RECONSTRUCTION AND INTERPOLATION OF MANIFOLDS II: INVERSE PROBLEMS WITH PARTIAL DATA FOR DISTANCES OBSERVATIONS AND FOR THE HEAT KERNEL

CHARLES FEFFERMAN, SERGEI IVANOV, MATTI LASSAS,
JINPENG LU, HARIHARAN NARAYANAN

DEDICATED TO THE MEMORY OF YAROSLAV V. KURYLEV

ABSTRACT. We consider how a closed Riemannian manifold M and its metric tensor g can be approximately reconstructed from local distance measurements. Moreover, we consider an inverse problem of determining (M, g) from limited knowledge on the heat kernel. In the part 1 of the paper, we considered the approximate construction of a smooth manifold in the case when one is given the noisy distances $\widetilde{d}(x, y) = d(x, y) + \varepsilon_{x,y}$ for all points $x, y \in X$, where X is a δ-dense subset of M and $|\varepsilon_{x,y}| < \delta$. In this part 2 of the paper, we consider a similar problem with partial data, that is, the approximate construction of the manifold (M, g) when we are given $\widetilde{d}(x, y)$ for $x \in X$ and $y \in U \cap X$, where U is an open subset of M. In addition, we consider the inverse problem of determining the manifold (M, g) with non-negative Ricci curvature from noisy observations of the heat kernel $G(y, z, t)$. We show that a manifold approximating (M, g) can be determined in a stable way, when for some unknown source points z_j in $X \setminus U$, we are given the values of the heat kernel $G(y, z_k, t)$ for $y \in X \cap U$ and $t \in (0, 1)$ with a multiplicative noise. We also give a uniqueness result for the inverse problem in the case when the data does not contain noise and consider applications in manifold learning. A novel feature of the inverse problem for the heat kernel is that the set $M \setminus U$ containing the sources and the observation set U are disjoint.

1. Introduction

Let (M, g) be a closed connected Riemannian manifold of dimension $n \geq 2$. We consider (M, g) in the following class of Riemannian manifolds with bounded geometry given by

$$\mathrm{diam}(M) \leq \Lambda, \quad \mathrm{inj}(M) \geq \Lambda^{-1}, \quad |\mathrm{Sec}_M| \leq \Lambda^2, \tag{1.1}$$

where $\Lambda \geq 1$, $\mathrm{diam}(M)$ denotes the diameter of M, $\mathrm{inj}(M)$ denotes the injectivity radius of M, and Sec_M denotes the sectional curvature of M.

Let $U \subset M$ be an open subset, and assume that U contains an open ball $B(x_0, R)$ of radius $R > \Lambda^{-1}$ centered at some $x_0 \in U$. We call the subset U the measurement domain. We say that $Y \subset U$ is an ε-net in U if the ε neighborhood of Y in M contains the set U. We also say that a set Y is ε-dense in U when it is an ε-net in U.

The goal of this paper is to show that when Y is an ε-net in U and X is an ε-net in M, then the approximate distances $d(x, y)$ between the points $x \in X$ and $y \in Y$ determine an approximation of the whole manifold M. The part 1 of the paper, [29], considers the case when the observation domain U is the whole manifold M.

1.0.1. *Formulation of data for the inverse problem: distance vector data.* Let $\varepsilon_0 > 0$ be a small parameter and let Y be a finite ε_0-net in U,

$$Y = \{y_j \in U : j = 0, 1, \dots, J\}. \tag{1.2}$$

Assume $y_0 \in Y$ is such that $d(x_0, y_0) < \varepsilon_0$, where $d(x, y) = \mathrm{dist}_M(x, y)$ is the distance between points $x, y \in M$.

Key words and phrases. Inverse problems, Riemannian manifolds, geodesic distances, heat kernel.

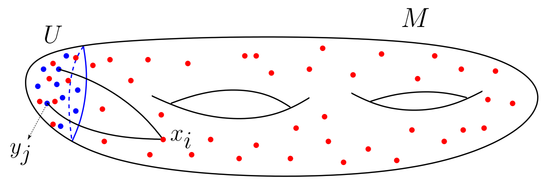

Figure 1. *On a closed manifold (M,g), we consider the distances $d(x_i,y_j)$ from the blue points $y_j \in Y$ in the open subset $U \subset M$ to the red points x_i filling the manifold M. These distances with measurement errors $\varepsilon_{i,j}$ define the noisy distance vectors $\vec{R}_i = [\widehat{R}_{i,j}]$, where $\widehat{R}_{i,j} = d(x_i,y_j) + \varepsilon_{i,j}$ and $|\varepsilon_{i,j}| < \varepsilon_1$. The inverse problem is to construct an approximation of the manifold (M,g) from these data.*

The *distance vector data* consist of a finite set of vectors $\vec{R}_i \in \mathbb{R}^{J+1}$, $i = 1,2,\dots,I$, given by

$$\vec{R}_i = [\widehat{R}_{i,j}]_{j\in\{0,\dots,J\}} \in \mathbb{R}^{J+1}. \tag{1.3}$$

We assume that the vectors $\vec{R}_i$ satisfy the following conditions for some small parameter $\varepsilon_1 > 0$:

(a1) For any $i = 1,2,\dots,I$, there exists a point $x \in M$ such that

$$|\widehat{R}_{i,j} - d(x,y_j)| < \varepsilon_1, \quad \text{for all} \;\; j = 0,1,\dots,J.$$

(a2) For any $x \in M$, there is $i \in \{1,2,\dots,I\}$ such that

$$|\widehat{R}_{i,j} - d(x,y_j)| < \varepsilon_1, \quad \text{for all } j = 0,1,\dots,J.$$

1.0.2. *An alterernative formulation of data: approximate interior distance function data.* For $x \in M$, we consider the distance function $r_x^U : U \to \mathbb{R}$ defined by

$$r_x^U(y) = d(x,y), \quad y \in U. \tag{1.4}$$

Let $\varepsilon_0 > 0$ and the set Y given in (1.2) be a finite ε_0-net in U. Let

$$\mathcal{R}_Y(M) := \{r_x^U|_Y : \; x \in M\} \subset \ell^\infty(Y) = \mathbb{R}^{J+1} \tag{1.5}$$

be the set of the restrictions of distance functions r_x^U onto the finite subset Y. The motivation of these functions is that they are discretizations of the distance functions

$$\mathcal{R}_U(M) := \{r_x^U : \; x \in M\} \subset C(U) \tag{1.6}$$

defined on the open set U.

The *approximate interior distance function data* consist of the finite set Y, and a finite set of functions on Y given by

$$\widehat{\mathcal{R}}_Y := \{\widehat{r}_i : Y \to \mathbb{R} \,|\, i = 1,2,\dots,I\} \subset \mathbb{R}^{J+1}. \tag{1.7}$$

We assume that the family $\widehat{\mathcal{R}}_Y$ satisfies

$$d_H(\widehat{\mathcal{R}}_Y, \mathcal{R}_Y(M)) < \varepsilon_1 \tag{1.8}$$

for some small parameter $\varepsilon_1 > 0$, that is, $\widehat{\mathcal{R}}_Y$ is an approximation of the set $\mathcal{R}_Y(M)$. Here d_H stands for the Hausdorff distance on $\mathbb{R}^{J+1}$, see [12]. The following lemma motivates the conditions (a1)-(a2).

Lemma 1.1. *Let $U \subset M$ be an open subset and Y be a finite ε_0-net in U. Then the approximate interior distance function data $\{Y, \widehat{\mathcal{R}}_Y\}$ satisfy the condition* (1.8) *if and only if the distance vector data $\vec{R}_i = [\widehat{R}_{i,j}]$, $i = 1, 2, \dots, I$, defined by $\widehat{R}_{i,j} = \widehat{r}_i(y_j)$, satisfy the conditions (a1) and (a2).*

The proof of Lemma 1.1 in given in Section 3.

1.1. **Main result.** Our main result is a global result on the determination of a smooth manifold (M, g) from the distance vector data consisting of the noisy distances of the points $x_i \in M$ to the points y_j in an open set $U \subset M$ (see Figure 1).

Theorem 1.2. *Let $n \in \mathbb{Z}_+$, $n \geq 2$, $\Lambda \geq 1$, $R > \Lambda^{-1}$. Then there exist $\widehat{\varepsilon}_1 > 0$, $C_0, C_1, C_2 > 1$ explicitly depending only on n, Λ, such that the following holds for all $0 < \varepsilon_1 < \widehat{\varepsilon}_1$ and $0 < \varepsilon_0 \leq \varepsilon_1$.*

Let (M, g) be a closed Riemannian manifold satisfying the bounds (1.1) *and $U \subset M$ be an open subset containing a ball $B(x_0, R)$. Let*

$$Y = \{y_j : j = 0, 1, \dots, J\} \subset U$$

be a finite ε_0-net in U, $d(x_0, y_0) < \varepsilon_0$, and $\varepsilon_2 = C_0 \varepsilon_1^{1/2}$.

Assume that we are given vectors $\vec{R}_i \in \mathbb{R}^{J+1}$, $i = 1, 2, \dots, I$, such that conditions (a1) and (a2) are valid with parameter ε_1. Then the following statements hold.

(1) *We can compute the numbers $\widehat{d}_{i,i'}$, $i, i' \in \{1, 2, \dots, I\}$ directly from the given data $\vec{R}_i \in \mathbb{R}^{J+1}$, $i = 1, 2, \dots, I$, such that there exists an ε_2-net $X = \{x_1, \dots, x_I\}$ in M for which*

$$\left|\widehat{d}_{i,i'} - d(x_i, x_{i'})\right| \leq C_1 \varepsilon_1^{\frac{1}{8}}, \quad \textit{for all } i, i' \in \{1, 2, \dots, I\}.$$

(2) *The given data $\vec{R}_i \in \mathbb{R}^{J+1}$, $i = 1, 2, \dots, I$, determine a smooth Riemannian manifold $(\widehat{M}, \widehat{g})$ that is diffeomorphic to M. Moreover, there is a diffeomorphism $F : \widehat{M} \to M$ such that*

$$\frac{1}{L} \leq \frac{d_M(F(x), F(x'))}{d_{\widehat{M}}(x, x')} \leq L, \quad \textit{for } x, x' \in \widehat{M},$$

where $L = 1 + C_2 \varepsilon_1^{1/12}$.

We will focus on proving the claim (1) in this paper, as the claim (2) essentially follows from the claim (1) and the part 1 of the paper, [29].

Remark 1.3. *In the form of the inverse problem formulated in Figure 1, Theorem 1.2 can be formulated in the following way. Let (M, g) be a closed Riemannian manifold satisfying the bounds* (1.1) *and $U \subset M$ be an open subset containing a ball $B(x_0, R)$. Suppose we are given an ε_1-net $Y = \{y_j : j = 1, \dots, J\}$ in U and an ε_2-net $X = \{x_i : i = 1, \dots, I\}$ in M, where $\varepsilon_2 = C_0 \varepsilon_1^{1/2}$. Then the noisy distance data $\widehat{R}_{i,j} = d(x_i, y_j) + \varepsilon_{i,j}$, where $|\varepsilon_{i,j}| < \varepsilon_1$, determine a smooth Riemannian manifold $(\widehat{M}, \widehat{g})$ that is diffeomorphic to M. Moreover, there is a bi-Lipschitz diffeomorphism $F : \widehat{M} \to M$ with Lipschitz constant $1 + C_2 \varepsilon_1^{1/12}$.*

2. Inverse problem for the heat kernel with partial data and the local reconstruction of the manifold

2.1. **Inverse problem for the heat kernel with noisy data.** Let $G(x, z, t)$ be the heat kernel of a Riemannian manifold (M, g), i.e., it satisfies

$$(\partial_t - \Delta_g) G(x, z, t) = 0, \quad \text{for } (x, t) \in M \times \mathbb{R}_+,$$
$$G(x, z, t)|_{t=0} = \delta_z(x),$$

where Δ_g is the Laplace-Beltrami operator on (M, g) that operates in the x-variable and δ_z is a point source at the point $z \in M$. Let

$$\widetilde{G}(x, z, t) = \eta(x, z, t)\, G(x, z, t) \tag{2.1}$$

be the values of the heat kernel with multiplicative noise $\eta(x, z, t)$ satisfying

$$\big|\log \eta(x, z, t)\big| \le \frac{\sigma}{t}, \quad \text{for } 0 < t < 1, \tag{2.2}$$

where $\sigma \in (0, 1)$ is small. We consider the stability of the following inverse problem on Riemannian manifolds with non-negative Ricci curvature.

Inverse problem for heat kernel with separated sources and observations. *Let (M, g) be a compact Riemannian manifold and $U \subset M$ be a non-empty open subset. Assume that we are given the set $(U, g|_U)$ as a Riemannian manifold and the heat kernel $G(y, z, t)$ at observation points $y \in U$ at all times $t \in (0, 1)$ with the source points $z \in M \setminus \overline{U}$. Do these data uniquely determine, the topology, the differentiable structure and the metric of the manifold (M, g)?*

In the case when the heat kernel $G(y, z, t)$ is sampled on the set $M \times M \times \mathbb{R}_+$, that is, sources and observations are on the whole manifold M, this problem in studied in the embedding of a manifold into an Euclidean space using the heat kernel [8, 71, 84], in diffusion distances in manifold learning [20, 21], and in manifold registration in shape theory [66, 67]. Also, an inverse problem where $G(y, z, t)$ is assumed to be known in the set $U \times U \times \mathbb{R}_+$ for an open subset $U \subset M$ was studied in [37, 38, 50], see also related studies [5, 7, 9, 24, 45, 77] for manifolds with boundary. In the methods used to study these problems it is essential that the set which contains the sources intersects the set where the solutions are observed. For partial data problems with separated sources and observations, the inverse problem for the wave equation on a non-trapping manifold is considered in [60]. Inverse problems for elliptic equations, where the sources and the observations are on sets that are small but intersect, have been studied under geometric convexity assumptions, see e.g. [25, 47, 51] and in the 2-dimensional case, see e.g. [36, 62, 63].

The following result proves that for manifolds with non-negative Ricci curvature, the Riemannian manifold structure depends in a stable way on the heat kernel with separated sources and observations when the data have multiplicative errors.

Theorem 2.1. *Let (M_1, g_1) and (M_2, g_2) be two closed Riemannian manifolds of dimension n satisfying the bounds (1.1) with parameter Λ, and let $U_l = B_{g_l}(y_0^l, R) \subset M_l$ be open balls of radius $R \ge \Lambda^{-1}$. Suppose the Ricci curvatures of the manifolds M_1 and M_2 are non-negative. Then there exist constants $\widehat{\sigma}, C_3 > 0$ explicitly depending only on n, Λ, such that the following holds for all $0 < \sigma < \widehat{\sigma}$ and $0 < h \le \sigma^{1/2}$.*

For $l = 1, 2$, let $\{z_i^l : i = 1, 2, \dots, I\}$ be an h-net in $M_l \setminus \overline{U}_l$, and $\{y_j^l : j = 0, 1, \dots, J\}$ be an h-net in the ball U_l. Suppose $\big|d_{M_1}(y_j^1, y_{j'}^1) - d_{M_2}(y_j^2, y_{j'}^2)\big| < h$ for all $j, j' = 0, 1, \dots, J$, and the heat kernels of M_1 and M_2 satisfy

$$e^{-\frac{\sigma}{t}} \le \frac{G_2(y_j^2, z_i^2, t)}{G_1(y_j^1, z_i^1, t)} \le e^{\frac{\sigma}{t}}, \tag{2.3}$$

for all $i = 1, \dots, I$, $j = 0, 1, \dots, J$ and $0 < t < 1$. Then M_1 and M_2 are diffeomorphic, and there is a diffeomorphism $F : M_1 \to M_2$ such that

$$\frac{1}{L} \le \frac{d_{M_2}(F(x), F(x'))}{d_{M_1}(x, x')} \le L, \quad \text{for } x, x' \in M_1,$$

where $L = 1 + C_3 \sigma^{1/24}$.

As a corollary we obtain the unique solvability of the inverse problem.

Corollary 2.2. *Let (M_1, g_1) and (M_2, g_2) be two closed Riemannian manifolds of dimension n with non-negative Ricci curvature. For $l = 1, 2$, let $U_l = B_{g_l}(y_0^l, R) \subset M_l$ be open balls of radius $R > 0$. If $\Phi : U_1 \to U_2$ is an isometry and $\Psi : M_1 \setminus U_1 \to M_2 \setminus U_2$ is a bijection such that*

$$G_1(y, z, t) = G_2(\Phi(y), \Psi(z), t), \quad \text{for all } y \in U_1,\ z \in M_1 \setminus U_1,\ 0 < t < 1,$$

then the Riemannian manifolds (M_1, g_1) and (M_2, g_2) are isometric.

We emphasize that in Corollary 2.2, the map $\Psi : M_1 \setminus U_1 \to M_2 \setminus U_2$ is only assumed to be a bijection and thus we do not a priori assume that M_1 and M_2 are homeomorphic.

Theorem 2.1 is proved by using the Cheeger-Yau asymptotics for heat kernel [17] that generalize Varadhan's classical formula [81].

2.2. **Reconstruction of local coordinates and metric tensor from partial distance data.** In this subsection, we consider how the local coordinates and the distances near a point $x_{i_0} \in M$ can be approximately constructed, using the approximate distance functions $\widehat{\mathcal{R}}_Y \subset \mathbb{R}^{J+1}$ that are in a neighborhood of the function $\widehat{r}_{i_0}$ corresponding to the point x_{i_0}. The studied question is closely related to a manifold learning problem where the local structure of the data set needs to be constructed from the distances to the marker points, see subsection 2.3.3.

Definition 2.3. *Consider the set $\widehat{\mathcal{R}}_Y$ satisfying* (1.8). *We say that $x_i \in M$ is a point corresponding to $\widehat{r}_i \in \widehat{\mathcal{R}}_Y$ if* (3.3) *holds. For each element $\widehat{r}_i$ in $\widehat{\mathcal{R}}_Y$, we choose one corresponding point $x_i \in M$ and denote the obtained set by X,*

$$X := \{x_i \in M \ : \ i = 1, 2, \dots, I\}. \tag{2.4}$$

Note that the set X of points are just known to exist, and they are not directly determined by the data $\widehat{\mathcal{R}}_Y$.

Let $U \subset M$ be an open subset and $Y = \{y_j : j = 0, 1, \dots, J\}$ be a finite ε_0-net in U. In the space of real-valued functions on Y, we denote the ℓ^∞-neighborhood of a function $\widehat{r} : Y \to \mathbb{R}$ by

$$\mathcal{B}_\infty(\widehat{r}, \rho) = \{f : Y \to \mathbb{R} \ : \ \|f - \widehat{r}\|_{\ell^\infty(Y)} < \rho\} \subset \mathbb{R}^{J+1}, \tag{2.5}$$

where $\rho > 0$ is the radius of the neighborhood. Suppose that we are given numbers $\widehat{d}^{Y}_{j,j'}$, $j, j' = 0, 1, \dots, J$ such that

$$\left|\widehat{d}^{Y}_{j,j'} - d(y_j, y_{j'})\right| \le 2\varepsilon_1, \quad \text{for all } j, j' = 0, 1, \dots, J. \tag{2.6}$$

For $x_{i_0} \in M$, the map $\exp_{x_{i_0}} : T_{x_{i_0}}M \to M$ is the Riemannian exponential map at x_{i_0}. Let $\{v_k\}_{k=1}^n$ be unit vectors that form a basis in $T_{x_{i_0}}M$, and $x_\ell \in B(x_{i_0}, r)$, $r < \mathrm{inj}(M)$. Then we say that

$$X(x_\ell) = (X_k(x_\ell))_{k=1}^n \in \mathbb{R}^n, \quad X_k(x_\ell) = \langle \exp_{x_{i_0}}^{-1}(x_\ell), v_k \rangle_g, \tag{2.7}$$

is the coordinate of the point x_ℓ in the Riemannian normal coordinates centered at the point x_{i_0}, associated to the (possibly non-orthogonal) basis $\{v_k\}_{k=1}^n$. Moreover,

$$g_{jk}(x_{i_0}) = \langle v_j, v_k \rangle_g, \quad j, k = 1, \dots, n, \tag{2.8}$$

are the components of the metric tensor in these Riemannian normal coordinates at the point x_{i_0}.

