RECONSTRUCTION OF A RIEMANNIAN MANIFOLD FROM NOISY INTRINSIC DISTANCES

CHARLES FEFFERMAN, SERGEI IVANOV,
MATTI LASSAS, HARIHARAN NARAYANAN

TO THE MEMORY OF YAROSLAV KURYLEV.

Abstract. We consider reconstruction of a manifold, or, invariant manifold learning, where a smooth Riemannian manifold M is determined from intrinsic distances (that is, geodesic distances) of points in a discrete subset of M. In the studied problem the Riemannian manifold (M, g) is considered as an abstract metric space with intrinsic distances, not as an embedded submanifold of an ambient Euclidean space. Let $\{X_1, X_2, \dots, X_N\}$ be a set of N sample points sampled randomly from an unknown Riemannian M manifold. We assume that we are given the numbers $D_{jk} = d_M(X_j, X_k) + \eta_{jk}$, where $j, k \in \{1, 2, \dots, N\}$. Here, $d_M(X_j, X_k)$ are geodesic distances, η_{jk} are independent, identically distributed random variables such that $\mathbb{E}e^{|\eta_{jk}|}$ is finite. We show that when N is large enough, it is possible to construct an approximation of the Riemannian manifold (M, g) with a large probability. This problem is a generalization of the geometric Whitney problem with random measurement errors. We consider also the case when the information on noisy distance D_{jk} of points X_j and X_k is missing with some probability. In particular, we consider the case when we have no information on points that are far away.

Keywords: Inverse problems, Manifold learning, Geometric Whitney problem.

1. Introduction

Let M be a manifold and g an intrinsic Riemannian metric on it. Assume that one is given distances, $d_M(X_j, X_k)$, with random measurement errors, between points in a randomly sampled set $\{X_1, X_2, \dots, X_N\}$ of points of M. In this paper we ask, how one can construct a Riemannian manifold (M^*, g^*) from these data so that with a large probability, the distance (in Lipschitz-sense) of the constructed manifold (M^*, g^*) to the original Riemannian manifold (M, g) can be estimated. The need of constructing the non-Euclidean, intrinsic metric is encountered in many applications, e.g., in medical and seismic imaging, discussed in Section 1.4.

In the traditional manifold learning, for instance by using the ISOMAP algorithm introduced in the seminal paper [53], one often aims to map points X_j to points $Y_j = F(X_j)$ in an Euclidean space $\mathbb{R}^m$, where $m \geq n$ is as small as possible, so that the Euclidean distances $\|Y_j - Y_k\|_{\mathbb{R}^m}$ are close to the intrinsic distances $d_M(X_j, X_k)$ and find a submanifold $\widetilde{M} \subset \mathbb{R}^m$ that is close to the points Y_j. However, even in the ideal case when one is given an infinite set of points X_j that form a dense subset of a smooth manifold M and one has no measurement errors, finding a map $F : M \to \mathbb{R}^m$ for which the embedded manifold $F(M) = \widetilde{M} \subset \mathbb{R}^m$ is isometric to (M, g) is numerically a very difficult task as it means finding a map which existence is proved by the Nash embedding theorem, see [39, 40] and [55] on numerical techniques based on Nash embedding theorem. One can overcome this difficulty by formulating the problem in a coordinate invariant way: Given the geodesic distances of points sampled from a Riemannian manifold (M, g), construct a manifold M^* with an intrinsic metric tensor g^* so that Lipschitz distance of (M^*, g^*) to the original manifold (M, g) is small. This problem was studied in [42, 29] using diffusion maps [9, 10]. In this paper we consider this problem when distances are given with random errors and use metric geometry to construct (M^*, g^*) so that the distance of (M^*, g^*) and (M, g) can be estimated with a large probability. We emphasise that we consider M^* as an abstract manifold, that is not isometricly embedded to an Euclidean space, but where the metric is given by a metric tensor g^* that is constructed from the above data.

