg](x).
\end{aligned}$$

Let's take $l = 2$ in (2.5). Then $h_1 = -A_1(g) = \frac{1}{3}\left(Ric_g - \frac{1}{2}S_g \cdot g\right)$, and $h_j = 0$ for all $j \ge 2$. We have

$$(9.8) \qquad \left(\tilde{\Psi}_t^* g_{can}\right) = G(t,t) = g + t^2 [A_{1,1}(h_1) + A_2(g)] + O(t^3),$$

where $A_2(g) = u_2(x,x) + \Sigma_{i,j=1}^n 2\partial_{\bar{j}}\partial_i u_1(x,y)|_{x=y}$ by the above lemma, and $A_{1,1}(h_1)$ (first-order variation of $A_1(g) = \frac{1}{3}\left(\frac{1}{2}S_g \cdot g - Ric_g\right)$) can be computed by the formulae of the variation of curvature tensors (cf. [**Be**, Theorem 1.174]):

$$\begin{aligned} Ric'_g h &= \frac{1}{2}\Delta_{(2)} h - \delta_g^* (\delta_g h) - \frac{1}{2}\nabla_g d(tr_g h), \\ S'_g h &= \Delta_g (tr_g h) + \delta_g (\delta_g h) - g(Ric_g, h), \end{aligned}$$

where δ_g is the divergence and δ_g^* is the formal adjoint and the notation "$'$" means the derivative with respect to the variation h of g.

If the derivation of (9.7) is rigorous, putting all these into (9.8), it appears that we can control the constant G in (6.2) by

$$G \leq C(n) \left(\|Ric_g\|_{C^{4,\alpha}(M)} + \|R_g\|^2_{C^{2,\alpha}(M)} \right) \tag{9.9}$$

for small $t > 0$, with a constant $C(n)$ only depending on n.

References

[A] H. Abdallah, *Embedding Riemannian manifolds via their eigenfunctions and their heat kernel,* Bull. Korean Math. Soc. **49** (2012), No. 5, pp. 939–947, MR 3012963, Zbl 1261.53051.

[BBG] P. Bérard, G. Besson & S. Gallot, *Embedding Riemannian Manifolds by Their Heat kernel*, Geom. and Func. Analysis, vol. 4, no. 4, (1994), MR 1280119, Zbl 0806.53044.

[Be] A. Besse, *Einstein Manifolds*, Reprint of 1987 ed., Springer 2002, MR 2371700, Zbl 1147.53001.

[BeGaM] M. Berger, P. Gauduchon & E. Mazet, *Le spectre d'une vari été riemannianne*, Springer Lecture Notes in Math. **194** (1971), MR 0282313, Zbl 0223.53034.

[Ch] I. Chavel, *Eigenvalues in Riemannian Geometry*, Vol. 115 of Pure and Applied Mathematics, Academic Press, 1984, MR 0768584, Zbl 0551.53001.

[DeFo] Y. Décanini & A. Folacci, *Off-diagonal coefficients of the DeWitt-Schwinger and Hadamard representations of the Feynman propagator,* Phys. Rev. D (3) **73** (2006), no. 4, 044027, 38 pp., MR 2214994.

[De] B. DeWitt, *The Global Approach to Quantum Field Theory*, Vol. 114, The International Series of Monographs on Physics, Oxford University Press, 2003, MR 1983836, Zbl 1044.81001.

[Don] S.K. Donaldson, *Scalar curvature and projective embeddings*, I. J. Differential Geom. **59** (2001), no. 3, 479–522, MR 1916953, Zbl 1052.32017.

[G1] M. Günther, *On the perturbation problem associated to isometric embeddings of Riemannian manifolds,* Ann. Global Anal. Geom. **7**, (1989) no. 1, 69-77, MR 1029846, Zbl 0691.53006.

[G2] M. Günther, *Isometric embeddings of Riemannian manifolds*, Proceedings of the International Congress of Mathematicians, Kyoto, Japan, 1990, MR 1159298, Zbl 0745.53031.

[Gil] P. Gilkey, *Asymptotic Formulae in Spectral Geometry* (Studies in Advanced Mathematics), Chapman and Hall/CRC; 1st edition (December 17, 2003), MR 2040963, Zbl 1080.58023.

[Gr1] M.L. Gromov, *Partial differential relations*, (1986), Springer-Verlag, Berlin, MR 0864505, Zbl 0651.53001.