Theorem 2.4. *Let $n \in \mathbb{Z}_+$, $n \ge 2$, $\Lambda \ge 1$, $R > \Lambda^{-1}$. Then there exist $\widehat{\varepsilon}_1 > 0$, $\rho_0 > 0$, $c_1 > 0$ and $C_4 > 1$, explicitly depending only on n and Λ, such that the following holds for all $0 < \varepsilon_1 < \widehat{\varepsilon}_1$, $0 < \varepsilon_0 \le \varepsilon_1$.*

Let (M, g) be a closed Riemannian manifold satisfying the bounds (1.1) *with parameter Λ, and $U \subset M$ be an open subset containing a ball $B(x_0, R)$. Assume that Y is a finite ε_0-net*

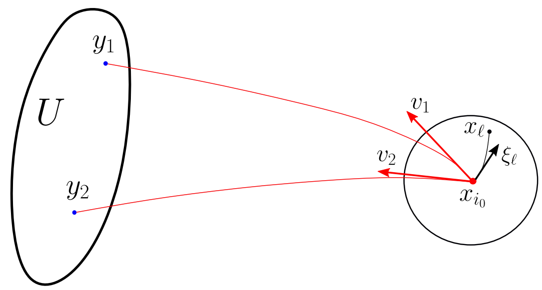

FIGURE 2. The Riemannian normal coordinates on the manifold M when $\dim(M) = 2$. These coordinates are centered at the point x_{i_0} and are associated to the basis $\{v_1, v_2\}$. We compute approximately the coordinates of the points $x_\ell = \exp_{x_{i_0}}(\xi_\ell)$, that is, $X(x_\ell) = (X_1(x_\ell), X_2(x_\ell)) \in \mathbb{R}^2$ and the metric tensor g_{jk} at the point x_{i_0}.

in U, $\widehat{d}^Y_{j,j'}$ are numbers satisfying (2.6), *and $\widehat{\mathcal{R}}_Y$ is an ε_1-approximation of $\mathcal{R}_Y(M)$ in the sense of* (1.8). *Let $x_i \in M$ be the points corresponding to $\widehat{r}_i \in \widehat{\mathcal{R}}_Y$ in the sense of Definition 2.3. Fix any $\widehat{r}_{i_0} \in \widehat{\mathcal{R}}_Y$. Then in the tangent space $T_{x_{i_0}}M$, there exist unit vectors $\{v_k\}_{k=1}^n$ satisfying $det([\langle v_j, v_k\rangle_g]_{j,k=1}^n) \geq c_1$, such that the following holds in the Riemannian normal coordinates centered at x_{i_0} and associated to the basis $\{v_k\}_{k=1}^n$.*

Assume that we are given Y, $\{\widehat{d}^Y_{j,j'}\}$, $\widehat{\mathcal{R}}_Y \cap \mathcal{B}_\infty(\widehat{r}_{i_0}, \varepsilon_1^{1/4})$, and any element $\widehat{r}_\ell \in \widehat{\mathcal{R}}_Y$ that satisfies

$$\widehat{r}_\ell \in \widehat{\mathcal{R}}_Y \cap \mathcal{B}_\infty(\widehat{r}_{i_0}, \rho_0), \tag{2.9}$$

see (2.5). *Then we can compute, directly from these data, numbers $\widehat{X}(x_\ell) \in \mathbb{R}^n$ and $\widehat{g}_{jk} \in \mathbb{R}$ such that*

$$\big|\widehat{X}(x_\ell) - X(x_\ell)\big| \leq C_4\big(d(x_\ell, x_{i_0})^{\frac{4}{3}} + \varepsilon_1\big), \tag{2.10}$$

where $X(x_\ell)$ is the coordinate of the point x_ℓ in the Riemannian normal coordinates given in (2.7), *and $\widehat{g}_{jk}$ satisfies*

$$\big|\widehat{g}_{jk} - g_{jk}(x_{i_0})\big| \leq C_4\varepsilon_1^{\frac{1}{8}}, \quad j,k = 1,\dots,n, \tag{2.11}$$

where $g_{jk}(x_{i_0})$ are the components of the metric tensor in the Riemannian normal coordinates at x_{i_0} given in (2.8).

The basis vectors $v_k \in T_{x_{i_0}}M$ in Theorem 2.4 are the directions to nearby points $x_{i(k)}$ such that the geodesics $\gamma_{x_{i_0},v_k}$ can be continued as distance-minimizing geodesics to points $y_{j(k)} \in U$, see Figure 2.

Theorem 2.4 has the following corollary.

Corollary 2.5. *Let $n, \Lambda, R, \widehat{\varepsilon}_1, \rho_0, c_1$ be as in Theorem 2.4. Then there exists $C_5 > 1$ explicitly depending only on n, Λ, such that the following holds for all $0 < \varepsilon_1 < \widehat{\varepsilon}_1$, $0 < \varepsilon_0 \leq \varepsilon_1$.*

Let $(M, g), U, Y, \widehat{\mathcal{R}}_Y$, $\widehat{r}_{i_0} \in \widehat{\mathcal{R}}_Y$ and $\widehat{r}_\ell \in \widehat{\mathcal{R}}_Y \cap \mathcal{B}_\infty(\widehat{r}_{i_0}, \rho_0)$ be as in Theorem 2.4. Then we can compute a number $\widehat{d}_{\ell,i_0}$ directly from the given data Y, $\{\widehat{d}^Y_{j,j'}\}$, $\widehat{r}_\ell$, and $\widehat{\mathcal{R}}_Y \cap \mathcal{B}_\infty(\widehat{r}_{i_0}, \varepsilon_1^{1/4})$ such that

$$\big|\widehat{d}_{\ell,i_0} - d(x_\ell, x_{i_0})\big| \leq C_5\big(d(x_\ell, x_{i_0})^{\frac{4}{3}} + \varepsilon_1^{\frac{1}{2}}\big), \tag{2.12}$$

where $x_{i_0}, x_\ell \in M$ *are the points in* X *corresponding to* $\widehat{r}_{i_0}, \widehat{r}_\ell$ *in the sense of Definition 2.3. More explicitly, the number* $\widehat{d}_{\ell,i_0}$ *can be computed as*

$$\widehat{d}_{\ell,i_0} := \Big(\sum_{j,k=1}^{n} \widehat{g}^{jk} \widehat{X}_j(x_\ell) \widehat{X}_k(x_\ell) \Big)^{\frac{1}{2}}, \tag{2.13}$$

where $\widehat{X}(x_\ell) = \big(\widehat{X}_k(x_\ell)\big)_{k=1}^n$ *and the inverse* $(\widehat{g}^{jk})$ *of the matrix* $(\widehat{g}_{jk})$ *are determined in Theorem 2.4.*

Theorem 2.4 and Corollary 2.5 show that in a neighborhood of $x_i \in X$, we can approximately find the local coordinates of nearby points $x_\ell \in X$ and the distances $d(x_\ell, x_i)$. We use this result to reconstruct distances in a finite net of M to prove Theorem 1.2(1). Similar results for manifolds with boundary have been studied in [46], assuming bounded derivative of the curvature tensor and using techniques that do not give explicit dependency on the geometric bounds. We emphasize that in Theorem 2.4 and Corollary 2.5 all the estimates can be made completely explicit in terms of Λ and n.

The method developed in the present paper replies on Toponogov's theorem. Suppose we do measurements in a ball B having the center y_0 and radius R. Any point $x \in M \setminus B$ can be connected to y_0 by some distance-minimizing geodesic $[y_0x]$. If $y_1 \in [y_0x]$ has distance $R/2 < d(y_1, y_0) < R$, then the geodesic segment $[y_1x]$ from y_1 to x is minimizing and has no cut points. Furthermore, one can perturb $[y_1x]$ slightly so that the perturbed geodesic has no cut points either, see Section 4 for a quantitative formulation. Technically, when we construct approximate values of the metric tensor, we handle this construction related to long geodesics by using only one-sided comparison estimates for large triangles, such as Toponogov's theorem, or by using comparison estimates for small triangles near the point x.

A situation considered in Theorem 2.4 is encountered in imaging applications, where the wave speed (the metric tensor g) needs to be reconstructed near a point x_i using the travel times of waves from nearby points x_ℓ to the points $y_j \in Y$. For example, consider the case when a measurement device located at the point y_j sends a wave at time $t = 0$ which reflects from a small scatterer (e.g. a detail in the material) at the point x_ℓ. If the reflected wave is observed at the point y_j at the time $t_{\ell,j}$, then the distance $d(x_\ell, y_j)$ is equal to $t_{\ell,j}/2$.

2.3. **Other applications and the stability of inverse problems under geometric a priori bounds.** The inverse problem of determining a Riemannian manifold (M, g) from the distance functions $r_x : U \to \mathbb{R}$, $r_x(y) = d(x, y)$, defined on an open subset $U \subset M$, is encountered in imaging problems that arise in geosciences, medical imaging and non-destructive testing. In these applications, the speed of waves defines a Riemannian metric on M so that the travel time of the waves from a point x to y is equal to the Riemannian distance $d(x, y)$. For instance, in the seismic imaging of the Earth, the inverse problem of finding the Riemannian metric in normal coordinates corresponds to finding the physical material parameters of the Earth in the travel time coordinates.

Geometric inverse problems have been studied on closed manifolds with data measured on an open subset of the manifold, see [37, 38, 50], as it is geometrically simpler to formulate the problems on a closed manifold than on a manifold with boundary. On the methods used to solve inverse problems for Riemannian manifold with boundary or a metric on it, see e.g. [5, 7, 44, 48, 49, 52, 56, 62, 63, 68, 77]. In many cases, it is also possible to reduce an inverse problem for a manifold with boundary to an inverse problem for a closed manifold, by means of extending the manifold with boundary to a closed manifold and extending measured boundary data to data on an open set.

2.3.1. *Inverse problems for linear equations.* Let us review a classical inverse problem that is related to the inverse problem studied in this paper.

1. Inverse interior spectral problem: Let (M, g) be a (unknown) Riemannian manifold and $U \subset M$ be an open set. Assume that we are given the following data,

$$\{U,\, (\lambda_j)_{j=1}^{\infty},\, (\phi_j|_U)_{j=1}^{\infty}\}. \tag{2.14}$$

Here λ_j are the eigenvalues of the Laplace-Beltrami operator Δ_g on M and ϕ_j are the corresponding orthonormal eigenfunctions. Do data (2.14) determine uniquely (up to an isometry) the Riemannian manifold (M, g)?

The methods used to solve the inverse problem *1* consist of two steps. First, the given data is used to construct the local distance function representation $\mathcal{R}_U(M)$ (for a detailed exposition, see e.g. [37, 45]). Second, the manifold (M, g) is reconstructed from $\mathcal{R}_U(M)$.

Analogous inverse problems for the wave equation and also for the Maxwell and Dirac systems are studied in [5, 6, 45, 53, 55]. The problem is closely related to inverse spectral problems where only eigenvalues are known, see [86]

2.3.2. *Inverse problems for non-linear equations.* The reconstruction of a manifold from partial distance measurements arises also in the study of the inverse problems for non-linear partial differential equations.

2. Inverse problem for a non-linear wave equation: Let (M, g) be a (unknown) Riemannian manifold and $U \subset M$ be an open set. Assume that we are given the source-to-solution map $L_U : C_0^\infty(U \times \mathbb{R}_+) \to C^\infty(U \times \mathbb{R}_+)$, $L_U(f) = v|_{U\times\mathbb{R}_+}$, where $f \in C_0^\infty(U \times \mathbb{R}_+)$ is a source and v is the solution of the following non-linear equation

$$\begin{aligned} &(\partial_t^2 - \Delta_g)v(x,t) + a(x,t)v(x,t)^2 = f(x,t), \quad \text{in } M \times \mathbb{R}_+, \\ &v|_{t=0} = 0, \quad \partial_t v|_{t=0} = 0, \end{aligned}$$

where $a(x, t) > 0$. Do the set U and the map L_U determine uniquely (up to an isometry) the Riemannian manifold (M, g) and the coefficient $a(x, t)$?

In the study of this problem, the non-linear interaction of linearized waves produced by suitable sources in $U \times \mathbb{R}_+$ can be used to produce "artificial" microlocal point sources at the points $y \in M$, including the unknown region $M \setminus U$ where the original source f vanishes, see [40, 54, 57, 59, 64, 80, 83]. The wave fronts that are produced by these point sources and are observed in U determine the distances $d(x, y)$ for the points $x \in M$ and $y \in U$. Thus the inverse problem *2* for the non-linear wave equation is reduced to the reconstruction of a manifold from partial distance measurements.

2.3.3. *Manifold learning.* In machine learning, an (invariant) manifold learning problem can be formulated as follows.

3. A manifold learning problem: Let (M, g) be a Riemannian manifold and $x_i \in M$, $i = 1, 2, \dots, I$, be an ε-dense set of sample points. We consider a small subset of these points, $x_1, \dots, x_J$, $J < I$, as the marker points. Assume that we are given the distances $d(x_i, x_j)$, $i = 1, 2, \dots, I$, $j = 1, 2, \dots, J$, between the sample points and the marker points. Can we obtain an approximation of the manifold (M, g) from these data?

In the case when the marker points $x_1, \dots, x_J$ belong in an open set $U \subset M$ and form a δ-dense set in U, this problem reduces to the problem studied in this paper. Machine learning problems analogous to the problem *3* were studied e.g. in [19, 20, 21, 30, 33, 72, 78, 87].

2.3.4. *A priori bounds and conditional stability of inverse problems.* In the inverse problems above, the problem of determining the metric from partial data measured on a subset is generally ill-posed in the sense of Hadamard: the map from the partial data to the metric is not continuous so that small change in the data can lead to huge errors in the reconstructed metric. One way out of this fundamental difficulty is to assume a priori bounds on the norms of the higher derivatives of coefficients. Results under this type of conditions are called conditional stability results and were known mostly for conformally Euclidean metric tensors (see e.g. [1, 2, 75, 76]). However, for inverse problems for general metric, this approach bears significant difficulties. The reason is that the usual C^m-norm bounds on coefficients are not invariant, and thus this type of conditions does not suit the invariance of the problems under diffeomorphisms. Moreover, if the structure of the manifold is not known a priori, this traditional approach cannot be used.

A natural way to overcome these difficulties is to impose a priori constraints in an invariant form and consider a class of manifolds that satisfy invariant a priori bounds, for instance on curvature, second fundamental form, injectivity radius, etc. Under such type of conditions, invariant stability results for various inverse problems have been proven in [5, 29, 75, 76]. In particular, for the inverse interior spectral problem on manifolds with non-trivial topology, stability results have been obtained in [5, 9], see also [14] for analogous results for manifolds with boundary. In the latter, it was shown that the convergence of the boundary spectral data implies the convergence of the manifolds with respect to the Gromov-Hausdorff distance. However, the stability for this problem is of $\log\log$ type.

In addition to being contaminated with errors, actual measurements typically provide only a finite set of data. An example of this is the classical Whitney problem on the extension of a function $f : X \to \mathbb{R}$, defined on $X \subset \mathbb{R}^n$, in an optimal way to a function $F \in C^m(\mathbb{R}^n)$, see [85]. This problem has been answered in the works of E. Bierstone, Y. Brudnyi, C. Fefferman, P. Milman, W. Pawluski, P. Shvartsman and others (see [10, 11, 26, 27, 28, 31]).

This paper is organized as follows. We review a few corollaries of Toponogov's theorem in Section 3. In Section 4, we derive the first variation type of estimates for almost minimizing paths as our main tool. Sections 5 and 6 are technical preparations for the reconstruction of local coordinates. Section 7 is devoted to proving the local result on the reconstruction of local coordinates and components of the metric tensor, that is, Theorem 2.4. We prove the global results, Theorem 1.2 and Theorem 2.1, in Sections 8 and 9.

Acknowledgement. We thank the anonymous referee for thoroughly reading our paper and many valuable comments. C. Fefferman was partially supported by AFOSR, grant DMS-1265524, and NSF, grant FA9550-12-1-0425. S. Ivanov was partially supported by RFBR, grants 14-01-00062 and 17-01-00128-A. M. Lassas and J. Lu were partially supported by AF, grants 284715 and 312110. H. Narayanan was partially supported by NSF grant DMS-1620102 and a Ramanujan Fellowship.

3. Preliminary constructions

3.1. **Notations.** Let (M, g) be a closed (that is, compact without boundary) connected Riemannian manifold of dimension $n \geq 2$ satisfying the bounds (1.1) with parameter Λ. Let $U \subset M$ be an open subset containing a ball $B(x_0, R)$ of radius $R > \Lambda^{-1}$ centered at $x_0 \in U$. By e.g. [15, Thm IX.6.1], if $r < \min\{\text{inj}(M)/2, \pi/(2K^{1/2})\}$ where K is the upper bound for the absolute value of sectional curvatures, the (open) metric balls $B(x, r)$ of (M, g) having radius r and center x are convex. Thus by making the ball $B(x_0, R) \subset U$ smaller, we can assume that the ball $B(x_0, R)$ is geodesically convex. We denote

$$\Sigma_r := \partial B(x_0, r).$$

Pairs (x, v), (y, u), etc. stand for points in the tangent bundle TM with v, u, etc. being tangent vectors. We identify the vector space $T_v(T_xM)$ with T_xM and $T_{(x,v)}(TM)$ with $T_xM \times T_xM$, denoting by $(u, w) \in T_xM \times T_xM$ a tangent vector in $T_{(x,v)}(TM)$. We denote by $\gamma_{x,v}(t) = \exp_x(tv)$ the geodesic emanating from x in the direction $v \in S_xM = \{v \in T_xM : |v|_g = 1\}$. Geodesics as well as other rectifiable curves are parametrized by the arclength.

Let $Y=\{y_j\}_{j=0}^J$ be a finite ε_0-net of U. Suppose we are given finite number of data

$$\widehat{\mathcal{R}}_Y := \{\widehat{r}_i : Y \to \mathbb{R} \,|\, i=1,\dots,I\}, \tag{3.1}$$

such that $\widehat{\mathcal{R}}_Y$ is an ε_1-approximation of $\mathcal{R}_Y(M)$ in the sense of (1.8). The given data essentially consists of $I(J+1)$ numbers. To shorten notations, we sometimes denote

$$\widehat{r}_{ij} := \widehat{r}_i(y_j), \quad y_j \in Y. \tag{3.2}$$

Next, using the above notations, we prove Lemma 1.1

Proof. (of Lemma 1.1) By the definition of the Hausdorff distance on $\mathbb{R}^{J+1}$, the condition (1.8) holds if and only if the following two conditions are satisfied.
(i) For any $\widehat{r}_i \in \widehat{\mathcal{R}}_Y$, there exists a point $x_i \in M$ such that

$$|\widehat{r}_i(y_j) - r^U_{x_i}(y_j)| < \varepsilon_1, \quad \text{for all } j=0,1,\dots,J. \tag{3.3}$$

(ii) For any $x\in M$, there is $i\in\{1,2,\dots,I\}$ such that

$$|\widehat{r}_i(y_j) - r^U_{x}(y_j)| < \varepsilon_1, \quad \text{for all } j=0,1,\dots,J.$$

Since we define $\widehat{R}_{i,j} = \widehat{r}_i(y_j)$, the conditions (i),(ii) are the same as the conditions (a1),(a2). □

In the proof of Theorem 1.2, we need to approximately determine the distances of the points in Y. To do that, we define approximate distances for points in Y as follows:

$$D^a_Y(y_j,y_k) = \inf_{i\in I}(\widehat{r}_{ij} + \widehat{r}_{ik}), \quad y_j,y_k \in Y. \tag{3.4}$$

Then by the triangular inequality and (3.3), we see that

$$D^a_Y(y_j,y_k) \ge d(y_j,y_k) - 2\varepsilon_1. \tag{3.5}$$

Let x be a point on the shortest geodesic from y_j to y_k. By (1.8), there is i such that $\|\widehat{r}_i - r^U_x|_Y\|_{\ell^\infty(Y)} < \varepsilon_1$. Then

$$D^a_Y(y_j,y_k) \le \widehat{r}_{ij} + \widehat{r}_{ik} \le d(y_j,x) + d(x,y_k) + 2\varepsilon_1.$$

Thus we see that

$$D^a_Y(y_j,y_k) \le d(y_j,y_k) + 2\varepsilon_1. \tag{3.6}$$

The above yields that

$$|D^a_Y(y_j,y_k) - d(y_j,y_k)| \le 2\varepsilon_1. \tag{3.7}$$

Note that (3.7) is independent of Y being a net of U. This shows we can find, up to an error of $2\varepsilon_1$, the distances between points in Y using only the given data $\widehat{\mathcal{R}}_Y$.

By (3.7), the numbers $D^a_Y(y_j,y_k)$ satisfy the inequality (2.6) that we required for the approximate distances $\widehat{d}^Y_{j,k}$ in the set Y. Thus, instead of using the notation $\widehat{d}^Y_{j,k}$ in the proof of Theorem 2.4, we identify these two notations and denote below

$$\widehat{d}^Y_{j,k} = D^a_Y(y_j,y_k). \tag{3.8}$$

In this paper, $C_1,C_2,\dots \in [1,\infty)$ and $c_1,c_2,\dots\in(0,1)$ denote uniform constants that explicitly depend only on n and Λ, unless specified. We also use a generic uniform constant $C>0$ that denotes a number that explicitly depends only on n,Λ, unless specified, but its exact value can be different in each appearance even inside one single formula.

3.2. **Implications of Toponogov's theorem.** To begin with, we introduce some notations that we will frequently use. We denote by $[ab]$ a minimizing geodesic (i.e. distance-minimizing curve) connecting the points a and b, and let $|ab| = d(a,b)$ denote the distance between the points a, b. Let β be the angle between the geodesics $[ab]$ and $[bc]$ at point b and $\theta = \pi - \beta$.

Let H be the rescaled hyperbolic plane with the constant sectional curvature $-\Lambda^2$, and $\overline{d}(\overline{a}, \overline{b})$ denotes the distance between the points $\overline{a}$ and $\overline{b}$ in H. Denote by $[\overline{a}\,\overline{b}]$ a minimizing geodesic connecting the points $\overline{a}$ and $\overline{b}$. For our considerations, we usually take $\overline{a}, \overline{b}$ in the following way. Let $\overline{a}, \overline{b}$, and $\overline{c}$ be points of H such that $\overline{d}(\overline{a}, \overline{b}) = d(a,b)$, $\overline{d}(\overline{b}, \overline{c}) = d(b,c)$ and the angle between the geodesics $[\overline{a}\,\overline{b}]$ and $[\overline{b}\,\overline{c}]$ at $\overline{b}$ is β. Then by Toponogov's theorem, the above triangle abc on M and the corresponding triangle $\overline{a}\overline{b}\overline{c}$ on H satisfy $d(a,c) \le \overline{d}(\overline{a}, \overline{c})$.

Now we present a corollary of Toponogov's theorem (e.g. [69, Thm. 79]). The analogous results to the first variation inequality (3.9) considered below are well-known in Alexandrov geometry. Similar types of formulae are used in Section 4.5 of [12] or Section 4 of [74] or Section 3.6 of [70]. However, we present the results in the form needed later and give the proof for the convenience of the reader.

Lemma 3.1. *There exist uniform constants $C_6, C_7 > 1$ such that the following holds.*

Let M be a closed Riemannian manifold with sectional curvature bounded below by $\mathrm{Sec}_M \ge -\Lambda^2$. Let $a, b, c \in M$ and β be the angle of the distance-minimizing geodesics $[ab]$ and $[bc]$ at b.

(i) Then

$$|ac| \le |ab| - |bc|\cos\beta + C_6 |bc|^2 / \min\{\Lambda^{-1}, |ab|\}. \tag{3.9}$$

(ii) In addition to the assumptions above, assume that $|ab| = |bc|$, $|ab| \le \Lambda$. Then

$$|ac| \le 2|ab|(1 - C_7\theta^2), \quad \textit{where } \theta = \pi - \beta. \tag{3.10}$$

Proof. (i) See Lemma A.2.