1.1. **The main result.** Let $n \geq 2$ be an integer, $\Lambda > 0$, $D > 0$, and $i_0 > 0$. Let (M, g) be a compact Riemannian manifold of dimension n such that

$$i)\ \|\mathrm{Sec}_M\|_{L^\infty(M)} \leq \Lambda^2, \quad ii)\ \mathrm{diam}\,(M) \leq D, \quad iii)\ \mathrm{inj}\,(M) \geq i_0, \tag{1.1}$$

where Sec_M is the sectional curvature of (M, g), $\mathrm{diam}\,(M)$ is the diameter of (M, g) and $\mathrm{inj}\,(M)$ is the injectivity radius of (M, g), that is, the minimal radius of Riemannian normal coordinates. Let $d_M(x, y)$ denote the intrinsic (or geodesic) distance of the points $x, y \in M$ determined by the metric tensor g corresponding to the line element $ds^2 = g_{jk}(x)dx^j dx^k$. Here and below, we use Einstein's summation convention and sum over indexes appearing as super and sub-indexes.

Let $(\Omega, \Sigma, \mathbb{P})$ be a complete probability space, $\mathcal{B}$ be the σ-algebra of Borel sets on M, and $\mu : \mathcal{B} \to [0, 1]$ be a probability measure on M. Let dV_g be Riemannian volume on (M, g). Assume that the Radon-Nikodym derivative of μ satisfies

$$0 < \rho_{min} \leq \frac{d\mu}{dV_g} \leq \rho_{max}, \quad \text{where } \rho_{min}, \rho_{max} \in \mathbb{R}_+. \tag{1.2}$$

Definition 1.1. Let X_j, $j = 1, 2, \dots, N$ be independent, identically distributed (i.i.d.) random variables having distribution μ. Let $\sigma > 0$, $\beta > 1$, and η_{jk} be random variables satisfying

$$\mathbb{E}\eta_{jk} = 0, \quad \mathbb{E}(\eta_{jk}^2) = \sigma^2, \quad \mathbb{E}e^{|\eta_{jk}|} = \beta. \tag{1.3}$$

We assume that all random variables η_{jk} and X_j are independent. Let

$$D_{jk} = d_M(X_j, X_k) + \eta_{jk}. \tag{1.4}$$

be the geodesic distances of points X_j and X_k measured with errors η_{jk}.

Note that the above assumptions are satisfied when $\eta_{jk} \sim N(0, \sigma^2)$ are i.i.d. Gaussian random variables and $\beta \leq 2e^{\sigma^2}$. We are mostly interested in a case when σ is fixed and N is large.

Definition 1.2. The partial data is given by

$$\overline{D}_{jk} = D_{jk}^{(\text{partial data})} = \begin{cases} D_{jk} & \text{if } Y_{jk} = 1, \\ \text{`missing'} & \text{if } Y_{jk} = 0, \end{cases}$$

where Y_{jk} are random variables taking values in $\{0, 1\}$ and $j, k \in \{1, 2, \dots, N\}$. We assume that Y_{jk} are independent of random variables $X_{j'}$ for all $j' \in \{1, 2, \dots, N\} \setminus \{j, k\}$ and of $\eta_{j''k''}$ for all j'' and k''. Above, the value 'missing', can be replaced by a large real value, e.g. by $2D$.

Assume that the conditional probability of the event $\{Y_{jk} = 1\}$, when X_j and X_k are known, is

$$\mathbb{P}(Y_{jk} = 1 \,|\, X_j, X_k) = \Phi(X_j, X_k). \tag{1.5}$$

More precisely, when $\mathcal{B}_{jk} \subset \Sigma$ is the σ-algebra generated by the random variables X_j and X_k, above in formula (1.5) we use notation $\mathbb{P}(Y_{jk} = 1 \,|\, X_j, X_k) = \mathbb{P}(Y_{jk} = 1 \,|\, \mathcal{B}_{jk})$. Here, $\Phi : M \times M \to [0, 1]$ is a measurable function such that there is a function $\Phi^1 : [0, \infty) \to [0, 1]$ so that $s \mapsto \Phi^1(s)$ is non-increasing and

$$\Phi^1(0) = \phi_0, \quad \|\Phi^1\|_{C^1(\mathbb{R})} \leq H, \quad c_1\Phi^1(d_M(x, y)) \leq \Phi(x, y) \leq c_2\Phi^1(d_M(x, y)), \quad x, y \in M, \tag{1.6}$$

where $0 < c_1 < 1 < c_2$ and $\phi_0 \in \mathbb{R}_+$.