[Gr2] M.L. Gromov, *Paul Levi's Isoperimeter inequality*, preprint, I.H.E.S., 1980.

[H1] L. Hörmander, *The spectral function of an elliptic operator*, Acta Math. **88** (1968), 341–370, MR 0609014, Zbl 0164.13201.

[H2] L. Hörmander, *The Analysis of Linear Partial Differential Operators III*, Springer Verlag, 1994, MR 2304165, Zbl 1115.35005.

[HH] Q. Han & J.X. Hong, *Isometric Embedding of Riemannian Manifolds in Euclidean Spaces,* (2006), American Mathematical Society, MR 2261749, Zbl 1113.53002.

[P] E. Potash, *Euclidean embeddings and Riemannian Bergman metrics*, preprint, arXiv:1310.4878.

[Po] J. Portegies, *Embeddings of Riemannian manifolds with heat kernels and eigenfunctions*, preprint, arXiv: 1311.7568.

[LY1] P. Li & S.T. Yau, *Estimates of eigenvalues of a compact Riemannian manifold,* Geometry of the Laplace operator, (Proc. Sympos. Pure Math., Univ. Hawaii, Honolulu, Hawaii, 1979), pp. 205–239, MR 0573435, Zbl 0441.58014.

[LY2] P. Li & S.T. Yau, *A New Conformal Invariant and Its Applications to the Willmore Conjecture and the First Eigenvalue of Compact Surfaces,* Invent. math. **69** (1982), 269–291, MR 0674407, Zbl 0503.53042.

[MS] D. McDuff & D. Salamon, *J-holomorphic Curves and Symplectic Topology*, Colloquim Publications, vol 52, AMS, Providence RI, 2004, MR 2954391, Zbl 1272.53002.

[N1] Nash, John, *C^1-isometric imbeddings*, Ann. of Math. **60** (1954) no. 3, 383–396, MR 0065993, Zbl 0058.37703.

[N2] Nash, John, *The imbedding problem for Riemannian manifolds*, Ann. of Math. **63** (1956) no. 1, 20–63, MR 0075639, Zbl 0070.38603.

[Ni] L.I. Nicolaescu, *Complexity of random smooth functions on compact manifolds.II*, preprint, arXiv:1209.0639.

[R] F. Rellich, *Störungstheorie der Spektralzerlegung. V*, Math. Ann. **118** (1942), 462–484, MR 0010791, JFM 68.0243.02.

[S] C.D. Sogge, *Concerning the L^p norm of spectral clusters for second-order elliptic operators on compact manifolds*, J. Funct. Anal. **77** (1988), no. 1, 123–134, MR 0930395, Zbl 0641.46011.

[SY] R. Schoen & S.-T. Yau, *Lectures on differential geometry*, Conference Proceedings and Lecture Notes in Geometry and Topology, I, International Press, Cambridge, MA, 1994, (Chinese version, 1988), MR 1333601, Zbl 0830.53001.

[SZ] B. Shiffman & S. Zelditch, *Asymptotics of almost holomorphic sections of ample line bundles on symplectic manifolds*, J. Reine Angew. Math. **544** (2002), 181–222, MR 1887895, Zbl 1007.53058.

[T] T. Sakai, *Riemannian geometry*, Vol. 149 of Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1996, MR 1390760, Zbl 0886.53002.

[Wa] X. Wang, *Riemannian moment map,* Comm. Anal. Geom. **16** (2008), no. 4, 837–863, MR 2471372, Zbl 1165.53058.

[WaZh] J. Wang & L. Zhou, *Gradient estimate for eigenforms of Hodge Laplacian*, Math. Res. Lett. **18** (2011), MR 2998141, Zbl 06194674.

[Wu] H. Wu, *Embedding Riemannian manifolds by the heat kernel of the connection Laplacian*, preprint, arXiv: 1305.4232.

[X] B. Xu, *Derivatives of the Spectral Function and Sobolev Norms of Eigenfunctions on a Closed Riemannian Manifold*, Ann. Global Anal. Geom. **26** (2004), no. 3, 231–252, MR 2097618, Zbl 1083.35072.

[Z] K. Zhu, *High-jet relations of the heat kernel embedding map and applications*, preprint, arXiv: 1308.0410.

Department of Mathematics & Computer Science
Rutgers University–Newark
Newark, NJ 07102
E-mail address: xiaowwan@rutgers.edu

Department of Mathematics
Harvard University
Cambridge, MA 02138
E-mail address: kzhu@math.harvard.edu