(ii) Denote $|ab| = |bc| = A \le \Lambda$ so that also $\overline{d}(\overline{a}, \overline{b}) = d(\overline{a}, \overline{b}) = A$. Let $B = d(a,c)$ and $\overline{B} = d(\overline{a}, \overline{c}) \le 2\Lambda$. Then by Toponogov's theorem $B \le \overline{B}$. Moreover, using the law of cosines (A.5), we can estimate B as follows: We have

$$\begin{aligned} \cosh(\Lambda B) &\le \cosh(\Lambda \overline{B}) = \cosh^2(\Lambda A) - \sinh^2(\Lambda A)\cos(\beta) \\ &\le 1 + \sinh^2(\Lambda A)(1 - \cos(\beta)) \end{aligned}$$

and as $\cosh(2t) = 1 + 2\sinh^2 t$, or

$$\cosh(\Lambda B) = 1 + 2\sinh^2(\frac{1}{2}\Lambda B),$$

we have,

$$2\sinh^2(\frac{1}{2}\Lambda B) \quad \le \quad \sinh^2(\Lambda A)(1 - \cos(\beta)). \tag{3.11}$$

Using the fact that there exists a uniform constant $C > 1$ so that

$$\frac{u}{w} - \frac{\sinh(u)}{\sinh(w)} \le C\left(1 - \frac{u}{w}\right), \quad \text{for all } u, w \in (0, \Lambda^2],\ u \le w,$$

we see using (3.11) that

$$\frac{B}{2A} \quad \le \quad \frac{1}{C+1}(C + \sqrt{\frac{(1-\cos(\beta))}{2}}). \tag{3.12}$$

Let $\theta = \pi - \beta$. Using (3.12), we see that there exists a uniform constant $C_7 > 1$ such that

$$\frac{B}{2A} \quad \le \quad 1 - \frac{1 - \sin(\beta/2)}{C+1} = 1 - \frac{1 - \cos(\theta/2)}{C+1} \le 1 - C_7\theta^2. \tag{3.13}$$

This proves (ii). □

The following lemma is a variation of Lemma 3.1 when $|ab|$ is small.

Lemma 3.2. *There exists a uniform constant $C_8 > 0$ such that the following holds.*

Let M be a closed Riemannian manifold with sectional curvature bounded below by $\mathrm{Sec}_M \geq -\Lambda^2$. Let $a, b, c \in M$ be such that $|ab| \geq |bc|$. Let β be the angle at b between (any pair of) shortest paths $[ab]$ and $[bc]$. If $|ab| \leq 1$ and $|bc| \leq \frac{1}{2}|ab|$, then

$$|ac| \leq |ab| - |bc|\cos\beta + C_8\beta\frac{|bc|^2}{|ab|}. \tag{3.14}$$

Proof. Let us first prove the claim in the case when $|ab| = 1$, $|bc| \leq 1/2$.

As above, let H be a rescaled hyperbolic plane of curvature $-\Lambda^2$ and $\overline{a}, \overline{b}, \overline{c} \in H$ be such that $|\overline{a}\overline{b}| = |ab|$, $|\overline{b}\overline{c}| = |bc|$ and $\angle\overline{a}\overline{b}\overline{c} = \beta$. Here by $|xy|$, we denote the distance between points x and y in whatever space they belong. Then by Toponogov's theorem, $|ac| \leq |\overline{a}\overline{c}|$.

It remains to prove (3.14) for $\overline{a}, \overline{b}, \overline{c} \in H$ in place of $a, b, c \in M$ and under the assumption that $|\overline{a}\overline{b}| = 1$. Let e_1, e_2 be an orthonormal basic of $T_{\overline{b}}H$ such that e_1 is tangent to the geodesic segment $[\overline{b}\overline{a}]$, i.e., $\overline{a} = \exp_{\overline{b}}(e_1)$. Let $r_0 = \frac{1}{2}$. For every $r \in [-r_0, r_0]$ and $\eta \in [-\pi, \pi]$ define a point $\xi(r, \eta) \in H$ by

$$\xi(r, \eta) = \exp_{\overline{b}}(r\cos(\eta)e_1 + r\sin(\eta)e_2).$$

Note that $\overline{c} = \xi(|bc|, \beta)$. Define

$$f(r, \eta) = \overline{d}(\overline{a}, \xi(r, \eta)) - 1 + r\cos(\eta).$$

Clearly $\xi : [-r_0, r_0] \times [-\pi, \pi] \to H$ is a smooth map and its image does not cover $\overline{a}$. Therefore f is a smooth function. (It can be written explicitly using the cosine law of the hyperbolic plane.) Observe that $f(r, 0) = 0$ for all r, hence

$$\left|\frac{\partial^2}{\partial r^2}f(r, \eta)\right| \leq 2C_8|\eta|, \quad r \in [-r_0, r_0],$$

where

$$C_8 = \frac{1}{2}\max_{r,\eta}\left|\frac{\partial^3}{\partial\eta\partial r^2}f(r, \eta)\right|.$$

By the first variation formula we have

$$\frac{\partial}{\partial r}\overline{d}(\overline{a}, \xi(r, \eta))\Big|_{r=0} = -\cos\eta,$$

and therefore

$$\frac{\partial}{\partial r}f(r, \eta)\Big|_{r=0} = 0.$$

This and the above estimate on $\partial^2 f/\partial r^2$ imply that

$$|f(r, \eta)| \leq C_8|\eta|r^2$$

for all $\eta \in [-\pi, \pi]$ and $r \in [-r_0, r_0]$. Substituting $\eta = \beta$ and $r = |\overline{b}\overline{c}|$ yields that

$$|\overline{a}\overline{c}| - 1 + |\overline{b}\overline{c}|\cos\beta \leq C_8\beta|\overline{b}\overline{c}|^2$$

or, equivalently,

$$|\overline{a}\overline{c}| \leq |\overline{a}\overline{b}| - |\overline{b}\overline{c}|\cos\beta + C_8\beta|\overline{b}\overline{c}|^2.$$

As explained above, the the claim follows from this inequality, Toponogov's theorem and the triangle inequality in M.

Thus we have proven the claim in the case when $|ab| = 1$. For the case when $|ab| < 1$ we can scale the metric g with the constant factor $|ab|^{-2}$. Observe that after this scaling the curvature is still bounded from below by $-\Lambda^2$. As the inequality (3.14) is invariant under metric scaling by a constant factor, we obtain the inequality (3.14) also on the case when $|ab| < 1$. □

4. First variation for almost minimizing paths

In this section, let M be a closed Riemannian manifold satisfying the bounds (1.1). We consider the first variation type of estimates for geodesics, with the help of corollaries of Toponogov's theorem in Section 3.2. To explain the idea of constructions we use, we first warm up by proving an improved version of the first variation formula for geodesics that can be continued as a minimizing geodesic.

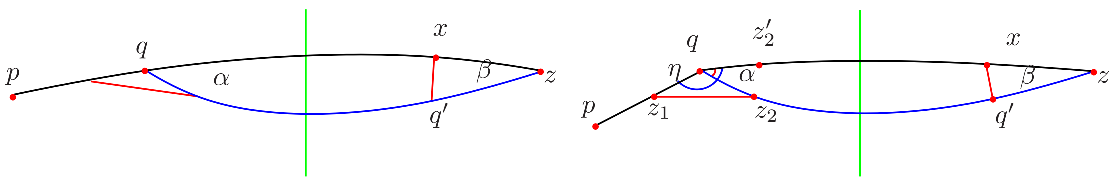

Figure 3. **Left:** *Setting of Lemma 4.1. The green vertical line corresponds to the boundary of the set U when the lemma is applied.* **Right:** *Setting of Lemma 4.2.*

Lemma 4.1. *Let M be a closed Riemannian manifold satisfying the bounds* (1.1). *Let $\gamma_{q,v}([0,\ell])$, $\ell = d(q,x) > \Lambda^{-1}$ be a distance-minimizing geodesic, parametrized by arc length, connecting $q, x \in M$. Assume that $p = \gamma_{q,v}(-\tau)$ and $\gamma_{q,v}([-\tau,\ell])$ is a distance-minimizing geodesic connecting p and x, where $\tau > \Lambda^{-1}/2$. Let $z = \gamma_{q,v}(\ell + r)$ be the point on the continuation of the geodesic $\gamma_{q,v}$, and α be the angle of $\gamma_{q,v}$ and $[qz]$ at q. Then there are uniform constants c_2, C_9 such that for all $0 < r < c_2$, we have*

$$\alpha \le C_9 r. \tag{4.1}$$

Remark. Lemma 4.1 can also be formulated in a scaling-invariant form. For example, assume that $\tau < \ell/2$ and $r < \ell/4$. Then we have

$$\alpha \le C_9 \frac{r}{\tau}. \tag{4.2}$$

This is also the case for the other two lemmas in this section. We note that an estimate of $\alpha \le C r^{1/2}$ can be obtained if one only assumes the lower sectional curvature bound.

Proof. Without loss of generality, let us assume $\tau > 1/2$, $\ell > 1$, since this only changes the constants by a factor of Λ due to scaling. Let β be the angle of $\gamma_{q,v}$ and $[zq]$ at z and denote by α the angle of $[qx]$ and $[qz]$ at q, that is, $\alpha = \angle xqz$. Let q' be a point on $[zq]$ such that $|zq'| = 2r$. Lemma 3.2, applied for the triangle xzq', implies that for $r < \min\{1/4, \mathrm{inj}(M)/4\}$,

$$|xq'| \le |zq'| - r\cos\beta + C_8 \beta r.$$

Let $z_1 \in [qp]$ and $z_2 \in [qz]$ be points such that $|z_1 q| = 1/4$ and $|z_2 q| = 1/4$. Then using a shortcut argument near the point q, see Fig. 3(Left) where the shortcut is the red segment, we can compare the distances of $|z_1 q| + |q z_2|$ and $|z_1 z_2|$, and using (3.10), we see that

$$\begin{aligned} |px| &\le |pz_1| + |z_1 z_2| + |z_2 q'| + |q' x| \\ &\le |pz_1| + (|z_1 q| + |q z_2| - \frac{1}{2} C_7 \alpha^2) + |z_2 q'| + (|q' z| - r\cos\beta + C_8 r \beta). \end{aligned}$$

Here, $|pz_1| + |z_1 q| = |pq|$ and $|q z_2| + |z_2 q'| + |q' z| = |qz|$ due to $\tau > 1/2$, $\ell > 1$, $r < 1/4$. Thus

$$|px| \le |pq| + |qz| - \frac{1}{2} C_7 \alpha^2 - r\cos\beta + C_8 r \beta. \tag{4.3}$$

As $|px| = |pq| + |qx|$, this yields

$$|qx| \le |qz| - \frac{1}{2}C_7\alpha^2 - r\cos\beta + C_8 r\beta. \tag{4.4}$$

Using (4.4) and the fact that $|qz| \le |qx| + r$, we see that

$$\begin{aligned} |qx| &\le |qx| + (1-\cos\beta)r - \frac{1}{2}C_7\alpha^2 + C_8 r\beta \\ &\le |qx| + \frac{1}{2}\beta^2 r - \frac{1}{2}C_7\alpha^2 + C_8 r\beta. \end{aligned} \tag{4.5}$$

Note that (4.5) yields only $\alpha^2 \le Cr$. We need the following improvement to obtain the desired estimate.

Pick the points $x_1 \in [qx]$, $x_2 \in [qz]$ such that $|zx_1| = |zx_2| = \Lambda^{-1}/2 \le \mathrm{inj}(M)/2$. Lemma A.1 yields that

$$d(x_1, x_2) \le C_A(\alpha + r).$$

On the other hand, by the Rauch comparison theorem for $\mathrm{Sec}_M \le \Lambda^2$,

$$C_A(\alpha + r) \ge d(x_1, x_2) \ge C(n,\Lambda)\sin\frac{\beta}{2} \ge \frac{1}{4}C(n,\Lambda)\beta.$$

This shows that for suitable $C_{10} > 1$, we have

$$\beta \le C_{10}(\alpha + r). \tag{4.6}$$

If $\alpha \le r$, the claim is proven. Thus we may assume that $r \le \alpha$. Then (4.6) becomes

$$\beta \le 2C_{10}\alpha. \tag{4.7}$$

Hence by (4.5), we see that

$$\begin{aligned} 0 &\le \frac{1}{2}4C_{10}^2\alpha^2 r - \frac{1}{2}C_7\alpha^2 + C_8 r\beta \\ &\le \frac{1}{2}(4C_{10}^2 r - C_7)\alpha^2 + 2C_8C_{10}r\alpha. \end{aligned}$$

Assuming that $r < c_2$ where $c_2 \le C_7C_{10}^{-2}/8$ we have $C_7 - 4C_{10}^2 r > \frac{1}{2}C_7$, and hence

$$0 \le -\frac{1}{4}C_7\alpha^2 + 2C_8C_{10}r\alpha$$

or

$$\alpha \le 8C_7^{-1}C_8C_{10}r =: C_9 r. \tag{4.8}$$

□

We remark that if there are multiple distance-minimizing geodesics from q to z, then (4.1) remains valid for each geodesic.

Next we modify the assumptions of Lemma 4.1 by replacing the minimizing geodesic by an almost minimizing path.

Lemma 4.2. *Let M be a closed Riemannian manifold satisfying the bounds* (1.1). *Let $\gamma_{q,v}([0,\ell])$, $\ell = d(q,x) > \Lambda^{-1}$ be a distance-minimizing geodesic, parametrized by arc length, connecting $q, x \in M$. Assume that there is a curve from p to x that goes through q that is almost distance-minimizing in the sense that*

$$|pq| + |qx| \le |px| + \delta. \tag{4.9}$$

Assume $|pq| > \Lambda^{-1}/2$. Let $z = \gamma_{q,v}(\ell + r)$ be the point on the continuation of the geodesic $\gamma_{q,v}$, and α be the angle of $\gamma_{q,v}$ and $[qz]$ at q. Then there are uniform constants $c_2, C_{11}, C_{12} > 1$ such that for all $0 < r, \delta < c_2$, we have

$$\alpha \le C_{11}(r^2 + \delta)^{1/2}. \tag{4.10}$$

Moreover,

$$|qx| + |xz| \le |qz| + C_{12} r (r^2 + \delta)^{1/2}. \tag{4.11}$$

Proof. Without loss of generality, let us assume $|pq| > 1/2$, $|qx| > 1$, since this only changes the constants by a factor of Λ due to scaling. Let β be the angle of $\gamma_{q,v}$ and $[zq]$ at z and let η be the angle of the (distance-minimizing) geodesic segments $[qp]$ and $[qx]$ at q. Denote $\alpha = \angle xqz$. Let q' be a point on $[zq]$ such that $|zq'| = 2r$. Then Lemma 3.2, applied for the triangle xzq' implies that for $r < \min\{1/4, \mathrm{inj}(M)/4\}$,

$$|xq'| \le |zq'| - r\cos\beta + C_8 \beta r. \tag{4.12}$$

Let $z_1 \in [qp]$, $z_2 \in [qz]$, $z_2' \in [qx]$ be the points such that $|z_1 q| = 1/4$, $|z_2 q| = 1/4$, $|qz_2'| = 1/4$, see Figure 3(Right). Then (3.10) applied to the triangle $z_1 z z_2'$ gives

$$|z_1 z_2'| \le |z_1 q| + |q z_2'| - 2|z_1 q| C_7 (\pi - \eta)^2. \tag{4.13}$$

Then we use a shortcut argument as follows.

$$\begin{aligned} |px| &\le |pz_1| + |z_1 z_2'| + |z_2' x| \\ &\le |pz_1| + \big(|z_1 q| + |q z_2'| - 2|z_1 q| C_7 (\pi-\eta)^2\big) + |z_2' x| \\ &= |pq| + |qx| - \frac{1}{2} C_7 (\pi - \eta)^2. \end{aligned}$$

As $|px| \ge |pq| + |qx| - \delta$, this yields

$$(\pi - \eta)^2 \le 2 C_7^{-1} \delta. \tag{4.14}$$

We use a shortcut argument near the point q, see Fig. 3(Right) where the shortcut is the red segment $[z_1 z_2]$. We compare the distances of $|z_1 q| + |q z_2|$ and $|z_1 z_2|$ using (3.10), and see that

$$|z_1 z_2| \le |z_1 q| + |q z_2| - 2|z_2 q| C_7 (\pi - \omega)^2, \tag{4.15}$$

where ω denotes the angle between $[qp]$ and $[qz]$ at q.

Inequality (4.15) implies then that

$$\begin{aligned} |px| &\le |pz_1| + |z_1 z_2| + |z_2 q'| + |q' x| \\ &\le |pz_1| + \big(|z_1 q| + |q z_2| - \frac{1}{2} C_7 (\pi-\omega)^2\big) + |z_2 q'| + \\ &\quad + (|q' z| - r\cos\beta + C_8 r\beta). \end{aligned}$$

Here, $|pz_1| + |z_1 q| = |pq|$ and $|q z_2| + |z_2 q'| + |q' z| = |qz|$. Thus

$$|px| \le |pq| + |qz| - \frac{1}{2} C_7 (\pi - \omega)^2 - r\cos\beta + C_8 r\beta. \tag{4.16}$$

As $|px| \ge |pq| + |qx| - \delta$, this yields

$$|qx| - \delta \le |qz| - \frac{1}{2} C_7 (\pi - \omega)^2 - r\cos\beta + C_8 r\beta. \tag{4.17}$$

Using (4.17) and the fact that $|qz| \le |qx| + r$, we see that

$$|qx| - \delta \quad \le \quad |qx| + (1 - \cos\beta) r - \frac{1}{2} C_7 (\pi - \omega)^2 + C_8 r\beta. \tag{4.18}$$

Note that (4.18) already implies $(\pi - \omega)^2 \le C(r + \delta)$, which combining with (4.14) yields $\alpha \le |\pi - \eta| + |\pi - \omega| \le C(r + \delta)^{1/2}$. We still need the following improvement to obtain the desired estimate.

Similar as in Lemma 4.1, by using the Rauch comparison theorem in a ball of radius $\Lambda^{-1}/2$ centered at z, for sufficiently small r, δ, we have

$$\beta \le C_{10} (\alpha + r). \tag{4.19}$$

If $\alpha \le r$, the claim is proven. Thus we may assume that $r \le \alpha$. Then (4.19) becomes

$$\beta \le 2C_{10}\alpha. \tag{4.20}$$

Hence using (4.18) and $\alpha \le |\pi - \omega| + (2C_7^{-1}\delta)^{1/2}$ by (4.14), we see that

$$\begin{aligned} -\delta &\le \frac{1}{2}4C_{10}^2\alpha^2 r - \frac{1}{2}C_7(\pi-\omega)^2 + 2C_8C_{10}r\alpha \\ &\le \frac{1}{2}(8C_{10}^2 r - C_7)(\pi-\omega)^2 + 2C_8C_{10}r|\pi-\omega| + C_{10}' r\delta + C_{10}' r\delta^{\frac{1}{2}}, \end{aligned}$$

for some constant $C_{10}' > 0$ depending on C_7, C_8, C_{10}. Assuming that $r < c_2$ where $c_2 \le C_7C_{10}^{-2}/16$, we have $C_7 - 8C_{10}^2 r > \frac{1}{2}C_7$. Hence,

$$\frac{1}{4}C_7(\pi-\omega)^2 - 2C_8C_{10}r|\pi-\omega| - (1 + C_{10}'r)\delta - C_{10}'r\delta^{\frac{1}{2}} \le 0.$$

Therefore,

$$\begin{aligned} \alpha &\le |\pi-\omega| + 2C_7^{-\frac{1}{2}}\delta^{\frac{1}{2}} \\ &\le \frac{2C_8C_{10}r + \sqrt{(2C_8C_{10}r)^2 + C_7(1+C_{10}'r)\delta + C_7C_{10}'r\delta^{1/2}}}{C_7/2} + 2C_7^{-\frac{1}{2}}\delta^{\frac{1}{2}} \\ &\le C_{11}(r^2+\delta)^{1/2}, \end{aligned}$$

with some uniform constant $C_{11} > 1$. This proves the first claim (4.10) of the lemma.

For the second claim, let $q' \in M$ be a point on $[zq]$ such that $|zq'| = 2r$. Lemma 3.2, applied for the triangle xzq' and its angle at z implies that

$$|q'x| \le |q'z| - |xz|\cos\beta + C_8 r\beta. \tag{4.21}$$

Adding $|qq'|$ on both sides of (4.21) and using the facts that $|qx| \le |qq'| + |q'x|$ and $|qz| = |qq'| + |q'z|$, we obtain

$$\begin{aligned} |qx| &\le |qq'| + |q'x| \\ &\le (|qq'| + |q'z|) - |xz|\cos\beta + C_8 r\beta \\ &= |qz| - |xz|\cos\beta + C_8 r\beta. \end{aligned} \tag{4.22}$$

Using the triangle inequality, (4.19), (4.10), and the facts that $|1-\cos\beta| \le \frac{1}{2}\beta^2$ and $|xz| = r$, we see that

$$\begin{aligned} 0 \le |qx| + |xz| - |qz| &\le C_8\beta r + \frac{1}{2}r\beta^2 \\ &\le C_{12}r(r^2+\delta)^{1/2} \end{aligned} \tag{4.23}$$

with some suitable uniform constant $C_{12} > 1$. □

Proposition 4.3. *Let M be a closed Riemannian manifold satisfying the bounds* (1.1). *Then there is a uniform constant $\widehat{\delta} \in (0, c_2^2)$ such that the following holds for all $0 < \delta < \widehat{\delta}$.*

Let $p, q, x \in M$ such that $|pq| > \Lambda^{-1}/2$, $|qx| > \Lambda^{-1}$ and

$$|pq| + |qx| \le |px| + \delta. \tag{4.24}$$

Denote by θ the angle of $\gamma_{q,v}$ and $[xy]$ at x. Let $y \in M$ be such that $d(y,x) < c_2^2$. Then there is a uniform constant $C_{13} > 1$ such that

$$\Big| |xy|\cos\theta - (|xq| - |yq|) \Big| \le C_{13}(|xy|\delta^{1/4} + |xy|^{4/3}).$$

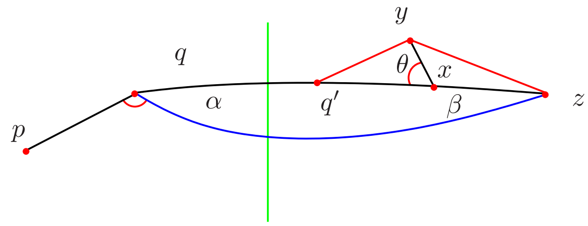

FIGURE 4. *An auxiliary figure for Proposition 4.3.*

Note that here $|xy|\cos\theta$ is the inner product of the vector $\xi \in T_xM$ for which $\exp_x \xi = y$ and the unit vector $v \in S_xM$ for which $\gamma_{x,v}$ coincides with the geodesic $[xq]$.