Below, for $t \in \mathbb{R}$ we denote by $\lfloor t \rfloor$ the largest integer m such that $m \leq t$.

We will show that probabilistic considerations involving the above data, combined with the deterministic results in [24] (where we considered small deterministic errors) and Appendix A, yield that one can construct a smooth manifold (M^*, g^*) that approximates the original manifold (M, g). The proofs of Theorem 1 below and the results in [24] give a procedure, which the output is a submanifold $M^* \subset \mathbb{R}^d$ (where d depends only on n, D, Λ, and i_0) and a metric tensor g^* on M^*.

Theorem 1. *Let $n \geq 2$, $D, \Lambda, i_0, \rho_{min}, \rho_{max}, \sigma, \beta, c_1, c_2, H, \phi_0 > 0$ be given. Then there are $\delta_0 > 0$, and $C_0 > 0$, depending on $n, D, \Lambda, i_0, \rho_{min}, \rho_{max}, \sigma, \beta, c_1, c_2, H, \phi_0$, and there is $C_1 > 0$, depending on n, such that the following holds for $\theta \in (0, \frac{1}{2})$.*

Let M be a compact n-dimensional manifold satisfying (1.1), $0<\delta<\delta_0$, and

$$N=\Big\lfloor C_0\delta^{-3n}\Big(\log^2(\frac{1}{\theta})+\log^8(\frac{1}{\delta})\Big)\Big\rfloor$$

and $\overline{D}_{jk}$, $j,k=1,2,\dots,N$ be as in Definitions 1.1 and 1.2. Suppose that one is given samples of the random variables $\overline{D}_{jk}$ for $j,k=1,2,\dots,N$. Then with a probability larger than $1-\theta$ one can construct a compact, smooth n-dimensional Riemannian manifold (M^*,g^*) that approximates the manifold (M,g) in the following way:

(1) *There is a diffeomorphism $F:M^*\to M$ satisfying*

$$\frac{1}{L}\le\frac{d_M(F(x),F(y))}{d_{M^*}(x,y)}\le L,\quad \textit{for all } x,y\in M^*, \tag{1.7}$$

where $L=1+C_1\delta$, that is, the Lipschitz distance of the metric spaces (M^*,g^*) and (M,g) satisfies $d_{Lip}((M^*,g^*),(M,g))\le\log L$.

(2) *The sectional curvature Sec_{M^*} of M^* satisfies $|Sec_{M^*}|\le C_1\Lambda^2$.*

(3) *The injectivity radius $inj(M^*)$ of M^* satisfies*

$$inj(M^*)\ge\min\{(C_1\Lambda)^{-1},(1-C_1\delta)\,inj(M)\}.$$

We note that the knowledge of the authors, the results are new also in the case when there is no missing data, that is, $\Phi(x,y)=1$ for all $x,y\in M$.

Remark 1.3. Theorem 1 concerns the regime where the noise level σ is a fixed constant, the number of points N is large, and we are interested in the situation where we want the probability θ of a wrong final reconstruction to be very small. This is reflected by the fact that the probability θ of obtaining a wrong reconstruction appears only in the logarithmic term $\log(\theta^{-1})$.

1.2. **Idea of the proof and three nets of points on the manifold.** Let us assume that $N=N_0+N_1+N_2$, where $N_0,N_1,N_2\in\mathbb{Z}_+$. We are interested in the case when $N_2>N_1>N_0$. We consider the random set $S_0=\{X_1,\dots,X_{N_0}\}$ as a coarse net on M and compute approximate distances between the points in the net S_0 by using the auxiliary nets $S_1=\{X_{N_0+1},\dots,X_{N_0+N_1}\}$ and $S_2=\{X_{N_0+N_1+1},\dots,X_{N_0+N_1+N_2}\}$. These random sets correspond to the index sets

$$I^{(0)}=\{1,2,\dots,N_0\},\quad I^{(1)}=\{N_0+1,\dots,N_0+N_1\},\quad I^{(2)}=\{N_0+N_1+1,\dots,N_0+N_1+N_2\}.$$

Recall that X_j are independent random variables, taking values on M, with the distribution μ.

Below, we say that a set $Y\subset M$ is $\delta-$dense in M if for all $p\in M$ there is $y\in Y$ such such that $d_M(p,y)<\delta$. A $\delta-$dense subset S of M is often called a $\delta-$net.

Let us give an overview of the ideas on the proof of Theorem 1.

First, we use the "densest net" S_2 to compute approximately the numbers

$$k_\Phi(y,z)=\int_M|d_M(y,x)-d_M(z,x)|^2\,\Phi(y,x)\Phi(x,z)\,d\mu(x), \tag{1.8}$$

for y and z in the "medium dense net" S_1, see Prop. 3.5. This corresponds to taking average of function $|d_M(y,x)-d_M(z,x)|^2$ over all those sample points $x\in S_2$ for which the distances $d_M(y,x)$ and $d_M(z,x)$ are not missing, see (3.12).

Note that when the product $\Phi(y,x)\Phi(x,z)$ is small, there are only a small amount of sample points $x\in S_2$ for which the value of the function $|d_M(y,x)-d_M(z,x)|^2$ can be computed, and then the estimator for the function $k_\Phi(y,z)$ is not reliable. Thus the reliability of the estimator for the function $k_\Phi(y,z)$ is measured by

$$A_\Phi(y,z)\;=\;\int_M\Phi(y,x)\Phi(x,z)\,d\mu(x). \tag{1.9}$$

Indeed, when $A_\Phi(y,z)$ is larger than some threshold value $b>0$, the obtained estimator for the function $k_\Phi(y,z)$ is reliable with a large probability. When we compute an estimator for the function $k_\Phi(y,z)$ using a sampling imitating the integral in (1.8), we can compute also an estimator for $A_\Phi(y,z)$, see (3.13).

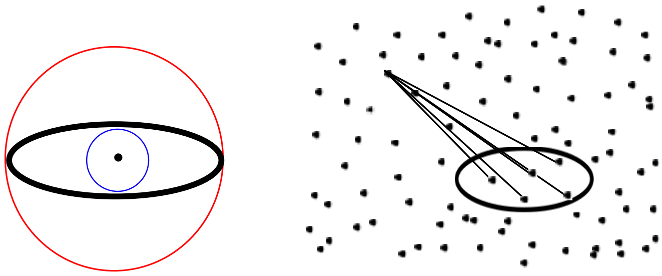

Figure 1. *Left: The set* $D_\Phi(y_1,\rho)$ *satisfies* $B_M(y,c_5^{-1}\rho)\subset D_\Phi(y_1,\rho)\subset B_M(y,\rho)$ *and thus* $D_\Phi(y_1,\rho)$ *can be considered as an approximate* ρ*-neighbourhood of the the point* y*. Right: The approximate distance* $d^{app}(y_1,y_2)$ *in the formula (1.12) is the average of distances from* y_2 *to the points in the neighbourhood* $D_\Phi(y_1,\rho)$*. Later, we approximate* $d^{app}(y_1,y_2)$ *by taking the average of distances of* y_2 *to the points in* $S_1\cap D_\Phi(y_1,\rho)$*, where* S_1 *is the medium dense net of sample points.*

Second, we are going to use the set S_1 to compute the approximate distances $d^{app}(y_1,y_2)$ of the points y_1 and y_2 in the "coarse net" S_0 using reliable distances $k_\Phi(y,z)^{1/2}$. We do this by computing estimators for the functions (see Definition 4.2)

$$
\begin{aligned}
(1.10)\qquad & Q(y_1,y_2)=\frac{V_\Phi(y_1,y_2)}{W_\Phi(y_1,y_2)},\\
& V_\Phi(y_1,y_2)=\int_M \beta_1(\frac{A_\Phi(y_1,z)}{b})\,\psi_\rho(k_\Phi(y_1,z))\Phi(z,y_2)d_M(z,y_2)d\mu(z),\\
& W_\Phi(y_1,y_2)=\int_M \beta_1(\frac{A_\Phi(y_1,z)}{b})\psi_\rho(k_\Phi(y_1,z))\Phi(z,y_2)d\mu(z),
\end{aligned}
$$