Proof. Let $v \in T_qM$ be the unit vector such that $[qx]$ is the path $\gamma_{q,v}([0,\ell])$, $\ell = |qx|$. Pick a number r such that $|xy| \le r < c_2$. Let $z = \gamma_{q,v}(\ell + r)$ and $q' = \gamma_{q,v}(\ell - r)$, so that $|xz| = r$ and $|xq'| = r$. Next we use Toponogov's theorem for the triangles $q'xy$ and zxy.

Below in this proof, we use the notation C denoting a generic constant whose exact value can change even inside one formula. The value value of each C can be computed as an explicit function of Λ and n. Consider the triangle $q'xy$, and we see from (3.9) that for $r < \Lambda^{-1}$,

$$|q'y| \le |xq'| - |xy|\cos\theta + C\frac{|xy|^2}{r},$$

where C is a uniform constant. Then, as $|qq'| + |xq'| = |qx|$,

$$|qy| \le |qq'| + |q'y| \le |qx| - |xy|\cos\theta + C\frac{|xy|^2}{r}. \tag{4.25}$$

Now considering triangle xyz, we see from (3.9) that

(4.26)
$$|yz| \le |xz| - |xy|\cos(\pi - \theta) + C\frac{|xy|^2}{r} = |xz| + |xy|\cos(\theta) + C\frac{|xy|^2}{r}.$$

By (4.11),

$$\begin{aligned} |qx| + |xz| &\le |qz| + Cr(r^2+\delta)^{1/2} \\ &\le |qy| + |yz| + Cr(r^2+\delta)^{1/2}. \end{aligned} \tag{4.27}$$

Now (4.26) and (4.27) yield

$$\begin{aligned} |qx| + |xz| &\le |qy| + |yz| + Cr(r^2+\delta^2)^{1/2} \\ &\le |qy| + (|xz| + |xy|\cos(\theta) + C\frac{|xy|^2}{r}) + Cr(r^2+\delta)^{1/2}, \end{aligned}$$

yielding that, after cancelation of $|xz|$,

$$|qx| \le |qy| + |xy|\cos(\theta) + C\frac{|xy|^2}{r} + Cr(r^2+\delta)^{1/2},$$

or

$$|qx| - |qy| \le |xy|\cos(\theta) + C\frac{|xy|^2}{r} + Cr(r^2+\delta)^{1/2}. \tag{4.28}$$

Comparing this with (4.25), we obtain

$$|xy|\cos\theta - C\frac{|xy|^2}{r} \le |qx| - |qy| \le |xy|\cos(\theta) + C\frac{|xy|^2}{r} + Cr(r^2+\delta)^{1/2}$$

and hence

$$\Big|(|qx|-|qy|)-|xy|\cos\theta\Big| \le C\frac{|xy|^2}{r} + Cr(r^2+\delta)^{1/2} =: E. \tag{4.29}$$

Now we optimize the value of r so that $|xy|^2/r \approx r(r^2+\delta)^{1/2}$ under the requirements that $|xy|\le r$ and $r<c_2$. First , we consider the case when $|xy|>\delta^{3/4}$. Then a good choice is $|xy|^2=r^3$, or $r=|xy|^{2/3}$ so that $|xy|<r$. Then with some uniform constant C, we have

$$E \le C\frac{|xy|^2}{|xy|^{2/3}} + C|xy|^{2/3}(2|xy|^{4/3})^{1/2} \le 4C|xy|^{4/3}.$$

As for the case when $|xy|\le\delta^{3/4}$, a good choice is $|xy|^2/r = r\delta^{1/2}$, or $r=|xy|\delta^{-1/4}$ so that $|xy| = r\delta^{1/4} < r$ and $r\le\delta^{1/2}$. Then we see that with some uniform constant C,

$$E \le C\frac{|xy|^2}{|xy|\delta^{-1/4}} + C(|xy|\delta^{-1/4})\delta^{1/2} = 2C|xy|\delta^{1/4}.$$

Note that the requirement $|xy|\le r$ is valid in both cases above. The requirement that $r<c_2$ is validated either by the condition $|xy|<c_2^2$ in the former case, or by the choice $\widehat{\delta}<c_2^2$ in the latter case. Thus using the above choices of r, we obtain the estimate

$$E \le C_{13}(|xy|\delta^{1/4} + |xy|^{4/3}) = C_{13}|xy|(\delta^{1/4}+|xy|^{1/3})$$

with some uniform constant $C_{13}>1$. □

5. Finding directions of minimizing paths

Let $U\subset M$ be an open subset of a closed Riemannian manifold M containing a ball $B(x_0,R)$, and $Y=\{y_j\}_{j=0}^J$ be an ε_0-net of U. Suppose we are given an ε_1-approximation $\widehat{\mathcal{R}}_Y$ of $\mathcal{R}_Y(M)$ in the sense of (1.8). By the definition of Hausdorff distance, for any $\widehat{r}_i\in\widehat{\mathcal{R}}_Y$ there exists $x_i\in M$ such that

$$|\widehat{r}_i(y_j)-d(x_i,y_j)|<\varepsilon_1, \quad \text{for all } y_j\in Y. \tag{5.1}$$

We choose for all $\widehat{r}_i\in\widehat{\mathcal{R}}_Y$ some point x_i satisfying (5.1), and call x_i a corresponding point to the approximate distance function $\widehat{r}_i$ (c.f. Definition 2.3). We denote by $X=\{x_i\}_{i=1}^I\subset M$ a set of points chosen so that each point x_i is a corresponding point to an approximate distance function $\widehat{r}_i\in\ell^\infty(Y)$. We also denote $\widehat{\mathcal{R}}_Y=\{\widehat{r}_i\}_{i=1}^I$.

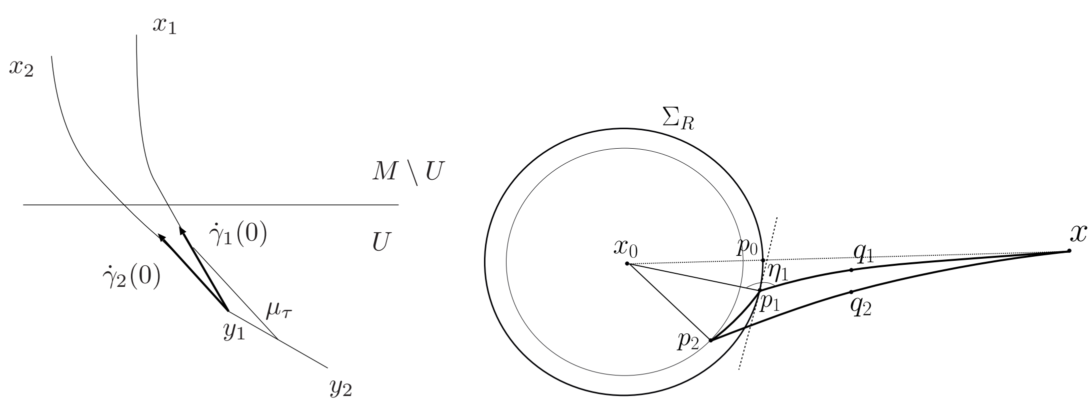

Figure 5. **Left:** *Distance-minimizing paths γ_i and a shortcut μ_τ in Lemma 5.1; the angles of μ and γ_i at y_1 are close to π.* **Right:** *Setting of Lemma 6.1.*

First, we prove a result that improves [46, Lemma 5.1]. Roughly speaking, the lemma states that we can identify the directions of the distance-minimizing path from $y_1 \in M$ to $x_1 \in M$ up to a small error, by considering the approximate distance function $\widehat{r}_1$ corresponding to the point x_1.

Lemma 5.1. *Let M be a closed Riemannian manifold with sectional curvature bounded below by $\mathrm{Sec}_M \geq -\Lambda^2$. Let $Y \subset U \subset M$ and $y_1, y_2 \in Y$. Suppose we are given an ε_1-approximation $\widehat{\mathcal{R}}_Y$ of $\mathcal{R}_Y(M)$ in the sense of (1.8). Let $\widehat{r}_i \in \widehat{\mathcal{R}}_Y$, $i = 1, 2$ and let $x_i \in M$ be the points corresponding to $\widehat{r}_i$ (c.f. Definition 2.3). Denote by $\gamma_i(t)$ some distance-minimizing path from y_1 to x_i, parametrized by arclength. Assume that*

$$d(y_1, y_2) \geq \Lambda^{-1}, \quad d(y_1, x_i) \geq \Lambda^{-1}, \quad i = 1, 2. \tag{5.2}$$

Suppose the following is true for some $\delta \in (0, 1]$ satisfying $\delta^2 \geq \varepsilon_1$:

$$|\widehat{r}_i(y_2) - \widehat{r}_i(y_1) - D_Y^a(y_1, y_2)| \leq \delta^2, \quad \textit{for } i = 1, 2. \tag{5.3}$$

Then there are uniform constants C_{14}, $C_{15} > 1$ such that

$$|\dot\gamma_1(0) - \dot\gamma_2(0)| \leq C_{14}\delta, \tag{5.4}$$

and

$$\Big| d(x_1, x_2) - |\widehat{r}_1(y_1) - \widehat{r}_2(y_1)| \Big| \leq C_{15}\delta. \tag{5.5}$$

Proof. Due to (5.1) and (3.7) and $\varepsilon_1 \leq \delta^2$, the condition (5.3) implies that

$$|d(x_i, y_2) - d(x_i, y_1) - d(y_1, y_2)| \leq 5\delta^2, \quad i = 1, 2. \tag{5.6}$$

Let $\mu(t)$ be a distance-minimizing geodesic of M from y_1 to y_2, i.e. $\mu(0) = y_1$, $\mu(d(y_1, y_2)) = y_2$. Denote by $\alpha_i > 0$ the angle between $\dot\gamma_i(0)$ and $\dot\mu(0)$, $i = 1, 2$. Now we show that α_i is close to π.

We restrict our attention to $i = 1$; the case of $i = 2$ follows in the same way. Let us use a shortcut argument. Pick $z_1 \in [y_1x_1]$, $z_2 \in [y_1y_2]$ such that $d(y_1, z_1) = d(y_1, z_2) = \Lambda^{-1}/2$. Applying Toponogov's theorem (3.10), we have

$$d(z_1, z_2) \leq d(z_1, y_1) + d(y_1, z_2) - \Lambda^{-1}C_7(\pi - \alpha_1)^2. \tag{5.7}$$

Then by the triangle inequality,

$$\begin{aligned} d(x_1, y_2) &\leq d(x_1, z_1) + d(z_1, z_2) + d(z_2, y_2) \\ &\leq d(x_1, z_1) + d(z_2, y_2) + d(z_1, y_1) + d(y_1, z_2) - \Lambda^{-1}C_7(\pi - \alpha_1)^2 \\ &= d(x_1, y_1) + d(y_1, y_2) - \Lambda^{-1}C_7(\pi - \alpha_1)^2. \end{aligned}$$

On the other hand, (5.6) gives

$$d(x_1, y_1) + d(y_1, y_2) \leq d(x_1, y_2) + 5\delta^2.$$

Hence,

$$|\pi - \alpha_i| \leq (5\Lambda C_7^{-1})^{\frac{1}{2}}\delta, \quad i = 1, 2. \tag{5.8}$$

Denote by α the angle between $\dot\gamma_1(0)$ and $\dot\gamma_2(0)$. Then for suitable $C_{14} > 1$,

$$|\dot\gamma_1(0) - \dot\gamma_2(0)| \leq \alpha \leq |\pi - \alpha_1| + |\pi - \alpha_2| \leq C_{14}\delta. \tag{5.9}$$

For the second claim, since $\gamma_i(d(x_i, y_1)) = x_i$, Lemma A.1 and (5.9) yield that

$$\begin{aligned} d(x_1, x_2) &\leq |d(x_1, y_1) - d(x_2, y_1)| + C_A\alpha \\ &\leq |d(x_1, y_1) - d(x_2, y_1)| + C_A C_{14}\delta. \end{aligned} \tag{5.10}$$

Then the second claim follows from (5.1) and the choice $\varepsilon_1 \leq \delta^2$. □

Observe that if there are several distance-minimizing paths from y_1 to x_i, the estimate (5.4) remains valid for each pair.

Let $\widehat{r}_{i_0}$ be given and let x_{i_0} be such that $\|r_{x_{i_0}} - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} < \varepsilon_1$. We consider the elements $\widehat{r}_\ell \in \widehat{\mathcal{R}}_Y$, $\ell = 1, 2, \dots, L$ for which

$$\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} \le \rho_0, \tag{5.11}$$

where ρ_0 is a sufficiently small uniform constant to be determined later.

We will frequently use the following notation

$$N_\epsilon(x_0; r) := B(x_0, r+\epsilon) \setminus B(x_0, r-\epsilon), \tag{5.12}$$

and we write $N_\epsilon(r)$ for short when the center is x_0.

Proposition 5.2. *Let M be a closed Riemannian manifold with sectional curvature bounded below by $\mathrm{Sec}_M \ge -\Lambda^2$, and $U \subset M$ be an open subset containing $B(x_0, R)$ with $R > \Lambda^{-1}$. Let Y be an ε_0-net of U, and $\widehat{\mathcal{R}}_Y$ be an ε_1-approximation of $\mathcal{R}_Y(M)$ in the sense of* (1.8)*. Then the following statements hold for $0 < \varepsilon_0 \le \varepsilon_1 < \min\{1/16, \Lambda^{-1}/32\}$.*

(1) *Let x_{i_0}, x_ℓ be the points corresponding to $\widehat{r}_{i_0}, \widehat{r}_\ell$ (c.f. Definition 2.3). Suppose $\widehat{r}_{i_0}, \widehat{r}_\ell$ satisfies the following conditions:*

$$\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} < \rho_0 < \min\{\frac{1}{4}, \frac{\Lambda^{-1}}{16}\}. \tag{5.13}$$

Assume we are given $y_0 \in Y$ such that $d(x_0, y_0) < \varepsilon_0$. Then there is a uniform constant $C_{15} > 1$ such that

$$d(x_\ell, x_{i_0}) \le 3C_{15}\big(\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} + 3\varepsilon_1\big)^{1/2}.$$

(2) *The set X is an ε_2-net of M, where $\varepsilon_2 = C_0 \varepsilon_1^{1/2}$ for some uniform constant $C_0 > 1$.*

Proof. (1) Let us keep the parameter R in the proof for clarity, and note that any dependency of R in the constants can be replaced by Λ using the condition $\Lambda^{-1} < R \le \Lambda$. We divide into two cases depending on where x_{i_0} lies.

- *Case 1: $\widehat{r}_{i_0}(y_0) > R/2$.*

Since $d(x_0, y_0) < \varepsilon_0 \le \varepsilon_1$, then $d(x_{i_0}, x_0) > R/2 - 2\varepsilon_1$. We take an arbitrary point $p \in N_{\varepsilon_1}(R/8) \cap Y$. The minimizing geodesic from x_{i_0} to p intersects with $\Sigma_{R/4}$ at some point q', and we take a point $q \in Y$ such that $d(q, q') < \varepsilon_0 \le \varepsilon_1$. As a consequence, $q \in N_{\varepsilon_1}(R/4) \cap Y$. Thus (5.13) and $R > \Lambda^{-1}$ yield

$$d(p, q) > \frac{R}{16}, \quad d(q, x_{i_0}) > \frac{R}{8}, \quad d(q, x_\ell) > \frac{R}{16}. \tag{5.14}$$

Moreover, by the triangle inequality,

$$d(p, x_{i_0}) = d(p, q') + d(q', x_{i_0}) \ge d(p, q) + d(q, x_{i_0}) - 2\varepsilon_1. \tag{5.15}$$

From (5.15), (5.1) and (3.7), we see that

$$|\widehat{r}_{i_0}(p) - \widehat{r}_{i_0}(q) - D_Y^a(p, q)| < 6\varepsilon_1. \tag{5.16}$$

Then pass to $\widehat{r}_\ell$,

$$|\widehat{r}_\ell(p) - \widehat{r}_\ell(q) - D_Y^a(p, q)| < 2\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} + 6\varepsilon_1. \tag{5.17}$$

Hence the assumptions of Lemma 5.1 are satisfied with

$$\delta^2 = 2\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} + 6\varepsilon_1,$$

which satisfies $\delta^2 > \varepsilon_1$ and $\delta < 1$ when $\rho_0 < 1/4$. Thus,

$$\begin{aligned} d(x_{i_0}, x_\ell) &\le |\widehat{r}_{i_0}(q) - \widehat{r}_\ell(q)| + C_{15}\big(2\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} + 6\varepsilon_1\big)^{1/2} \\ &\le 3C_{15}(\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} + 3\varepsilon_1)^{1/2}. \end{aligned}$$

- *Case 2: $\widehat{r}_{i_0}(y_0) \le R/2$.*

In this case, $d(x_{i_0}, x_0) \le R/2 + 2\varepsilon_1$. To keep distances bounded away from zero, one can choose points p, q from the outer layer $B(x_0, R) \setminus B(x_0, 3R/4)$. More precisely, we take an arbitrary point $p \in N_{\varepsilon_1}(R) \cap Y$. The minimizing geodesic from x_{i_0} to p intersects with $\Sigma_{3R/4}$ at some point q', and we take a point $q \in Y$ such that $d(q, q') < \varepsilon_0 \le \varepsilon_1$. As a consequence, $q \in N_{\varepsilon_1}(3R/4) \cap Y$. Thus the bounds in (5.14) still hold. Then the exact proof of Case 1 works in this case. This concludes the proof of the first claim.

(2) For any $x \in M$, there exists $\widehat{r} \in \widehat{\mathcal{R}}_Y$ such that $\|r_x - \widehat{r}\|_{\ell^\infty(Y)} \le \varepsilon_1$ by (1.8). Let $x' \in M$ be a point corresponding to $\widehat{r}$, i.e. satisfying $\|r_{x'} - \widehat{r}\|_{\ell^\infty(Y)} \le \varepsilon_1$. Then it follows that

$$\|r_x - r_{x'}\|_{\ell^\infty(Y)} \le 2\varepsilon_1. \tag{5.18}$$

If $x \in U$, then there exists $y \in Y$ such that $d(x, y) < \varepsilon_0 \le \varepsilon_1$, and hence $d(x', y) < 3\varepsilon_1$ by (5.18). Thus $d(x, x') < 4\varepsilon_1$.

If $x \in M \setminus U$, we take an arbitrary point $p \in N_{\varepsilon_1}(R/4) \cap Y$. The minimizing geodesic from x to p intersects with $\Sigma_{R/2}$ at some point q', and we take a point $q \in Y$ such that $d(q, q') < \varepsilon_0 \le \varepsilon_1$. As a consequence, $q \in N_{\varepsilon_1}(R/2) \cap Y$. Similarly as (1), by the triangle inequality,

$$d(p, x) \ge d(p, q) + d(q, x) - 2\varepsilon_1, \tag{5.19}$$

and by (5.18),

$$\begin{aligned} d(p, x') \ge d(p, x) - 2\varepsilon_1 &\ge d(p, q) + d(q, x) - 4\varepsilon_1 \\ &\ge d(p, q) + d(q, x') - 6\varepsilon_1. \end{aligned} \tag{5.20}$$

Thus the condition (5.6) is satisfied. Since $d(p, q),\ d(q, x),\ d(q, x')$ are all bounded below by $R/8$ by construction, one can repeat the proof of Lemma 5.1 from (5.19) and (5.20). In the end, we get the same conclusion as (5.10), namely

$$d(x, x') \le |d(x, q) - d(x', q)| + C_{15}\varepsilon_1^{1/2}, \tag{5.21}$$

which shows $d(x, x') \le 2\varepsilon_1 + C_{15}\varepsilon_1^{1/2}$ by (5.18). □

Remark. One also has the other direction of Proposition 5.2(1):

$$\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} \le d(x_\ell, x_{i_0}) + 2\varepsilon_1, \tag{5.22}$$

which holds without any assumption on the parameters. The proof is straightforward by the triangle inequality.

6. Existence of a frame associated with measurement points

We start with the following observation on geodesics connecting a point $x \in M$ to points near spheres $\Sigma_r := \partial B(x_0, r)$. We denote

$$N_\epsilon(r) := B(x_0, r + \epsilon) \setminus B(x_0, r - \epsilon). \tag{6.1}$$

For $x, y \in M$, we denote $|xy| = d(x, y)$, and denote by $[xy]$ a distance-minimizing geodesic from x to y. The angle between $[xy]$ and $[xz]$ at x is denoted by $\angle yxz$.

Lemma 6.1. *Let M be a closed Riemannian manifold with diameter bound $diam(M) \le \Lambda$ and sectional curvature bounded by $Sec_M \ge -\Lambda^2$. Let $B(x_0, R)$ be an open ball with $R < \mathrm{inj}(M)/2$. Then there exist uniform constants $\widehat{\varepsilon}, c_3, C_{16} > 0$ explicitly depending on Λ, R, such that the following holds for $0 < \varepsilon < \widehat{\varepsilon}$.*

Given a point x with $d(x, x_0) \ge 2R$, take p_0 to be a nearest point in $B(x_0, R)$ from x. Let $p_1, p_2 \in N_\varepsilon(R)$ such that

$$|p_1p_2| \ge C_{17}\varepsilon, \quad |p_0p_1| < c_3, \quad |p_0p_2| < c_3, \tag{6.2}$$

for some $C_{17} \ge 5$. Then

(1) *$\angle p_1xp_2 > C_{16}C_{17}\varepsilon$.*
(2) *Let $q_1 \in [xp_1]$, $q_2 \in [xp_2]$ such that $|xq_1| > R/2$ and $|xq_2| > R/2$. Then $|q_1q_2| > RC_{16}C_{17}\varepsilon/4$.*

Proof. (1) Suppose $|xp_1| \le |xp_2|$. By the triangle inequality,

$$|xx_0| = |xp_0| + R \ge |xp_1| - c_3 + |x_0p_1| - \varepsilon.$$

Then by a shortcut argument using Lemma 3.1(2), similar to (4.14), we have

$$(\pi - \eta_1)^2 \le C_7^{-1}(c_3 + \varepsilon), \tag{6.3}$$

where $\eta_1 \in [0,\pi]$ is the angle between $[p_1x]$ and $[p_1x_0]$ at p_1. We choose c_3, ε such that $\pi - \eta_1 \in [0, \pi/6]$.

Moreover, we can choose sufficiently small c_3 depending on Λ, R, such that the angle $\angle x_0p_1p_2$ is bounded below by $\pi/3$. This can be proved as follows. Applying Lemma 3.1(1) to the triangle $x_0p_1p_2$, one has

$$2\varepsilon \ge |x_0p_1| - |x_0p_2| \ge |p_1p_2|\cos(\angle x_0p_1p_2) - C_6\max\{\Lambda, R^{-1}\}|p_1p_2|^2.$$

Then using the condition (6.2),

$$\cos(\angle x_0p_1p_2) \le 2\varepsilon|p_1p_2|^{-1} + C_6\max\{\Lambda, R^{-1}\}|p_1p_2| < \frac{2}{5} + C(\Lambda, R)2c_3. \tag{6.4}$$

The above shows that we can choose sufficiently small c_3, ε, such that the angle $\angle xp_1p_2$ satisfies

$$\angle xp_1p_2 \le (\frac{\pi}{2} + \pi - \eta_1) + (\frac{\pi}{2} - \frac{\pi}{3}) \le \frac{5}{6}\pi. \tag{6.5}$$

Next, we apply Lemma 3.1(1) to the triangle xp_1p_2 and use (6.5) as follows:

$$\begin{aligned} |xp_2| &\le |xp_1| - |p_1p_2|\cos(\angle xp_1p_2) + C_6\max\{\Lambda, R^{-1}\}|p_1p_2|^2 \\ &\le |xp_1| + \frac{\sqrt{3}}{2}|p_1p_2| + C_6\max\{\Lambda, R^{-1}\}|p_1p_2|^2. \end{aligned}$$

Using $|p_1p_2| < 2c_3$, we obtain for sufficiently small c_3,

$$|xp_2| - |xp_1| \le \frac{\sqrt{3}}{2}|p_1p_2| + C_6\max\{\Lambda, R^{-1}\}|p_1p_2|^2 < \frac{9}{10}|p_1p_2|. \tag{6.6}$$

Thus by Lemma A.1 and (6.6), we obtain

$$\begin{aligned} |p_1p_2| &\le |xp_2| - |xp_1| + C_A\angle p_1xp_2 \\ &< \frac{9}{10}|p_1p_2| + C_A\angle p_1xp_2, \end{aligned} \tag{6.7}$$

which proves part (1) due to $|p_1p_2| \ge C_{17}\varepsilon$, with $C_{16} = (10C_A)^{-1}$.

(2) Consider the triangle xp_1p_2 and denote its comparison triangle in the hyperbolic plane of constant sectional curvature $-\Lambda^2$ by $\overline{xp_1p_2}$, i.e. $|xp_1| = |\overline{xp_1}|$, $|xp_2| = |\overline{xp_2}|$, $|p_1p_2| = |\overline{p_1p_2}|$. The angle at $\overline{x}$ of this comparison triangle is denoted by $\overline{\angle}p_1xp_2$. Since the inequalities (6.6) and (6.7) are also valid for the comparison triangle $\overline{xp_1p_2}$, it also holds that $\overline{\angle}p_1xp_2 > C_{16}C_{17}\varepsilon$ in the comparison triangle. By the angle-sidelength monotonicity version of Toponogov's theorem (e.g. [4, Theorem 7.3.2]), the function $(|xq_1|, |xq_2|) \mapsto \overline{\angle}q_1xq_2$ is decreasing in both arguments. Hence

$$\overline{\angle}q_1xq_2 \ge \overline{\angle}p_1xp_2 > C_{16}C_{17}\varepsilon. \tag{6.8}$$

Then in the comparison triangle $\overline{xq_1q_2}$, by (A.5) and (6.8),

$$\begin{aligned}\cosh(\Lambda|\overline{q}_1\overline{q}_2|) &= \cosh(\Lambda|\overline{x}\overline{q}_1|)\cosh(\Lambda|\overline{x}\overline{q}_2|) - \sinh(\Lambda|\overline{x}\overline{q}_1|)\sinh(\Lambda|\overline{x}\overline{q}_2|)\cos(\overline{\angle}q_1xq_2)\\ &= \cosh\big(\Lambda|xq_1| - \Lambda|xq_2|\big) + \sinh(\Lambda|xq_1|)\sinh(\Lambda|xq_2|)\big(1-\cos(\overline{\angle}q_1xq_2)\big)\\ &> 1 + \frac{\Lambda^2R^2}{4}\big(1-\cos(C_{16}C_{17}\varepsilon)\big)\\ &\geq 1 + \frac{\Lambda^2R^2}{4}\frac{1}{4}(C_{16}C_{17}\varepsilon)^2.\end{aligned}$$

Since $\cosh(\Lambda|\overline{q}_1\overline{q}_2|) = \cosh(\Lambda|q_1q_2|) \leq 1 + \Lambda^2|q_1q_2|^2$ for small $|q_1q_2|$, the second claim follows. □

Given a point $x \in M$ outside the ball $B(x_0, 2R)$ and an ε-net of $N_\varepsilon(R)$, we consider all the unit initial vectors of minimizing geodesics from x to points in the ε-net. The following lemma shows that there exist n such unit vectors so that the corresponding determinant is bounded away from zero.

Lemma 6.2. *Let M be a closed Riemannian manifold with diameter bound $diam(M) \leq \Lambda$ and sectional curvature bounded by $|Sec_M| \leq \Lambda^2$. Let $B(x_0,R)$ be an open ball for $R < \mathrm{inj}(M)/2$, and Y be an ε-net of $B(x_0,R)$. Given a point x with $d(x,x_0) \geq R$, take p_0 to be a nearest point in $B(x_0,R/2)$ from x. Then for every $C_{17} \geq 5$, there are uniform constants $\widehat{\varepsilon}, c_1, c_3 > 0$ explicitly depending on n, Λ, R, C_{17}, such that the following holds for $0 < \varepsilon < \widehat{\varepsilon}$.*

We can find a separated set $\{p_{j(l)}:\ l = 1,2,\dots,L\} \subset Y \cap N_\varepsilon(R/2)$ satisfying

$$|p_{j(l)}p_{j(m)}| \geq C_{17}\varepsilon, \quad \text{for } l \neq m, \text{ and } \quad |p_0p_{j(l)}| < c_3, \tag{6.9}$$

such that the following statement is true: there exist n unit vectors $w_{l_1}, \dots, w_{l_n}$ such that the volume of the simplex in T_xM with the vertices $0, w_{l_1}, \dots, w_{l_n}$ is larger than c_1, where $w_l \in S_xM$, $l = 1,2,\dots,L$ is the unit initial vector of a minimizing geodesic in M from x to $p_{j(l)}$.

Proof. Let us take $C_{17} = 10$ for the sake of argument. First, we choose a suitable separated set. We can choose a separated set of $Y \cap N_\varepsilon(R/2)$ satisfying the condition (6.9) such that it is also a 11ε-net. Namely, first pick any point, say p_1, from $Y \cap N_\varepsilon(R/2)$, and then take all points $p_2, \dots, p_l$ in $Y \cap N_\varepsilon(R/2)$ with distance at least 10ε away from p_1. For the second step, pick all points with distance at least 10ε from all of $p_1, p_2, \dots, p_l$. Repeat this procedure and the procedure stops in finite steps. The chosen points form a 11ε-net of $N_\varepsilon(R/2)$. Indeed, for any $z \in N_\varepsilon(R/2)$, there exists a point $y \in Y \cap N_\varepsilon(R/2)$ such that $d(y,z) < \varepsilon$ by the ε-net condition for Y. If y is also chosen by our procedure above, then we are done. If y is not chosen, it means that there must be at least one chosen point p_m such that $d(y,p_m) < 10\varepsilon$ otherwise y would have been chosen. In such case, we have $d(z,p_m) \leq d(z,y) + d(y,p_m) < 11\varepsilon$.

The separated set of our choice above satisfying the condition (6.9) has cardinality at least

$$L \geq C(n,\Lambda)c_3\varepsilon^{-(n-1)}. \tag{6.10}$$

Indeed, the separated set that we chose above is a 11ε-net of the set $N_\varepsilon(R/2) \cap B(p_0,c_3)$, and the latter set has volume bounded below. More precisely,

$$C(n)(11\varepsilon)^nL \geq \mathrm{vol}_n\big(N_\varepsilon(R/2) \cap B(p_0,c_3)\big) \geq C(n,\Lambda)c_3\varepsilon,$$

which yields the lower bound (6.10).

On the other hand, by Lemma 6.1(1), the set $\{w_l\}_{l=1}^L \subset S_xM$ of unit vectors is $C_{16}\varepsilon$-separated. We claim that we can choose sufficiently small h explicitly depending on n, Λ, c_3, such that for any unit vector $\xi \in S_xM \setminus \{0\}$,

$$\{w_l\}_{l=1}^L \not\subset \{v \in S_xM : \langle v, \xi\rangle_g \in (-h,h)\}. \tag{6.11}$$

The claim (6.11) can be proved as follows. Suppose (6.11) is not true: there exists ξ such that $\{w_l\}_{l=1}^L \subset A(\xi,h)$, where $A(\xi,h)$ denotes the set on the right-hand side of (6.11). Since $\{w_l\}_{l=1}^L \subset S_xM$ is $C_{16}\varepsilon$-separated, this implies

$$C(n)(C_{16}\varepsilon)^{n-1}L \le \mathrm{vol}_{n-1}(A(\xi,h)).$$

However, the $(n-1)$-dimensional volume of $A(\xi,h)$ is bounded above by $C(n)h$. Hence (6.10) yields

$$C(n,\Lambda)c_3 \le C(n)(C_{16}\varepsilon)^{n-1}L \le \mathrm{vol}_{n-1}(A(\xi,h)) \le C(n)h,$$

which cannot be true if $h < C(n,\Lambda)c_3$. The claim (6.11) is proved.

Thus we can construct the desired $w_{l_1},\dots,w_{l_n}$ as follows. Let w_{l_1} be chosen arbitrarily. Then when $w_{l_1},\dots,w_{l_k}$ are chosen, let ξ be a unit vector orthogonal to $w_{l_1},\dots,w_{l_k}$. From the claim (6.11), there exists $w_{l_{k+1}}$ such that $w_{l_{k+1}} \notin A(\xi,h)$ with h properly chosen as above. In other words, the angle between $w_{l_{k+1}}$ and the linear subspace spanned by $w_{l_1},\dots,w_{l_k}$ is bounded below by h. Hence the simplex with vertices $0, w_{l_1},\dots,w_{l_k},w_{l_{k+1}}$ has volume bounded below by $C(n,h)$. □

In other words, there exist n unit vectors in $\{w_l\}_{l=1}^L$ such that the corresponding determinant $\det([\langle w_k,w_m\rangle_g]_{k,m=1}^n) > c_1$.

The next lemma is a technical modification of the previous lemma.

Lemma 6.3. *Under the setting of Lemma 6.2, there are uniform constants $\widehat{\varepsilon}, c_1, c_3, c_4 > 0$ explicitly depending on n,Λ,R, such that the following holds for $0<\varepsilon<\widehat{\varepsilon}$.*

Let $\{p_{j(l)}:\ l=1,2,\dots,L\} \subset Y\cap N_\varepsilon(R/2)$ be a choice of points satisfying the condition (6.9) for sufficiently large C_{17}. For each l, suppose a minimizing geodesic $[xp_{j(l)}]$ intersects with $\partial B(R/2+c_4)$ at $q'_{j(l)}$, and we take $q_{j(l)} \in Y\cap N_\varepsilon(R/2+c_4)$ to be a point in Y such that $|q'_{j(l)}q_{j(l)}| < \varepsilon$.

Then we have

$$|q_{j(l)}q_{j(m)}| \ge 10\varepsilon, \quad \text{for } l\neq m, \text{ and } \quad |p_0q_{j(l)}| < 2c_3. \tag{6.12}$$

As a consequence, there exist n unit vectors $v_{l_1},\dots,v_{l_n}$ such that the volume of the simplex in T_xM with the vertices $0, v_{l_1},\dots,v_{l_n}$ is larger than c_1, where $v_l\in S_xM$, $l=1,2,\dots,L$ is the unit initial vector of a minimizing geodesic in M from x to $q_{j(l)}$.

Proof. Fix the parameters $\widehat{\varepsilon}, c_3$ as chosen in Lemma 6.2. Let us consider two points p_1,p_2 in the maximal set constructed in Lemma 6.2 with C_{17} to be determined later. The intersection of $[xp_1]$, $[xp_2]$ with $\partial B(R/2+c_4)$ is q'_1,q'_2. By Lemma 6.1(2), $|q'_1q'_2| > RC_{16}C_{17}\varepsilon/4$. When we take the points q_1,q_2 in the ε-net, we have

$$|q_1q_2| > \Big(\frac{RC_{16}C_{17}}{4} - 2\Big)\varepsilon. \tag{6.13}$$

Then we can choose $C_{17} = 48R^{-1}C_{16}^{-1}$ so that $|q_1q_2| > 10\varepsilon$. Furthermore, since the incident angle is bounded by (6.3), we can choose sufficiently small $c_4>0$ such that $|p_1q'_1|$ is bounded by $c_3/2$. Hence $|p_0q_1| < 2c_3$. This proves the claim (6.12).

Let q_0 be the intersection of $[xp_0]$ with $\partial B(R/2+c_4)$, and hence q_0 is a nearest point in $B(R/2+c_4)$ from x. Moreover, $|q_0q_1| < 3c_3$ if we choose $c_4<c_3$. Thus the condition (6.2) in Lemma 6.1 is satisfied by the triangle xq_1q_2, which gives $\angle q_1xq_2 > C_{16}\varepsilon$. Observe that the total number of points $\{q_{j(l)}\}$ is equal to the total number L of points $\{p_{j(l)}\}$, because of $\{q_{j(l)}\}$ being 10ε-separated. The total number L is bounded below by (6.10). Then the same argument yields (6.11) for $\{v_l\}_{l=1}^L$, and the second claim follows from the last part of the proof of Lemma 6.2. □

Remark 6.4. *Lemma 6.1(1) is also true if x is inside the ball, say $x \in B(x_0, R/2)$, assuming two-side bounds on the sectional curvature $|Sec_M| \le \Lambda^2$. The proof is similar and can be found in Lemma A.3. As a consequence, Lemma 6.2 is still valid by the same argument if $x \in B(x_0, R/2)$, in which case we can find the desired separated set in $Y \cap N_\varepsilon(R)$. We will use this observation in the next section.*

7. Local reconstructions from partial distance data

This section is the proof of Theorem 2.4 and consequently Corollary 2.5. Let $x_i \in M$ be the points corresponding to $\widehat{r}_i \in \widehat{\mathcal{R}}_Y$, $i = 1, 2, \dots, I$, i.e. satisfying (5.1). Let us fix one element $\widehat{r}_{i_0} \in \widehat{\mathcal{R}}_Y$ and the corresponding point $x_{i_0} \in M$. The basic idea of the proof is to find appropriate points in $B(x_0, R) \subset U$, and apply geometric lemmas in previous sections to approximate the inner product. One important point is to keep distances of points bounded away from zero, as required by previous lemmas. This is possible because we assumed the knowledge of a point $y_0 \in Y$ such that $d(x_0, y_0) < \varepsilon_0$. This assumption enables us to determine where x_{i_0} lies in reference to $B(x_0, R)$ up to a small error. Furthermore, it is possible to use only part of all measurement points in the ball. This allows us to simply take, for example $R = (4\Lambda)^{-1}$, and only consider the measurement points in Y within this smaller ball. In the proof, we keep the parameter R for clarity, and note that any dependency of R in the constants can be replaced by Λ.

We divide the proof of Theorem 2.4 into two cases depending on where x_{i_0} lies, in a similar way as we considered in Proposition 5.2(1). We will focus on the first case, as the second case is a simple modification from the first case.

Case 1: $\widehat{r}_{i_0}(y_0) > R/2$.

Let us set

$$0 < \varepsilon_0 \le \varepsilon_1 < \min\{\frac{1}{16}, \frac{R}{32}\}. \tag{7.1}$$

We consider the elements $\widehat{r}_\ell \in \widehat{\mathcal{R}}_Y$ in the neighborhood of $\widehat{r}_{i_0}$,

$$\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} < \rho_0, \quad \rho_0 < \min\{\frac{1}{4}, \frac{R}{16}\}, \tag{7.2}$$

with the parameter ρ_0 to be determined later. Let $x_\ell \in M$ be a point corresponding to $\widehat{r}_\ell$.

- *Step 1: Applying first variation formula.*

The condition $\widehat{r}_{i_0}(y_0) > R/2$ implies that $d(x_{i_0}, x_0) > R/2 - 2\varepsilon_1$. For an arbitrary point $p \in N_{\varepsilon_1}(R/8) \cap Y$, we pick a point $q \in N_{\varepsilon_1}(R/4) \cap Y$ such that

$$\left|\widehat{r}_{i_0}(p) - \widehat{r}_{i_0}(q) - D_Y^a(p,q)\right| < 6\varepsilon_1. \tag{7.3}$$

Recall D_Y^a defined in (3.4). It is clear that condition (7.3) can be tested using the given data $\widehat{\mathcal{R}}_Y$ only. Observe that the set of points q (for each p) satisfying (7.3) is nonempty. This is because the point in Y within ε_0-distance from the intersection of a minimizing geodesic $[x_{i_0} p]$ with $\Sigma_{R/4}$ satisfies the condition (7.3) due to (5.15) and (5.16). In particular, the following bounds are valid:

$$d(p,q) > \frac{R}{16}, \quad d(q, x_{i_0}) > \frac{R}{8}, \quad d(q, x_\ell) > \frac{R}{16}. \tag{7.4}$$

Moreover, by (7.3), (5.1), (3.7), and (3.8) we see that

$$\left|d(x_{i_0}, p) - d(x_{i_0}, q) - d(p,q)\right| \le 10\varepsilon_1. \tag{7.5}$$

Hence the assumptions of Proposition 4.3 are satisfied when $\varepsilon_1 < \widehat{\delta}/10$.

By Proposition 5.2(1), we know

$$d(x_\ell, x_{i_0}) \le 3C_{15}(\|\widehat{r}_\ell - \widehat{r}_{i_0}\|_{\ell^\infty(Y)} + 3\varepsilon_1)^{\frac{1}{2}}. \tag{7.6}$$

Hence if we choose ρ_0 in (7.2) such that

$$3C_{15}(\rho_0 + 3\varepsilon_1)^{1/2} < c_2^2\,, \tag{7.7}$$

we can apply Proposition 4.3 (with $x = x_{i_0}$, $y = x_\ell$) and obtain

$$\Big|\langle \xi_\ell, v\rangle_g - \big(|x_{i_0}q| - |x_\ell q|\big)\Big| \le C_{13}|x_{i_0}x_\ell|\big(\varepsilon_1^{1/4} + |x_{i_0}x_\ell|^{1/3}\big), \tag{7.8}$$

where $\xi_\ell = \exp_{x_{i_0}}^{-1}(x_\ell)$ (of length $|x_{i_0}x_l|$), and v is the unit initial vector of $[x_{i_0}q]$. Then using (5.1), we have

$$\Big|\langle \xi_\ell, v\rangle_g - \big(\widehat{r}_{i_0}(q) - \widehat{r}_\ell(q)\big)\Big| \le C_{13}|x_{i_0}x_\ell|\big(\varepsilon_1^{1/4} + |x_{i_0}x_\ell|^{1/3}\big) + 2\varepsilon_1. \tag{7.9}$$

- *Step 2: Finding the length* $|x_{i_0}x_\ell|$.

We aim to construct an approximate inner product to the actual one, i.e. the first term in (7.9). However, the first term $\langle \xi_\ell, v\rangle_g$ involves the length $|\xi_\ell|_g = |x_{i_0}x_\ell|$, which cannot be exactly computed from the data $\widehat{\mathcal{R}}_Y$. What we can do is to use Lemma 5.1 to approximate it.

Let us take a small parameter s whose value is determined later:

$$s \in (\varepsilon_1^{1/2}, \frac{\rho_0}{2}). \tag{7.10}$$

Let $p \in N_{\varepsilon_1}(R/8) \cap Y$ be arbitrary, and let $q \in N_{\varepsilon_1}(R/4) \cap Y$ be chosen according to the condition (7.3). Now we choose an element $\widehat{r}_\ell \in \widehat{\mathcal{R}}_Y$ satisfying (7.2) such that the following conditions (7.11) and (7.12) hold:

$$\big|\widehat{r}_\ell(p) - \widehat{r}_\ell(q) - D_Y^a(p,q)\big| \le 9\varepsilon_1, \tag{7.11}$$

$$\big|\widehat{r}_{i_0}(q) - (\widehat{r}_\ell(q) + s)\big| \le 9\varepsilon_1. \tag{7.12}$$

In fact, as we will later specify our choice $s = \varepsilon_1^{3/8}$, we can actually choose this $\widehat{r}_\ell$ from $\widehat{\mathcal{R}}_Y \cap \mathcal{B}_\infty(\widehat{r}_{i_0}, \varepsilon_1^{1/4})$. Indeed, considering (5.22), it is straightforward to check that the element in $\widehat{\mathcal{R}}_Y \cap \mathcal{B}_\infty(\widehat{r}_{i_0}, \varepsilon_1^{1/4})$ corresponding to $\gamma_{x_{i_0},w}(s)$ satisfies these two conditions above, where w is the unit initial vector of $[x_{i_0}p]$. Essentially, these conditions can be understood as a test to search for $\gamma_{x_{i_0},w}(s)$ up to a small error, see Figure 6.