where $\beta_1\in C^\infty(\mathbb{R})$ is a cut-off function such that $\beta_1(t)=0$ for $t<1$ and $\beta_1(t)=1$ for $t>2$, and $\psi_\rho\in C^\infty(\mathbb{R})$ is a cut-off function such that $\psi_\rho(s)=1$ for $s<\rho^2$ and $\psi_\rho(s)=0$ for $s>2\rho^2$. Here, $\beta_1(A_\Phi(y_1,z)/b)\psi_\rho(k_\Phi(y,z))$ is the smoothened version of the indicator function of the set

$$
(1.11)\qquad D_\Phi(y_1,\rho)=\{z\in M:\ k_\Phi(y,z)<\rho^2,\ A_\Phi(y_1,z)\geq b\}.
$$

The set is $D_\Phi(y_1,\rho)$ is a Lipschitz approximation the union of the ball $B_M(y_1,\rho)$.

Then, roughly speaking, we compute an estimator for the function $V_\Phi(y_1,y_2)$ computing averages of function $\beta_1(A_\Phi(y_1,z)/b)d_M(z,y_2)$ over all sample points z the medium net S_1 that are in the set $D_\Phi(y_1,\rho)$ and for which data on the distance $d_M(z,y_2)$ is not missing. At the same time, we compute an estimator for the function $W_\Phi(y_1,y_2)$ by computing averages of function $\beta_1(A_\Phi(y_1,z)/b)\psi_\rho(k_\Phi(y,z))$ over the same sample points. The idea is that when $W_\Phi(y_1,y_2)$ is larger than some threshold $u>0$, the estimators computed from random data for the functions $V_\Phi(y_1,y_2)$, $W_\Phi(y_1,y_2)$, and $Q(y_1,y_2)$ are reliable with a large probability. Then, we define for $y_1,y_2\in S_0$

$$
(1.12)\qquad d^{app}(y_1,y_2)=\begin{cases} Q(y_1,y_2), & \text{if } W_\Phi(y_1,y_2)>u,\\ D, & \text{otherwise.}\end{cases}
$$

Then there is r_1 such that $d^{app}(y_1,y_2)$ approximate the true distance $d_M(y_1,y_2)$ with a small error ε_1 when $d_M(y_1,y_2)<r_1$, and moreover, if $d_M(y_1,y_2)\geq r_1$, then $d^{app}(y_1,y_2)>r_1-\varepsilon_1$. In other words, with a large probability, we can construct the distances $d_M(y_1,y_2)$ with small errors for all points y_1 and y_2 in S_0 that close to each other.

After the above constructions, we will use Proposition 4.10 in Appendix A, concerning a reconstruction of a Riemannian manifold when we are given distances with small (deterministic) errors. This result is an improved version of the earlier results given in [24, Corollary 1.10].

1.3. **Earlier results for submanifolds of $\mathbb{R}^n$ and graphs of functions.** In dimensionality reduction and in the traditional manifold learning, the aim is to transform data, consisting of points in a d-dimensional space that are near an n-dimensional submanifold M, where $d>>n$ into a set of points in a low dimensional space $\mathbb{R}^m$ close to an n-dimensional submanifold, where $d>m\geq n$.

During transformation all of them try to preserve some geometric properties, such as appropriately measured distances between points of the original data set, see [7, 8, 53]. Perhaps the most basic of such methods is 'Principal Component Analysis' (PCA), [44, 50] where one projects the data points onto the span of the n eigenvectors corresponding to the n largest eigenvalues of the $(d \times d)$ covariance matrix of the data points.