Next, let us discuss the properties of such $\widehat{r}_\ell$ satisfying the criteria (7.11) and (7.12). Due to Lemma 5.1, (7.3) and (7.11) imply that

$$\Big||x_{i_0}x_\ell| - |\widehat{r}_{i_0}(q) - \widehat{r}_\ell(q)|\Big| \le 3C_{15}\varepsilon_1^{1/2}. \tag{7.13}$$

Hence (7.13), (7.12) and (5.1) yield for some suitable $C_{18} > 1$,

$$\big||x_{i_0}x_\ell| - s\big| \le C_{18}\varepsilon_1^{1/2}, \tag{7.14}$$

$$\big||x_{i_0}x_\ell| - (|x_{i_0}q| - |x_\ell q|)\big| \le C_{18}\varepsilon_1^{1/2}. \tag{7.15}$$

Furthermore, let $\gamma_{x_{i_0},v}(\cdot)$ be a minimizing geodesic from x_{i_0} to q, and $\xi_\ell = \exp_{x_{i_0}}^{-1}(x_\ell)$. We claim that for some uniform constant $C_{19}' > 1$,

$$d(\gamma_{x_{i_0},v}(s), x_\ell) \le C_{19}'\varepsilon_1^{1/2}. \tag{7.16}$$

This claim can be proved as follows. Denote $z' = \gamma_{x_{i_0},v}(s)$ and observe that z' is on a minimizing geodesic from x_{i_0} to q. Due to the fact that

$$d(x_{i_0}, p) - d(x_{i_0}, q) \le d(x_{i_0}, z') + d(z', p) - d(x_{i_0}, z') - d(z', q) = d(z', p) - d(z', q),$$

the inequality (7.5) is still valid after replacing x_{i_0} with $z' = \gamma_{x_{i_0},v}(s)$, that is,

$$\big|d(z',p) - d(z',q) - d(p,q)\big| \le 10\varepsilon_1. \tag{7.17}$$

Thus Lemma 5.1 (with $x_1 = z',\ x_2 = x_\ell$) and (7.11) yield that

$$\big|d(z',x_\ell) - |d(z',q) - d(x_\ell,q)|\big| \le C_{15}\varepsilon_1^{1/2}.$$

Since

$$\begin{aligned} d(z',q) - d(x_\ell,q) &= |z'q| + |x_{i_0}z'| - (|x_\ell q| + |x_{i_0}x_\ell|) + |x_{i_0}x_\ell| - |x_{i_0}z'| \\ &= |x_{i_0}q| - (|x_\ell q| + |x_{i_0}x_\ell|) + (|x_{i_0}x_\ell| - s), \end{aligned}$$

then the claim (7.16) follows from (7.14) and (7.15).

As a consequence, by applying the Rauch comparison theorem (see e.g. [69]) for $\mathrm{Sec}_M \le \Lambda^2$ (in a neighborhood of x_{i_0}), (7.16) yields that for some $C_{19} > 1$,

$$|\xi_\ell - sv| \le C_{19}\varepsilon_1^{1/2}. \tag{7.18}$$

- *Step 3: Approximating the inner product.*

Denote by s_0, s_1 the lower and upper bounds for $|x_{i_0}x_\ell|$. From (7.14), we set

$$s_0 := s - C_{18}\varepsilon_1^{1/2}, \quad s_1 := s + C_{18}\varepsilon_1^{1/2}. \tag{7.19}$$

We require that $s > 2C_{18}\varepsilon_1^{1/2}$ so that $s_0 > 0$ and $s_1 < 2s$. Observe that (7.14) yields

$$\Big|\langle \xi_\ell, v\rangle_g - s\langle \frac{\xi_\ell}{|\xi_\ell|_g}, v\rangle_g\Big| = \Big||x_{i_0}x_\ell|\langle \frac{\xi_\ell}{|\xi_\ell|_g}, v\rangle_g - s\langle \frac{\xi_\ell}{|\xi_\ell|_g}, v\rangle_g\Big| \le C_{18}\varepsilon_1^{1/2}. \tag{7.20}$$

Then (7.9) gives

$$\Big|s\langle \frac{\xi_\ell}{|\xi_\ell|_g}, v\rangle_g - \big(\widehat{r}_{i_0}(q) - \widehat{r}_\ell(q)\big)\Big| \le C_{13}s_1(\varepsilon_1^{1/4} + s_1^{1/3}) + (C_{18}+2)\varepsilon_1^{1/2}. \tag{7.21}$$

Hence dividing by s gives

$$\begin{aligned} \Big|\langle \frac{\xi_\ell}{|\xi_\ell|_g}, v\rangle_g - \frac{1}{s}\big(\widehat{r}_{i_0}(q) - \widehat{r}_\ell(q)\big)\Big| &\le C_{13}s_1(\varepsilon_1^{1/4} + s_1^{1/3})s^{-1} + (C_{18}+2)\varepsilon_1^{1/2}s^{-1} \\ &\le 4(C_{13}+C_{18}+2)s^{-1}(\varepsilon_1^{1/2} + \varepsilon_1^{1/4}s + s^{4/3}). \end{aligned}$$

We obtain a good estimate when we choose

$$s = \varepsilon_1^{3/8}. \tag{7.22}$$

Note that when ε_1 is sufficiently small, the requirement that $s > 2C_{18}\varepsilon_1^{1/2}$ is satisfied. In such case, we have the estimate

$$\Big|\langle \frac{\xi_\ell}{|\xi_\ell|_g}, v\rangle_g - \frac{1}{s}\big(\widehat{r}_{i_0}(q) - \widehat{r}_\ell(q)\big)\Big| \le C_{20}\varepsilon_1^{1/8}. \tag{7.23}$$

- *Step 4: Approximating the metric.*

Let us find a proper frame in $T_{x_{i_0}}M$ and apply the estimate (7.23) to each vector in the frame to approximate the metric. First, we search for a point $p_0 \in N_{\varepsilon_1}(R/8) \cap Y$ such that

$$\widehat{r}_{i_0}(p_0) = \min_{y \in N_{\varepsilon_1}(R/8)\cap Y} \widehat{r}_{i_0}(y). \tag{7.24}$$

Let $p_{j(k)} \in N_{\varepsilon_1}(R/8) \cap Y,\ j(k) \in \{0,\dots,J\},\ k = 1,\dots,n$, be arbitrary n points in Y satisfying the condition (6.9) (with $\varepsilon = \varepsilon_1$), which we will vary later. For each $p_{j(k)}$, we search for $q_{j(k)} \in N_{\varepsilon_1}(R/4) \cap Y$ such that (7.3) hold, i.e.

$$\big|\widehat{r}_{i_0}(p_{j(k)}) - \widehat{r}_{i_0}(q_{j(k)}) - D_Y^a(p_{j(k)}, q_{j(k)})\big| < 6\varepsilon_1. \tag{7.25}$$

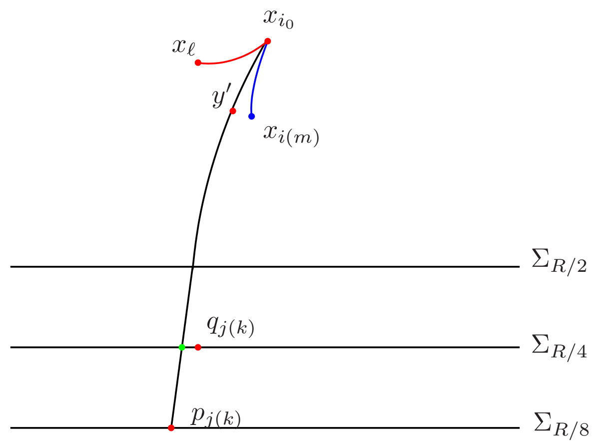

FIGURE 6. *The geodesic $[x_{i_0}p_{j(k)}]$ intersects $\Sigma_{R/4} = \partial B(x_0, R/4)$ at a green point that is close to the red point $q_{j(k)}$. The other red point is $y' = \gamma_{x_{i_0},w}(s)$, where $w \in S_{x_{i_0}}M$ is the direction vector of the geodesic $[x_{i_0}p_{j(k)}]$ and $s = d(x_{i_0}, y')$. When the criteria (7.11) and (7.12) are satisfied for $p_{j(k)}, q_{j(k)}$, the point $x_{i(m)}$ corresponding to $\widehat{r}_{i(m)}$ is close to y'. The blue curve is the geodesic $[x_{i_0}x_{i(m)}]$ and $\xi_{i(m)} = \exp^{-1}_{x_{i_0}}(x_{i(m)})$. Similarly, the red curve is the geodesic $[x_{i_0}x_\ell]$ and $\xi_\ell = \exp^{-1}_{x_{i_0}}(x_\ell)$. In the proof we show that we can find indexes $i(m)$, $m = 1, 2, \dots, n$, such that the unit vectors $\overline{\xi}_{i(m)} := |\xi_{i(m)}|_g^{-1}\xi_{i(m)}$ form a good basis of $T_{x_{i_0}}M$. Moreover, we show that we can approximately compute the inner product of the vectors $\overline{\xi}_{i(m)}$ and ξ_ℓ, that is, approximately find the coordinates of the points x_ℓ in the normal coordinates at x_{i_0}.*

Then we choose an element $\widehat{r}_{i(k)} \in \widehat{\mathcal{R}}_Y \cap \mathcal{B}_\infty(\widehat{r}_{i_0}, \varepsilon_1^{1/4})$, $i(k) \in \{1, \dots, I\}$ such that (7.2), (7.11) and (7.12) hold for $p_{j(k)}, q_{j(k)}$, namely

$$\big|\widehat{r}_{i(k)}(p_{j(k)}) - \widehat{r}_{i(k)}(q_{j(k)}) - D^a_Y(p_{j(k)}, q_{j(k)})\big| \le 9\varepsilon_1, \tag{7.26}$$

$$\big|\widehat{r}_{i_0}(q_{j(k)}) - (\widehat{r}_{i(k)}(q_{j(k)}) + s)\big| \le 9\varepsilon_1. \tag{7.27}$$

By the discussion following (7.12), the set of $\widehat{r}_{i(k)}$ satisfying the conditions above is non-empty, as we set $s = \varepsilon_1^{3/8}$.

Now we apply the estimate (7.23) for each index $j(k)$. Denote by $v_{j(k)}$ the unit initial vectors of minimizing geodesics $[x_{i_0}q_{j(k)}]$. For every $k, m \in \{1, \dots, n\}$, we have

$$\Big|\langle \frac{\xi_{i(m)}}{|\xi_{i(m)}|_g}, v_{j(k)}\rangle_g - \frac{1}{s}\big(\widehat{r}_{i_0}(q_{j(k)}) - \widehat{r}_{i(m)}(q_{j(k)})\big)\Big| \le C_{20}\varepsilon_1^{1/8}, \tag{7.28}$$

where $\xi_{i(m)} = \exp^{-1}_{x_{i_0}}(x_{i(m)})$. By (7.18),

$$\big|\langle \xi_{i(m)}, v_{j(k)}\rangle_g - s\langle v_{j(m)}, v_{j(k)}\rangle_g\big| \le C_{19}\varepsilon_1^{1/2}, \quad \forall\, k, m = 1, \dots, n. \tag{7.29}$$

Hence (7.20), (7.22) and (7.28) yield

$$\Big|\langle v_{j(m)}, v_{j(k)}\rangle_g - \frac{1}{s}\big(\widehat{r}_{i_0}(q_{j(k)}) - \widehat{r}_{i(m)}(q_{j(k)})\big)\Big| \le C_{21}\varepsilon_1^{1/8},\ \forall\, k, m = 1, \dots, n. \tag{7.30}$$

The formula (7.30) shows that we can compute the numbers

$$G_{k,m} := \frac{1}{s}\big(\widehat{r}_{i_0}(q_{j(k)}) - \widehat{r}_{i(m)}(q_{j(k)})\big), \quad \text{where } s = \varepsilon_1^{3/8}, \tag{7.31}$$

such that

$$\big|\langle v_{j(k)}, v_{j(m)}\rangle_g - G_{k,m}\big| \le C_{21}\varepsilon_1^{1/8}, \quad \forall\, k,m = 1,\dots,n. \tag{7.32}$$

As above, we have considered any n indices $j(k)$, the points $p_{j(k)}, q_{i(k)}$ and the unit initial vectors $v_{j(k)} \in S_{x_{i_0}}M$ for geodesics $[x_{i_0}q_{j(k)}]$, $k = 1,2,\dots,n$. In view of Lemma 6.3[1], the formula (7.32) shows that when ε_1 is smaller than some uniform constant, there exist some indices $j(k)$ and points $p_{j(k)}, q_{j(k)}$ such that

$$\det([G_{k,m}]_{k,m=1}^n) > \frac{3}{4}c_1. \tag{7.33}$$

Hence we can search for such indices so that (7.33) is satisfied by computing the determinant $\det([G_{k,m}]_{k,m=1}^n)$ for each choice of indices.

Reconstruction of metric. Now let us summarize our procedure for the reconstruction of metric using the given data $\widehat{\mathcal{R}}_Y$ only. Let $\widehat{r}_{i_0}$ be given satisfying $\widehat{r}_{i_0}(y_0) > R/2$. Fix sufficiently small $\varepsilon_0 \le \varepsilon_1$ explicitly depending only on n, Λ. First, we search for a point $p_0 \in N_{\varepsilon_1}(R/8) \cap Y$ by (7.24). Due to (3.7), see also (3.8), and the assumption that $d(x_0, y_0) < \varepsilon_0$, the point p_0 can be chosen using the given data (up to an error of $3\varepsilon_1$).

For $k = 1,\dots,n$, we arbitrarily choose n points $p_{j(k)} \in N_{\varepsilon_1}(R/8) \cap Y$ satisfying the condition (6.9). The points $q_{j(k)} \in N_{\varepsilon_1}(R/4) \cap Y$ are chosen according to (7.25). For each k, choose one element $\widehat{r}_{i(k)} \in \widehat{\mathcal{R}}_Y \cap \mathcal{B}_\infty(\widehat{r}_{i_0}, \varepsilon_1^{1/4})$ such that (7.26) and (7.27) are satisfied. Thus we can compute the numbers $G_{k,m}$ defined by (7.31). We test all possible choices of n points $p_{j(k)}$ satisfying (6.9), and find one choice such that (7.33) is satisfied. Note that by (7.32), we know that the metric corresponding to the basis that $G_{k,m}$ approximates must satisfy $\det([\langle v_{j(k)}, v_{j(m)}\rangle_g]_{k,m=1}^n) > c_1/2$.

The discussion above shows that in the Riemannian normal coordinates, denoted below by $X : B(x_{i_0}, r) \to \mathbb{R}^n$, at x_{i_0} associated with the specific basis $\{v_{j(k)}\}_{k=1}^n$ that $G_{k,m}$ approximates, we can find the metric tensor $g_{km}(x_{i_0}) = \langle v_{j(k)}, v_{j(m)}\rangle_g$ up to a uniformly bounded error, that is, we can find numbers $\widehat{g}_{km} := G_{k,m}$ such that

$$\big|g_{km}(x_{i_0}) - \widehat{g}_{km}\big| \le C_{21}\varepsilon_1^{1/8}, \quad \forall\, k,m = 1,2,\dots,n. \tag{7.34}$$

In particular, since the basis $\{v_{j(k)}\}_{k=1}^n$ are unit, we have $|g_{km}| \le 1$, $|\widehat{g}_{km}| \le 2$ for all k,m.

Reconstruction of normal coordinates. Suppose we have already picked indices $j(k)$ and points $p_{j(k)}, q_{j(k)}$ such that the metric is approximated as above. We can also find the coordinates of the points corresponding to elements $\widehat{r}_\ell$ in a neighborhood of $\widehat{r}_{i_0}$ in the a normal coordinate $X : B(x_{i_0}, r) \to \mathbb{R}^n$, up to a uniformly bounded error. Indeed, for any $\widehat{r}_\ell$ satisfying (7.2), we can compute the numbers

$$\widehat{X}_k(x_\ell) := \widehat{r}_{i_0}(q_{j(k)}) - \widehat{r}_\ell(q_{j(k)}), \quad k = 1,2,\dots,n. \tag{7.35}$$

Note that due to Proposition 5.2, we can choose sufficiently small ρ_0 such that $d(x_\ell, x_{i_0}) < \mathrm{inj}(M)$. Then by (7.9), we have

$$\begin{aligned} \big|X_k(x_\ell) - \widehat{X}_k(x_\ell)\big| &\le C_{13}|x_{i_0}x_\ell|\big(\varepsilon_1^{1/4} + |x_{i_0}x_\ell|^{1/3}\big) + 2\varepsilon_1 \\ &\le 3C_{13}\big(|x_{i_0}x_\ell|^{4/3} + \varepsilon_1\big), \end{aligned}$$

where

$$X_k(x_\ell) := \langle \xi_\ell, v_{j(k)}\rangle_g = \langle \exp_{x_{i_0}}^{-1}(x_\ell), v_{j(k)}\rangle_g \tag{7.36}$$

[1]Note that Lemma 6.3 is applicable here if $R/8 \le c_4$. However if $R/8 > c_4$, one can replace the radius $R/4$ with $R/8 + c_4$, and instead find $q_{j(k)} \in N_{\varepsilon_1}(R/8 + c_4) \cap Y$. All relevant distances would be bounded below depending on c_4, which again depends on n, Λ, R.

are the true values of the coordinates of the point x_ℓ in the X-coordinates. Here we have used Young's inequality $ab \le a^{4/3} + b^4$ in the last inequality. This concludes the proof for Case 1.

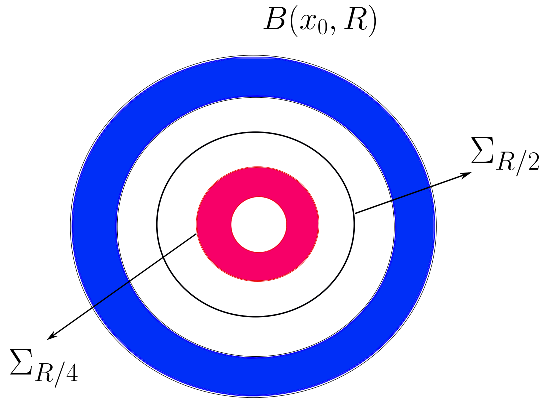

Figure 7. *When $x_{i_0} \in B(x_0, R/2)$, we choose measurement points p, q from the outer (blue) band such that the distances between x_{i_0}, p, q are bounded away from zero. When $x_{i_0} \notin B(x_0, R/2)$, we choose measurement points in the inner (red) band.*

- **Case 2:** $\widehat{r}_{i_0}(y_0) \le R/2$.

In this case, $d(x_{i_0}, x_0) \le R/2 + 2\varepsilon_1$. To keep distances bounded away from zero, one can choose points $q_{j(k)}, p_{j(k)}$ from the outer layer $B(x_0, R) \setminus B(x_0, 3R/4)$, see Figure 7. More precisely, we first choose arbitrary n points $\{q_{j(k)}\}$ in $N_{\varepsilon_1}(3R/4) \cap Y$ satisfying (6.9). Note that Lemma 6.2 still holds in this case due to Lemma A.3, see Remark 6.4. For each $q_{j(k)}$, we can choose one point $p_{j(k)} \in N_{\varepsilon_1}(R) \cap Y$ such that (7.25) is valid. Observe that the set of points satisfying (7.25) (for each $q_{j(k)}$) is nonempty. This is because one can extend the minimizing geodesic $[x_{i_0} q_{j(k)}]$ further until it intersects with Σ_R, since we are within the injectivity radius. Thus the point in $N_{\varepsilon_1}(R) \cap Y$ within ε_0-distance from the intersection point satisfies (7.25). In particular, the bounds in (7.4) still hold. From this point, the exact proof of Case 1 works in this case.

Now we prove Corollary 2.5.

Proof of Corollary 2.5. Let us fix the basis $\{v_k\}_{k=1}^n$ for which the metric has been approximated in Theorem 2.4. Observe that

$$d(x_\ell, x_{i_0}) = \big|\exp^{-1}_{x_{i_0}}(x_\ell)\big|_g = \Big(\sum_{j,k=1}^n g^{jk} X_j X_k\Big)^{1/2},$$

where X_k is defined in (2.7) and (g^{jk}) is the matrix inverse of (g_{jk}). Hence we define an approximate distance by

$$\widehat{d}_{\ell,i_0} = \Big(\sum_{j,k=1}^n \widehat{g}^{jk} \widehat{X}_j \widehat{X}_k\Big)^{1/2}. \tag{7.37}$$

We use the following notation for convenience:

$$|\xi|_g^2 = \sum_{j,k=1}^n g^{jk} X_j X_k, \quad |\xi|_{\widehat{g}}^2 = \sum_{j,k=1}^n \widehat{g}^{jk} X_j X_k, \quad \xi = (X_1, \dots, X_n). \tag{7.38}$$

Denote $\widehat{\xi}=(\widehat{X}_1,\ldots,\widehat{X}_n)$. Then

$$\begin{aligned}\big|\widehat{d}_{\ell,i_0}-d(x_\ell,x_{i_0})\big| &\le \big||\widehat{\xi}|_{\widehat{g}}-|\xi|_{\widehat{g}}\big|+\big||\xi|_{\widehat{g}}-|\xi|_g\big| \\ &\le \big|\widehat{\xi}-\xi\big|_{\widehat{g}}+|\xi|_g^{-1}\big||\xi|_{\widehat{g}}^2-|\xi|_g^2\big|.\end{aligned}$$

Due to (2.11) and $\det([\langle v_j,v_k\rangle_g]_{j,k=1}^n)\ge c_1$, all components $g^{jk},\widehat{g}^{jk}$ are bounded above by $C(n,c_1)$. Hence by (2.10),

$$\big|\widehat{\xi}-\xi\big|_{\widehat{g}}\le C(n,c_1)C_4\big(d(x_{i_0},x_\ell)^{4/3}+\varepsilon_1\big). \tag{7.39}$$

Moreover, the largest eigenvalue of the matrix (g_{jk}) is bounded above by $C(n)$, and thus the eigenvalues of (g^{jk}) are bounded below by $C(n)^{-1}$. Hence by (2.11),

$$|\xi|_g^{-1}\big||\xi|_{\widehat{g}}^2-|\xi|_g^2\big| \le C(n)|X|_{\mathbb{R}^n}^{-1}\sum_{j,k=1}^n\big(\widehat{g}_{jk}-g_{jk}\big)X_jX_k\le nC(n)C_4\varepsilon_1^{\frac{1}{8}}|\xi|_g. \tag{7.40}$$

Thus by (7.39) and (7.40), we obtain

$$\begin{aligned}\big|\widehat{d}_{\ell,i_0}-d(x_\ell,x_{i_0})\big| &\le C(n,c_1)C_4\big(d(x_{i_0},x_\ell)^{4/3}+\varepsilon_1\big)+nC(n)C_4\varepsilon_1^{\frac{1}{8}}d(x_{i_0},x_\ell)\\ &\le C_5\big(d(x_{i_0},x_\ell)^{4/3}+\varepsilon_1^{1/2}\big).\end{aligned}$$

□

8. Global constructions

In this section we prove Theorem 1.2.

Proof of Theorem 1.2. (1) The first claim has been proved in Proposition 5.2(2), and we prove the second claim here. Let $\rho>0$ be a parameter which is determined later. Given two indices $i,j\in\{1,\ldots,I\}$, we consider the following minimization problem

$$\widehat{d}_{i,j}:=\min\left(\widehat{d}_{i,\sigma(1)}+\sum_{k=1}^{N-1}\widehat{d}_{\sigma(k),\sigma(k+1)}+\widehat{d}_{\sigma(N),j}\right) \tag{8.1}$$

over all chain of N indices $\sigma(k)\in\{1,\ldots,I\}$ with $N\le 1+\Lambda/\rho$, and with

$$\|\widehat{r}_{\sigma(k)}-\widehat{r}_{\sigma(k+1)}\|_{\ell^\infty(Y)}\le\rho+C_0\varepsilon_1^{1/2}+2\varepsilon_1<\rho_0,\quad \text{for } k=0,\ldots,N, \tag{8.2}$$

where we denote $\sigma(0):=i$ and $\sigma(N+1):=j$. In this minimization problem, the numbers $\widehat{d}_{\sigma(k),\sigma(k+1)}$ are determined by the data $\widehat{\mathcal{R}}_Y\cap\mathcal{B}_\infty(\widehat{r}_{\sigma(k)},\varepsilon_1^{1/4})$ in Corollary 2.5 since the condition (8.2) is valid for all k. Thus the solution of the minimization problem can be found using the given data $\widehat{\mathcal{R}}_Y$ only. Note that the numbers $\widehat{d}^{\,Y}_{j,j'}$ assumed in Corollary 2.5 can be determined by $\widehat{\mathcal{R}}_Y$ due to (3.7).

Now let us analyze the property of the solution of the minimization problem (8.1). Given any pair of points in X, consider a shortest path γ connecting these two points. On this shortest path γ, one can choose a chain of points with at most $N\le 1+\Lambda/\rho$ points such that each pair of adjacent points has distance at most ρ. Since the set X is an ε_2-net of M, we can replace this chain of points on γ by points in X, and thus each pair of adjacent points has distance at most $\rho+\varepsilon_2=\rho+C_0\varepsilon_1^{1/2}$. We require $\rho+C_0\varepsilon_1^{1/2}<\rho_0/2$ such that the condition (2.9) is satisfied by the elements in $\widehat{\mathcal{R}}_Y$ corresponding to this chain of points in X, by virtue of (5.22). Let us relabel this chain of points by $x_1,\ldots,x_N$ with endpoints x_1,x_N for convenience. Then the estimate (2.12) shows that for this particular chain of points, we have

$$\begin{aligned}E:=\Big|d(x_1,x_N)-\sum_{i=k}^{N-1}\widehat{d}_{k,k+1}\Big| &\le NC_5\big(\varepsilon_1^{1/2}+(\rho+\varepsilon_1^{1/2})^{4/3}\big)\\ &< 8\Lambda C_5\big(\varepsilon_1^{1/2}\rho^{-1}+\rho^{1/3}\big).\end{aligned}$$

We obtain a good estimate when we choose $\rho = \varepsilon_1^{3/8}$, namely

$$E < 16\Lambda C_5 \varepsilon_1^{1/8}. \tag{8.3}$$

This particular chain of points above satisfies the conditions of the minimization problem (8.1), and thus the solution of the minimization problem also satisfies (8.3) when we choose $\rho = \varepsilon_1^{3/8}$. Therefore, the solution of the minimization problem (8.1) gives us an approximate distance $\widehat{d}(x_i, x_j) := \widehat{d}_{i,j}$ that satisfies

$$\left|\widehat{d}(x_i, x_j) - d(x_i, x_j)\right| < 16\Lambda C_5 \varepsilon_1^{1/8}, \quad \forall\, i, j \in \{1, \dots, I\}. \tag{8.4}$$

This proves part (1) of Theorem 1.2.

Part (2) is a direct consequence of the estimate (8.4) and [29, Corollary 1.10]. □

9. Reconstruction of the manifold from the noisy heat kernel

In this section, we consider the reconstruction of a manifold from noisy heat kernel measurements (2.1) satisfying (2.2).

Theorem 9.1. *Let M be a closed Riemannian manifold of dimension n satisfying the bounds (1.1) with parameter Λ, and let $U = B(y_0, R)$ be a ball of radius $R > \Lambda^{-1}$. Suppose the Ricci curvature of M is non-negative. Then there exist constants $\widehat{\sigma}, C_{22} > 0$ explicitly depending only on n, Λ, such that the following holds for $0 < \sigma < \widehat{\sigma}$ and $0 < h \leq \sigma^{1/2}$.*

Let $Y = \{y_j : \ j = 0, 1, \dots, J\}$ be an h-net in the ball U. Assume that either

(i) $\{z_i : \ i = 1, \dots, I\}$ is an h-net in $M \setminus \overline{U}$, and we are given $\widehat{d}^{Y}_{j,j'}$, $j, j' = 0, 1, \dots, J$ such that $|\widehat{d}^{Y}_{j,j'} - d(y_j, y_{j'})| < h$,

or

(ii) $\{z_i : \ i = 1, \dots, I\}$ is an h-net in M.

Moreover, assume that we are given the data

$$\left\{\widetilde{G}(y_j, z_i, t) : \ i = 1, \dots, I, \ j = 0, 1, \dots, J, \ 0 < t < 1\right\} \tag{9.1}$$

which satisfy

$$e^{-\frac{\sigma}{t}} \leq \frac{\widetilde{G}(y_j, z_i, t)}{G(y_j, z_i, t)} \leq e^{\frac{\sigma}{t}}, \quad \text{for all } i = 1, \dots, I, \ j = 0, \dots, J, \ 0 < t < 1. \tag{9.2}$$

Then the given data (9.1) determine a smooth Riemannian manifold $(\widehat{M}, \widehat{g})$ that is diffeomorphic to M. Moreover, there is a diffeomorphism $F : \widehat{M} \to M$ such that

$$\frac{1}{L} \leq \frac{d_M(F(x), F(x'))}{d_{\widehat{M}}(x, x')} \leq L, \quad \text{for } x, x' \in \widehat{M}, \tag{9.3}$$

where $L = 1 + C_{22}\sigma^{1/24}$.

Proof. By Corollary 3.1 and Theorem 4.1 in [65] (see also [73]), for every $\epsilon \in (0, 1)$, $t > 0$,

$$G(y, z, t) \leq C_\epsilon v_{z,t} \exp\Big(-\frac{d^2(y, z)}{(4+\epsilon)t}\Big), \tag{9.4}$$

and

$$G(y, z, t) \geq C_\epsilon^{-1} v_{z,t} \exp\Big(-\frac{d^2(y, z)}{(4-\epsilon)t}\Big), \tag{9.5}$$

where $v_{z,t} = \mathrm{vol}^{-1}(B(z, \sqrt{t}))$, and $C_\epsilon \to \infty$ as $\epsilon \to 0$.

Taking log on both sides of (9.4), we have

$$\frac{d^2}{4} + t\log G \leq t\log(C_\epsilon v_{z,t}) + \frac{d^2}{4} - \frac{d^2}{4+\epsilon} \leq t\log(C_\epsilon v_{z,t}) + C(\Lambda)\epsilon.$$

Similarly, from (9.5),

$$\frac{d^2}{4} + t\log G \geq t\log(C_\epsilon^{-1} v_{z,t}) + \frac{d^2}{4} - \frac{d^2}{4-\epsilon} \geq t\log(C_\epsilon^{-1} v_{z,t}) - C(\Lambda)\epsilon.$$

Combining the two inequalities above, we obtain

$$\left|d^2(y,z) + 4t\log G(y,z,t)\right| \leq 4t|\log C_\epsilon| + 4t|\log v_{z,t}| + C(\Lambda)\epsilon. \tag{9.6}$$

From the noisy measurements $\widetilde{G}(y,z,t) = \eta(y,z,t)G(y,z,t)$, (9.6) and (2.2) yield that

$$\begin{aligned} \left|d^2(y,z) + 4t\log \widetilde{G}(y,z,t)\right| &\leq \left|d^2(y,z) + 4t\log G(y,z,t)\right| + 4t|\log\eta| \\ &\leq 4\sigma + 4t|\log C_\epsilon| + 4t|\log v_{z,t}| + C(\Lambda)\epsilon. \end{aligned} \tag{9.7}$$

Now we pick sufficiently small $\epsilon > 0$ such that $C(\Lambda)\epsilon < \sigma$. For small $t > 0$, we know

$$t|\log v_{z,t}| = t|\log \mathrm{vol}(B(z,\sqrt{t}))| \leq C(n)t|\log t|.$$

Thus one can pick sufficiently small $t > 0$ such that $4t|\log C_\epsilon| + 4t|\log v_{z,t}| < \sigma$. Hence,

$$\left|d^2(y,z) + 4t\log \widetilde{G}(y,z,t)\right| < 6\sigma,$$

which yields that

$$\left|d(y,z) - \sqrt{4t|\log \widetilde{G}(y,z,t)|}\right| < 6\sigma^{\frac{1}{2}}. \tag{9.8}$$

First, consider the case when the condition (i) is valid. Since $\{z_i\}$ is an h-net in $M \setminus \overline{U}$ for $h \leq \sigma^{1/2}$, (9.8) shows that the data $\widetilde{G}(y_j, z_i, t)$ for some suitable choice of t (depending on σ) give a $7\sigma^{1/2}$-approximation of the distances of the pairs (y_j, z_i). Moreover, we are already given an h-approximation $\widehat{d}^Y_{j,j'}$ of the distances of the pairs $(y_j, y_{j'})$ in Y. Since the set $X = \{z_i : \ i = 1,\dots,I\} \cup Y$ is a $(2h)$-net in M, thus the sets X, Y and the given data satisfy the conditions (a1) and (a2) with parameter $\varepsilon_1 = 7\sigma^{1/2}$.

Second, consider the case when the condition (ii) is valid. Since $\{z_i\}$ is an h-net in M for $h \leq \sigma^{1/2}$, (9.8) shows that the data $\widetilde{G}(y_j, z_i, t)$ for some suitable choice of t (depending on σ) give a $7\sigma^{1/2}$-approximation of the interior distance functions on Y. Thus the sets $\{z_i\}, Y$ and the given data satisfy the conditions (a1) and (a2) with parameter $\varepsilon_1 = 7\sigma^{1/2}$.

These considerations imply that in both cases (i) and (ii) the claim follows by applying Theorem 1.2(2). □

Finally, we obtain the uniqueness and the stability for the inverse problem for heat kernel.

Proof of Theorem 2.1. Let $X_1 = \{z_1^1,\dots,z_I^1\} \cup \{y_0^1,\dots,y_J^1\} \subset M_1$, $Y_1 = \{y_0^1,\dots,y_J^1\} \subset M_1$ and $X_2 = \{z_1^2,\dots,z_I^2\} \cup \{y_0^2,\dots,y_J^2\} \subset M_2$, $Y_2 = \{y_0^2,\dots,y_J^2\} \subset M_2$.

Due to the condition (2.3), the heat kernel data $G_2(y_j^2, z_i^2, t)$ of M_2 can be used as the noisy observations of the heat kernel of M_1 at (y_j^1, z_i^1, t). Then by the proof of Theorem 9.1, the heat kernel data $G_2(y_j^2, z_i^2, t)$ of M_2 and $d_{M_2}(y_j^2, y_{j'}^2)$ determine the distances $d_{M_1}(x,y)$ of M_1 for $(x,y) \in X_1 \times Y_1$ up to an error $7\sigma^{1/2}$. Similarly, the heat kernel data $G_2(y_j^2, z_i^2, t)$ of M_2 and $d_{M_2}(y_j^2, y_{j'}^2)$ also determine the distances $d_{M_2}(x,y)$ of M_2 for $(x,y) \in X_2 \times Y_2$ up to an error $7\sigma^{1/2}$. Thus, by enumerating the points in X_l as $\{x_i^l : \ i = 0,1,\dots,I'\}$ and the points in Y_l as $\{y_j^l : \ j = 0,1,\dots,J\}$, we see that

$$|d_{M_1}(x_i^1, y_j^1) - d_{M_2}(x_i^2, y_j^2)| < 14\sigma^{\frac{1}{2}}, \quad i = 0,1,\dots,I', \ j = 0,1,\dots J.$$

Thus, the numbers $d_{i,j} = d_{M_1}(x_i^1, y_j^1)$ can be used as the noisy distance data both for the manifold M_1 and for the manifold M_2 with an error $14\sigma^{\frac{1}{2}}$. Hence, by Theorem 1.2 we have

$$|d_{M_1}(x_i^1, x_{i'}^1) - d_{M_2}(x_i^2, x_{i'}^2)| \leq C_1(14\sigma^{\frac{1}{2}})^{\frac{1}{8}} < 2C_1\sigma^{\frac{1}{16}}, \quad i, i' = 0,1,\dots,I',$$

and there is a manifold $\widehat{M}$ such that there are L_1-bi-Lipschitz diffeomorphisms $F_1 : M_1 \to \widehat{M}$ and $F_2 : M_2 \to \widehat{M}$ with L_1 given in Theorem 1.2. Hence there is an (L_1^2)-bi-Lipschitz diffeomorphism $F = F_2^{-1} \circ F_1 : M_1 \to M_2$, which proves the claim. □

Proof of Corollary 2.2. Let $\Lambda > R^{-1}$ be such that both M_1 and M_2 satisfy the geometric bounds bounds (1.1). Consider an arbitrary $\sigma > 0$, $h = \sigma^{1/2}$. Let $\{y_j^1 : j = 0, 1, \ldots, J\}$ be an h-net in U_1 and $\{y_j^2 = \Phi(y_j^1) : j = 0, 1, \ldots, J\}$ be an h-net in U_2.

We recall that the map $\Psi : M_1 \setminus U_1 \to M_2 \setminus U_2$ is assumed only to be a bijection, and thus we need to do some additional considerations to obtain suitable h-nets on sets $M_1 \setminus U_1$ and $M_2 \setminus U_2$. To that end, let $\{\widetilde{z}_i^1,\, i = 1, 2, \ldots, I_1\}$ be an h-net in $M_1 \setminus U_1$ and $\{\widetilde{z}_i^2,\, i = 1, 2, \ldots, I_2\}$ be an h-net in $M_2 \setminus U_2$. Then we define

$$z_i^1 = \begin{cases} \widetilde{z}_i^1, & \text{for } i = 1, 2, \ldots, I_1, \\ \Psi^{-1}(\widetilde{z}_{i-I_1}^2), & \text{for } i = I_1 + 1, I_1 + 2, \ldots, I_1 + I_2, \end{cases}$$

and

$$z_i^2 = \begin{cases} \Psi(\widetilde{z}_i^1), & \text{for } i = 1, 2, \ldots, I_1, \\ \widetilde{z}_{i-I_1}^2, & \text{for } i = I_1 + 1, I_1 + 2, \ldots, I_1 + I_2. \end{cases}$$

Then $\{z_i^1 :\ i = 1, \ldots, I_1 + I_2\}$ is an h-net in $M_1 \setminus U_1$ and $\{z_i^2 :\ i = 1, \ldots, I_1 + I_2\}$ is an h-net in $M_2 \setminus U_2$, and we have

$$G_1(y_j^1, z_i^1, t) = G_2(y_j^2, z_i^2, t), \quad i = 1, \ldots, I_1 + I_2,\ j = 0, \ldots, J,\ 0 < t < 1.$$

By applying Theorem 2.1, it follows that the Gromov-Hausdorff distance of the metric spaces (M_1, d_{g_1}) and (M_2, d_{g_2}) is smaller than $C_{23}\sigma^{1/24}$, where $C_{23} > 0$ depends only on n and Λ, see e.g. Corollary 7.3.28 in [12]. Letting $\sigma \to 0$, we see that Gromov-Hausdorff distance of (M_1, d_{g_1}) and (M_2, d_{g_2}) is zero, which implies that (M_1, d_{g_1}) and (M_2, d_{g_2}) are isometric as (compact) metric spaces, see e.g. [12, 69]. By the Myers-Steenrod theorem, there is a diffeomorphism $F : M_1 \to M_2$ between Riemannian manifolds such that $g_2 = F_* g_1$. This proves the claim. □

Appendix A. Auxiliary Lemmas

Lemma A.1. *Let M be a closed Riemannian manifold with sectional curvature bounded below by $\mathrm{Sec}_M \geq -\Lambda^2$. Suppose $\gamma_{x,v_1}(t),\ \gamma_{x,v_2}(t)$ are two distance-minimizing geodesics emanating from $x \in M$ with unit initial vectors $v_1, v_2 \in S_x M$. Denote by α the angle between v_1 and v_2. Then there is a uniform constant $C_A > 1$, explicitly depending only on Λ, such that*

$$d(\gamma_{x,v_1}(t_1),\ \gamma_{x,v_2}(t_2)) \leq |t_1 - t_2| + C_A \alpha, \quad \forall\, t_1, t_2 \in [0, \Lambda]. \tag{A.1}$$

Proof. Assume that $t_1 \leq t_2$. Let us denote $a = \gamma_{x,v_1}(t_1)$ and $b = \gamma_{x,v_2}(t_1)$. We can compare the triangle axb with a triangle $\overline{axb}$ in the rescaled hyperbolic plane H with constant sectional curvature $-\Lambda^2$, satisfying that $d(x, a) = \overline{d}(\overline{x}, \overline{a}) = t_1$, $d(x, b) = \overline{d}(\overline{x}, \overline{b}) = t_1$, $\alpha = \angle \overline{axb}$. Then Toponogov's theorem yields $d(a, b) \leq \overline{d}(\overline{a}, \overline{b})$.

On the hyperbolic plane H, the exponential map is smooth everywhere, and its differential is uniformly bounded. Hence,

$$\overline{d}(\overline{a}, \overline{b}) \leq C(\Lambda)|t_1 v_1 - t_1 v_2| \leq C(\Lambda) t_1 \alpha. \tag{A.2}$$

Then,

$$\begin{aligned} d(\gamma_{x,v_1}(t_1),\ \gamma_{x,v_2}(t_2)) &\leq d(a, b) + d(\gamma_{x,v_2}(t_1),\ \gamma_{x,v_2}(t_2)) \\ &\leq \overline{d}(\overline{a}, \overline{b}) + |t_2 - t_1| \leq |t_2 - t_1| + C(\Lambda)\Lambda\alpha. \end{aligned}$$

□

Lemma A.2. *There exists a uniform constant $C_6 > 1$ such that the following holds. Let N be a compact Riemannian manifold with boundary ∂N with sectional curvature bounded below by $Sec_N \geq -\Lambda^2$. Let $a, b, c \in N$ and β be the angle of the length minimizing curves $[ab]$ and $[bc]$ at b. Then we have*

$$|ac| \leq |ab| - |bc|\cos\beta + C_6|bc|^2/\min\{\Lambda^{-1}, |ab|, d(b, \partial N)\}. \tag{A.3}$$

Proof. To prove the statement, we apply Toponogov's Theorem (e.g. [69, Thm. 79]) to the triangle abc. Below, let H be the rescaled hyperbolic plane of constant sectional curvature $-\Lambda^2$, and $\overline{a}, \overline{b}, \overline{c} \in H$ be such that $|\overline{a}\overline{b}| = |ab|$, $|\overline{b}\overline{c}| = |bc|$ and $\angle\overline{a}\overline{b}\overline{c} = \angle abc = \beta$. Here by $|xy|$, we denote the distance between points x and y in whatever space they belong.

(i) Let us first consider the case when

$$|ab| = \frac{1}{4}, \quad |bc| < \frac{1}{4}, \quad \text{and} \quad d(b, \partial N) \geq 1. \tag{A.4}$$

Then $|ac| \leq |ab| + |bc| \leq \frac{1}{2}$. Applying Toponogov's theorem to triangle abc implies that $|ac| \leq |\overline{a}\overline{c}|$.

By [69, Prop. 48], the law of cosines on H gives

(A.5)

$$\cosh(\Lambda|\overline{a}\overline{c}|) \;=\; \cosh(\Lambda|\overline{a}\overline{b}|)\cosh(\Lambda|\overline{b}\overline{c}|) - \sinh(\Lambda|\overline{a}\overline{b}|)\sinh(\Lambda|\overline{b}\overline{c}|)\cos\beta$$

and thus

$$\cosh(\Lambda|ac|) \;\leq\; \cosh(\Lambda|ab|)\cosh(\Lambda|bc|) - \sinh(\Lambda|ab|)\sinh(\Lambda|bc|)\cos\beta.$$

Using Taylor series, the above yields

$$\begin{aligned} & \cosh(\Lambda|ac|) - \cosh(\Lambda|ab|) \\ \leq\; & \cosh(\Lambda|ab|)\,(\cosh(\Lambda|bc|) - 1) - \sinh(\Lambda|ab|)\sinh(\Lambda|bc|)\cos\beta \\ =\; & -V|bc|\cos\beta + E_1, \end{aligned}$$

where $V = \Lambda\sinh(\Lambda|ab|) > 0$ and E_1 satisfies $|E_1| \leq C|bc|^2$, where C is a uniform constant.