In the case of 'Multi Dimensional Scaling' (MDS) [11], the pairwise distances between points are attempted to be preserved. One minimizes a certain 'stress function' which captures the total error in pairwise distances between the data points and between their lower dimensional counterparts. For instance, given points $(x_j)_{j=1}^N$, $x_j \in \mathbb{R}^d$, one tries to find $(y_j)_{j=1}^N$, $x_j \in \mathbb{R}^m$, that is an (approximate) minimizer of a stress function

$$\min_{y_j \in \mathbb{R}^m} \big(\sum_{i,j=1}^N (\|y_i - y_j\|_{\mathbb{R}^m} - d_{ij})^2 \big), \tag{1.13}$$

where $d_{ij} = \|x_i - x_j\|_{\mathbb{R}^d}$ are the Euclidean distances of points x_i and x_j.

'ISOMAP' [53] attempts to improve on MDS by trying to capture geodesic distances between points while projecting. For each data point x_i in the data set $\mathcal{X} = (x_j)_{j=1}^N$, $x_j \in \mathbb{R}^d$, a 'neighbourhood graph' is constructed using the K-neighbours of x_i, that is, the K nearest points of $\mathcal{X}$ to x_i, the edges carrying the length between points. Now the shortest distance between points is computed in the resulting global graph containing all the neighbourhood graphs using a standard graph theoretic algorithm such as Dijkstra's. Let $D^G = [d_{ij}^G]$ be the $N \times N$ matrix of graph distances. Then MDS is used to find $(y_j)_{j=1}^N$, $x_j \in \mathbb{R}^m$ that (approximately) solve the minimization problem (1.13) with distances d_{ij} replaced by d_{ij}^G. If the data set $\mathcal{X} = (x_j)_{j=1}^N$ consists of δ-dense set points of a submanifold $M \subset \mathbb{R}^d$ with small δ, then ISOMAP tries to find an approximation for isometric embedding, that is, a map $F : M \to \mathbb{R}^m$ for which

$$\|F(x) - F(y)\|_{\mathbb{R}^m} \approx d_M(x,y), \quad x, y \in M, \tag{1.14}$$

where $d_M(x,y)$ is the intrinsic distance of the points x and y of the isometrically embedded manifold $M \subset \mathbb{R}^d$. Observe that one can not have equality $\|F(x) - F(y)\|_{\mathbb{R}^m} = d_M(x,y)$ in (1.14) unless all unit speed geodesics of M are mapped to Euclidean lines (parametrized by the Euclidean length) in the map F, implying that M has to be a flat manifold. Thus, even if the data points are contained in a submanifold M and the number N of data points grows and if d_{ij} are equal to the intrinsic distances $d_M(x_i, x_j)$, the ISOMAP algorithm does not produce a manifold which intrinsic distances are the same as those of the original manifold M when the manifold M has non-zero curvature. The convexity and flatness conditions that guarantee that the ISOMAP algorithm reconstructs the original manifold are studied in [3, 16, 18] for ISOMAP and in [56] for the continuum ISOMAP.

In the seminal papers [9, 10] on 'Diffusion Maps', a complete graph is built on the data points sampled from a manifold (M, g) and each edge is assigned a weight $a(x,y)$ that is a Gaussian function of the distance of the points x and y. The normalized version of the kernel $a(x,y)$ defines a diffusion operator on M and this operator has eigenfunctions $\phi_j(x)$. These functions can be used to construct a non-isometric embedding $x \mapsto (\phi_j(x))_{j=1}^m$ of the manifold M into $\mathbb{R}^m$. This construction is continued in [42] by computing an approximation the metric tensor g by using finite differences to find the Laplacian of the products of the local coordinate functions. When there are no errors in the data, this construction is shown in [29] to converge to a correct limit as the number of points sampled from the manifold M tends to infinity. Other topological embedding methods for manifolds, based on heat kernels and eigenfunctions, have been developed in [2, 17, 30]. Moreover, locally linear construction methods are studied in [4, 36, 37, 48, 57].

The construction of a surface approximating a set of points in $\mathbb{R}^m$ is closely related to the classical Whitney's problem. This problem is the construction of a function $F(x) \in C^m(\overline{S})$, where $S \subset \mathbb{R}^n$ is open, which is equal to a given function $f(x)$ on K, where $K \subset S$. This problem has been studied in different norms in [20, 21, 22, 25] and the interpolation results on Whitney problem have been applied for manifold of a submanifold of $\mathbb{R}^m$ in [23, 26].