By triangle inequality, $-|bc| \leq |ac| - |ab| \leq |bc|$. Then one can show that

$$\cosh(\Lambda|ac|) - \cosh(\Lambda|ab|) \geq V(|ac| - |ab|) + E_2,$$

where E_2 satisfies $|E_2| \leq C(|ac| - |ab|)^2 \leq C|bc|^2$, where C is a uniform constant. Combining these we see that

$$V(|ac| - |ab|) + E_2 \leq -V|bc|\cos\beta + E_1,$$

or

$$|ac| - |ab| \leq -|bc|\cos\beta + V^{-1}(E_1 - E_2),$$

which yields the inequality

$$|ac| \leq |ab| - |bc|\cos\beta + C_6|bc|^2, \tag{A.6}$$

where C_6 is a uniform constant. We can assume that $C_6 > 8$.

Consider next the case when

$$|ab| = \frac{1}{4}, \quad \text{and} \quad d(b, \partial N) \geq 1. \tag{A.7}$$

If it holds that $|bc| \geq \frac{1}{4}$, then $C_6 > 8$ yields that $C_6|bc|^2 > 2|bc|$. Considering $|ac| \leq |ab| + |bc|$, we see that (A.6) automatically holds. Since we have already proven (A.6) when $|bc| < \frac{1}{4}$, we can conclude that (A.6) holds under assumptions (A.7).

Next, consider the case when

$$|ab| \geq \frac{1}{4}, \quad d(b, \partial N) \geq 1. \tag{A.8}$$

Let a' be the point on $[ab]$ with $|a'b| = \frac{1}{4}$. By the triangle inequality we have $|ac| \le |aa'| + |a'c|$. Moreover, we have $|aa'| = |ab| - |a'b|$ and thus (A.6) for the triangle $a'bc$ implies

$$|ac| - |ab| \le |aa'| + |a'c| - |ab| \le |a'c| - |a'b| \le -|bc|\cos\beta + C_6|bc|^2.$$

Hence, (A.6) holds under assumption (A.8).

The inequality (A.6) yields that the inequality (A.3) holds. □

Lemma A.3. *Let M be a closed Riemannian manifold with sectional curvature bounded by $|\mathrm{Sec}_M| \le \Lambda^2$. Let $B(x_0, R)$ be an open ball for $R \le \min\{\mathrm{inj}(M)/2, \pi/(4\Lambda)\}$. Then there exist uniform constants $\widehat{\varepsilon}, c_3, C_{16} > 0$ explicitly depending on Λ, R, such that the following holds for $0 < \varepsilon < \widehat{\varepsilon}$.*

Given a point $x \in B(x_0, R/2)$, take z_0 to be the nearest point on $\partial B(x_0, R)$ from x. Let $z_1, z_2 \in N_\varepsilon(R)$ such that

$$|z_1z_2| \ge C_{17}\varepsilon, \quad |z_0z_1| < c_3, \quad |z_0z_2| < c_3, \tag{A.9}$$

for some $C_{17} \ge 32(\Lambda R)^{-1}$. Then $\angle z_1xz_2 > C_{16}C_{17}\varepsilon$.

Proof. The proof is similar to Lemma 6.1(1): we use the upper bound for the angle $\angle xz_1z_2$ to derive an lower bound for $\angle z_1xz_2$. First, we show that the incident angle of $[xz_1]$, i.e. the angle of $[xz_1]$ with the tangent space $T_{z_1}\Sigma_{|x_0z_1|}$ is bounded from below. Suppose $|x_0z_1| \le |x_0z_2|$.

We take the point $z_1' \in \Sigma_{3R/2}$ such that $|x_0z_1'| - |x_0z_1| = |z_1z_1'|$. Let z_0' be the nearest point in $\Sigma_{3R/2}$ from x. Then for sufficiently small c_3, ε, we have

$$|xz_1'| \ge |xz_0'| - |z_0'z_1'| \ge |xz_0| + \frac{R}{2} - 2c_3 \ge |xz_1| + |z_1z_1'| - 3c_3.$$

Hence the same argument as (6.3) gives $(\pi - \angle xz_1z_1')^2 \le C_7^{-1}3c_3$, which is

$$\angle xz_1x_0 \le C(\Lambda)c_3^{1/2}. \tag{A.10}$$

For the upper bound for the angle $\angle x_0z_1z_2$, one can apply Rauch comparison theorem to compare with the sphere of constant sectional curvature Λ^2. Namely, we take the triangle $\overline{x_0z_1z_2}$ on the sphere such that $|x_0z_1| = |\overline{x_0z_1}|$, $|z_1z_2| = |\overline{z_1z_2}|$, $\angle x_0z_1z_2 = \angle\overline{x_0z_1z_2}$. Due to Rauch comparison theorem, $|x_0z_2| \ge |\overline{x_0z_2}|$. Hence,

$$\cos(\Lambda|x_0z_2|) \le \cos(\Lambda|x_0z_1|)\,\cos(\Lambda|z_1z_2|) + \sin(\Lambda|x_0z_1|)\,\sin(\Lambda|z_1z_2|)\,\cos(\angle x_0z_1z_2).$$

Assume $\angle x_0z_1z_2 > \pi/2$. Since $|x_0z_1| \le R + \varepsilon < \pi/(2\Lambda)$, then

$$\begin{aligned}\cos(\Lambda|x_0z_2|) - \cos(\Lambda|x_0z_1|) &\le \cos(\Lambda|x_0z_1|)\big(\cos(\Lambda|z_1z_2|) - 1\big)\\ &\quad + \sin(\Lambda|x_0z_1|)\,\sin(\Lambda|z_1z_2|)\,\cos(\angle x_0z_1z_2)\\ &\le \frac{\Lambda^2R}{8}|z_1z_2|\angle x_0z_1z_2.\end{aligned}$$

On the other hand,

$$\cos(\Lambda|x_0z_2|) - \cos(\Lambda|x_0z_1|) \ge \Lambda(|x_0z_1| - |x_0z_2|).$$

Hence,

$$2\varepsilon \ge |x_0z_2| - |x_0z_1| \ge -\frac{\Lambda R}{8}|z_1z_2|\angle x_0z_1z_2,$$

which yields $\angle x_0z_1z_2 \le 2\pi/3$ due to the condition (A.9). Thus combining with (A.10), we can choose sufficiently small c_3 such that

$$\angle xz_1z_2 \le \angle xz_1x_0 + \angle x_0z_1z_2 \le \frac{5}{6}\pi. \tag{A.11}$$

Then the lemma follows from the exact same argument in the last part of Lemma (6.1)(1). □

References

[1] G. Alessandrini, *Stable determination of conductivity by boundary measurements*, Appl. Anal. **27** (1988), 153–172.

[2] G. Alessandrini, J. Sylvester, *Stability for a multidimensional inverse spectral theorem*, Comm. PDE. **15** (1990), 711–736.

[3] R. Alexander, S. Alexander, *Geodesics in Riemannian manifolds-with-boundary*, Indiana Univ. Math. J. **30** (1981), 481–488.

[4] S. Alexander, V. Kapovitch, A. Petrunin, *Alexandrov geometry: preliminary version no. 1*, arXiv:1903.08539.

[5] M. Anderson, A. Katsuda, Y. Kurylev, M. Lassas, M. Taylor, *Geometric Convergence, and Gel'fand's Inverse Boundary Problem*, Invent. Math. **158** (2004), 261–321.

[6] M. Belishev, *An approach to multidimensional inverse problems for the wave equation*, (Russian) Dokl. Akad. Nauk SSSR **297** (1987), 524–527

[7] M. Belishev, Y. Kurylev, *To the reconstruction of a Riemannian manifold via its spectral data (BC-method)*, Comm. PDE **17** (1992), 767–804.

[8] P. Bérard, G. Besson, S. Gallot, *Embedding Riemannian manifolds by their heat kernel*, Geom. Funct. Anal. **4** (1994), 373–398.

[9] R. Bosi, Y. Kurylev, M. Lassas, *Reconstruction and stability in Gel'fand's inverse interior spectral problem*, To appear in Analysis and PDE.

[10] Y. Brudnyi, P. Shvartsman, *A linear extension operator for a space of smooth functions defined on closed subsets of* $\mathbb{R}^n$, Dokl. Akad. Nauk SSSR **280** (1985), 268–270. English transl. in Soviet Math. Dokl. **31**, No. 1 (1985), 48–51.

[11] Y. Brudnyi, P. Shvartsman, *Generalizations of Whitney's extension theorem*, Int. Math. Research Notices **3** (1994), 129–139.

[12] D. Burago, Y. Burago, S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics **33**, AMS, 2001. xiv+415 pp.

[13] D. Burago, S. Ivanov, Y. Kurylev, *A graph discretization of the Laplace-Beltrami operator*, J. Spectr. Theory **4** (2014), no. 4, 675–714.

[14] D. Burago, S. Ivanov, M. Lassas, J. Lu, *Quantitative stability of Gel'fand's inverse boundary problem*, arXiv:2012.04435.

[15] I. Chavel, *Riemannian geometry–a modern introduction.* Cambridge U. Press, 1993.

[16] J. Cheeger, *Finiteness theorems for Riemannian manifolds*, Amer. J. Math. **92** (1970), 61–75.

[17] J. Cheeger, S. T. Yau, *A lower bound for the heat kernel*, Comm. Pure Appl. Math. **34** (1981), 465–480.

[18] X. Chen, M. Lassas, L. Oksanen, G. Paternain, *Detection of Hermitian connections in wave equations with cubic non-linearity*, To appear in JEMS. arXiv:1902.05711.

[19] S. Cheng, T. Dey, E. Ramos, *Manifold reconstruction from point samples*, SODA (2005), 1018–1027.

[20] R. Coifman, S. Lafon, *Diffusion maps*, Appl. Comp. Harm. Anal. **21** (2006), 5–30.

[21] R. Coifman, S. Lafon, A. Lee, M. Maggioni, B. Nadler, F. Warner, S. Zucker, *Geometric diffusions as a tool for harmonic analysis and structure definition of data: Multiscale methods*, Proc. of Nat. Acad. Sci. **102** (2005), 7432–7438.

[22] M. de Hoop, G. Uhlmann, Y. Wang, *Nonlinear responses from the interaction of two progressing waves at an interface*, Ann. de l'Inst. Henri Poincaré C, Anal. non lin. **36** (2019), 347–363.

[23] M. de Hoop, G. Uhlmann, Y. Wang, *Nonlinear interaction of waves in elastodynamics and an inverse problem*, Math. Ann. **376** (2020), 765–795.

[24] M. Di Cristo, L. Rondi, S. Vessella, *Stability properties of an inverse parabolic problem with unknown boundaries.* Mat. Pura Appl. (4) **185** (2006), 223–255.

[25] D. Dos Santos Ferreira, C. Kenig; J. Sjštrand, G. Uhlmann, *Determining a magnetic Schrödinger operator from partial Cauchy data.* Comm. Math. Phys. **271** (2007), 467–488.

[26] C. Fefferman, *A sharp form of Whitney's extension theorem*, Ann. of Math. **161** (2005), 509–577.

[27] C. Fefferman, *Whitney's extension problem for* C^m, Ann. of Math. **164** (2006), 313–359.

[28] C. Fefferman, C^m*-extension by linear operators*, Ann. of Math. **166** (2007), 779–835.

[29] C. Fefferman, S. Ivanov, Y. Kurylev, M. Lassas, H. Narayanan, *Reconstruction and interpolation of manifolds I: The geometric Whitney problem*, Found. Comp. Math. **20** (2020), 1035–1133.

[30] C. Fefferman, S. Ivanov, M. Lassas, H. Narayanan: *Reconstruction of a Riemannian manifold from noisy intrinsic distances*, SIAM J. Math. Data Science **2** (2020), No. 3, 770–808.

[31] C. Fefferman, B. Klartag, *Fitting* C^m*-smooth function to data I*, Ann. of Math. **169** (2009), 315–346.

[32] C. Fefferman, B. Klartag, *Fitting* C^m*-smooth function to data II*, Rev. Mat. Iberoam. **25** (2009), 49–273.

[33] C. Fefferman, S. Mitter, H. Narayanan, *Testing the manifold hypothesis*, JAMS **29** (2016), 983–1049.

[34] I. Gel'fand, *Some aspects of functional analysis and algebra*, Proc. Intern. Cong. Math. **1** (1954), 253–277.

[35] M. Gromov, *Filling Riemannian manifolds*, J. Diff. Geom. **18** (1983), 1–147.

[36] C. Guillarmou, L. Tzou, *Calderon inverse problem with partial data on Riemann surfaces.* Duke Math. J. **158** (2011), 83–120.

[37] T. Helin, M. Lassas, L. Oksanen, T. Saksala, *Correlation based passive imaging with a white noise source*, J. Math. Pures et Appl. **116** (2018), 132–160.

[38] T. Helin, M. Lassas, L. Ylinen, Z. Zhang, *Inverse problems for heat equation and space-time fractional diffusion equation with one measurement.* J. Diff. Eq. **269** (2020), no. 9, 7498–7528.

[39] P. Hintz, G. Uhlmann, J. Zhai, *An inverse boundary value problem for a semilinear wave equation on Lorentzian manifolds*, Int. Math. Res. Not., rnab088, 2021.

[40] P. Hintz, G. Uhlmann, J. Zhai, *The Dirichlet-to-Neumann map for a semilinear wave equation on Lorentzian manifolds*, arXiv:2103.08110.

[41] S. Ivanov, *Distance difference representations of Riemannian manifolds*, Geom. Dedicata **207** (2020), 167–192.

[42] H. Karcher, *Riemannian Comparison Constructions.* In Global differential geometry (Ed. S. S. Chern), Studies in Mathematics, Vol. 27 (1989) MAA., pp. 170–222.

[43] A. Kasue, *Convergence of Riemannian manifolds and Laplace operators. I*, Ann. Inst. Fourier **52** (2002), 1219–1257.

[44] A. Katchalov, Y. Kurylev, *Multidimensional inverse problem with incomplete boundary spectral data*, Comm. Part. Diff. Eq. **23** (1998), 55–95.

[45] A. Katchalov, Y. Kurylev, M. Lassas, *Inverse Boundary Spectral Problems.* Chapman Hall/CRC, Pure and Applied Mathematics **123**. (2001), 290pp.

[46] A. Katsuda, Y. Kurylev, M. Lassas, *Stability of boundary distance representation and reconstruction of Riemannian manifolds*, Inverse Probl. Imag. **1** (2007), 135–157.

[47] C. Kenig; J. Sjštrand, G. Uhlmann, *The Calderón problem with partial data.* Ann. of Math. **165** (2007), 567-591.

[48] R. Kohn, M. Vogelius, *Determining conductivity by boundary measurements.* Comm. Pure Appl. Math.**37** (1984),289-298.

[49] R. Kohn, M. Vogelius, *Determining conductivity by boundary measurements. II. Interior results.* Comm. Pure Appl. Math. **38** (1985), 643-667.

[50] K. Krupchyk, Y. Kurylev, M. Lassas, *Inverse spectral problems on a closed manifold*, J. Math. Pures Appl. **90** (2008), 42–59.

[51] K. Krupchyk, G. Uhlmann, *Inverse problems for advection diffusion equations in admissible geometries.* Comm. Part. Diff. Eq. **43** (2018), 585–615.

[52] Y. Kurylev, *Inverse boundary problems on Riemannian manifolds*, Contemp. Math. **173** (1994), 181–192.

[53] Y. Kurylev, M. Lassas, *Inverse problem for a Dirac-type equation on a vector bundle*, Adv. Math. **221** (2009), 170-216.

[54] Y. Kurylev, M. Lassas, L. Oksanen, G. Uhlmann, *Inverse problem for Einstein-scalar field equations.* To appear in Duke Math. J.

[55] Y. Kurylev, M. Lassas, E. Somersalo, *Maxwell's equations with a polarization independent wave velocity: direct and inverse problems*, J. Math. Pures Appl. **86** (2006), 237–270.

[56] Y. Kurylev, M. Lassas, G. Uhlmann, *Rigidity of broken geodesic flow and inverse problems*, Amer. J. Math. **132** (2010), 529–562.

[57] Y. Kurylev, M. Lassas, G. Uhlmann, *Inverse problems for Lorentzian manifolds and non-linear hyperbolic equations*, Invent. Math. **212** (2018), 781–857.

[58] Y. Kurylev, L. Oksanen, G. Paternain, *Inverse problems for the connection Laplacian*, J. Diff. Geom. **110** (2018), 457–494.

[59] M. Lassas, *Inverse problems for linear and non-linear hyperbolic equations.* Proc. Int. Congress of Math. ICM 2018, Rio de Janeiro, Brazil, Vol III, 3739–3760, 2018.

[60] M. Lassas, L. Oksanen, *Inverse problem for the Riemannian wave equation with Dirichlet data and Neumann data on disjoint sets.* Duke Math. J. **163** (2014), 1071–1103.

[61] M. Lassas, T. Saksala, *Determination of a Riemannian manifold from the distance difference functions.* Asian J. Math. **23** (2019), 173–200.

[62] M. Lassas, M. Taylor, G. Uhlmann, *The Dirichlet-to-Neumann map for complete Riemannian manifolds with boundary*, Comm. Anal. Geom. **11** (2003), 207–221.

[63] M. Lassas, G. Uhlmann, *On determining a Riemannian manifold from the Dirichlet-to-Neumann map*, Ann. Sci. Ecole Norm. Sup. **34** (2001), 771–787.

[64] M. Lassas, G. Uhlmann, Y. Wang, *Inverse Problems for Semilinear Wave Equations on Lorentzian Manifolds*, Comm. in Math. Phys. **360** (2018), 555–609.

[65] P. Li, S. T. Yau, *On the parabolic kernel of the Schrödinger operator*, Acta Math. **156** (1986), 153–201.

[66] F. Memoli. *Spectral Gromov-Wasserstein distances for shape matching.* In Proc. NORDIA, 2009.

[67] M. Ovsjanikov et al., *One Point Isometric Matching with the Heat Kernel.* Computer Graphics Forum 29 (2010):1555-1564.

[68] L. Pestov, G. Uhlmann, *Two dimensional compact simple Riemannian manifolds are boundary distance rigid*, Ann. of Math. **161** (2005), 1093–1110.
[69] P. Petersen, *Riemannian geometry*, 1st Ed. Springer, 1998. xvi+432pp.
[70] C. Plaut, *Metric spaces of curvature* $\geq k$, Handbook of geometric topology, 819–898, North-Holland, Amsterdam, 2002.
[71] J. Portegies, *Embeddings of Riemannian manifolds with heat kernels and eigenfunctions*, Comm. Pure Appl. Math. **69** (2016), 478–518.
[72] S. Roweis, L. Saul, *Nonlinear dimensionality reduction by locally linear embedding*, Science **290** (2000), 2323–2326.
[73] R. Schoen, S. T. Yau, *Lectures on differential geometry*, International Press, 1994.
[74] K. Shiohama, *An introduction to the geometry of Alexandrov spaces*, Lecture Notes Series, 8. Seoul National University, Seoul, 1993. ii+78 pp.
[75] P. Stefanov, G. Uhlmann, *Stability estimates for the hyperbolic Dirichlet to Neumann map in anisotropic media*, J. Funct. Anal. **154** (1998), 330–358.
[76] P. Stefanov, G. Uhlmann, *Boundary rigidity and stability for generic simple metrics*, J. Amer. Math. Soc. **18** (2005), 975–1003.
[77] C. Stolk, M. de Hoop, *Microlocal analysis of seismic inverse scattering in anisotropic elastic media*, Comm. Pure Appl. Math. **55** (2002), 261–301.
[78] J. Tenenbaum, V. de Silva, J. Langford, *A global geometric framework for nonlinear dimensionality reduction*, Science **290** (2000), 2319–2323.
[79] G. Uhlmann, *Inverse boundary value problems for partial differential equations*, Proceedings of the International Congress of Mathematicians, Vol. III (Berlin, 1998). Doc. Math., 77–86.
[80] G. Uhlmann, Y. Wang. *Determination of space-time structures from gravitational perturbations*, To appear in Comm. Pure App. Math.
[81] S. Varadhan, *On the behaviour of the fundamental solution of the heat equation with variable coefficients.* Comm. Pure Appl. Math. **20** (1967), 431–455.
[82] N. Varopoulos, *The Poisson kernel on positively curved manifolds*, J. Funct. Anal. **44** (1981), 359–380.
[83] Y. Wang, T. Zhou, *Inverse problems for quadratic derivative nonlinear wave equations*, Comm. PDE **44** (2019), 1140–1158.
[84] X. Wang, K. Zhu, *Isometric embeddings via heat kernel*. J. Diff. Geom. **99** (2015), 497–538.
[85] H. Whitney, *Functions differentiable on the boundaries of regions,* Ann. of Math. **35** (1934), 482–485.
[86] S. Zelditch, *Survey on the inverse spectral problem.* ICCM Not. 2 (2014), no. 2, 1-20.
[87] H. Zha, Z. Zhang, *Continuum Isomap for manifold learnings*, Comp. Stat. Data Anal. **52** (2007), 184–200.

Charles Fefferman, Princeton University, Mathematics Department, Fine Hall, Washington Road, Princeton NJ, 08544-1000, USA.

Sergei Ivanov, St. Petersburg Department of Steklov Institute of Mathematics, Russian Academy of Sciences, 27 Fontanka, 191023 St. Petersburg, Russia.

Matti Lassas and Jinpeng Lu, University of Helsinki, Department of Mathematics and Statistics, P.O. Box 68, 00014, Helsinki, Finland.

Hariharan Narayanan, School of Technology and Computer Science, Tata Institute for Fundamental Research, Mumbai 400005, India.