1.4. **Applications of new results with intrinsic distances.** *A. Submanifolds of Banach spaces.* In several imaging problems the images are considered as elements of Banach spaces that have no inner product structure. For example, in many applications the 2-dimensional images (corresponding photographs) or 3-dimensional objects, are modeled by functions $u : D \to \mathbb{R}$, $D \subset \mathbb{R}^m$, $m = 2, 3$ that are in the space of functions of bounded variation, $BV(D)$ or in Besov spaces $B^s_{11}(D)$, [49, 34]. For example, in medical imaging, such functions are used to model piecewise constant or piecewise smooth functions that correspond to the structure of the human body with internal organs with sharp jumps of density at the boundaries of the organs.

B. Physical models with non-Euclidean metric. Many inverse problems can be formulated as geometric problems where the goal is to determine the underlying manifold structure. For example, by probing a medium with waves one can measure the travel times between points. This defines a non-Euclidean metric called the travel time metric g. Recovering the wave speed function inside the medium is equivalent to the determination of a Riemannian manifold from external or boundary measurements. The relation of the boundary measurements to the distances between the points in an δ-net in the interior of the manifold is considered e.g. in [1, 15, 32, 33]. Determining the wave speed of elastic waves inside a body is a central problem in seismic imaging of the Earth, [52, 54]. An example, in medical imaging, where the physical structures are represented using abstract Riemannian manifolds is ultrasound imaging where the acoustic properties of the inside of a body is imaged. The typical ultrasound images correspond to image of the body represented in the non-Euclidean travel time coordinates, of more precisely, in the Riemannian normal coordinates of the travel time metric [46]. In these coordinates, the image rays, that is, the geodesics from the location of the source device are straight lines. Theorem 1 can be applied when we can measure travel times of waves between points. For instance, for the earthquakes the travel times of the surface waves can be directly measured and one can use these data to deduce the properties of the Earth close to the surface. Another example is the Magnetic Resonance Imaging (MRI) based Elastography in medical imaging, where the elastic properties of the body of a patient are determined by observing the propagation of elastic waves sent into the body [28].

Models involving local and missing data are encountered in sensor technology, e.g. in the radio frequency identification (RFID) or in smart dust sensors, where a large number of low quality sensor send signals, either to receivers or to each others. On the earlier results on missing data in manifold learning, see e.g. [6, 19].

1.5. **Notations.** For a Riemannian manifold (M, g) satisfying (1.1), let vol_g be the Riemannian volume on (M, g), and $p \in M$, the exponential map $\exp_p : T_pM \to M$ defines a smooth surjective map $\exp_p : \{\xi \in T_pM : \ \|\xi\|_g < D + 1\} \to M$. By Bishop-Gromov inequality [45, Chapter 9, Lem. 1.6], the function $r \mapsto \mathrm{vol}_g(B(x, r))/v(n, -\Lambda^2, r)$ is non-increasing and bounded by 1, where $\mathrm{vol}_g(B(x, r))$ is the volume of the ball $B(x, r) \subset M$ and $v(n, -\Lambda^2, R)$ is the volume of the ball of radius r in the hyperbolic space of dimension n having curvature $-\Lambda^2$. Hence

$$\mathrm{vol}_g(M) \le V_0 = v(n, -\Lambda^2, D) \le \omega_n \Big(\frac{\sinh(\Lambda D)}{\Lambda D}\Big)^{n-1} D^n \tag{1.15}$$

where ω_n is the volume of the unit ball in $\mathbb{R}^n$, see [45, Ch. 6, Cor. 2.4]. Moreover,

$$\frac{\mathrm{vol}_g(B(x,\rho))}{\mathrm{vol}_g(M)} \ge \frac{v(n, -\Lambda^2, \rho)}{v(n, -\Lambda^2, D)} \ge \widehat{c}_3 \rho^n, \quad \mu(B(x,\rho)) \ge c_3\rho^n, \quad \widehat{c}_3 = \frac{\omega_n}{V_0}, \quad c_3 = \frac{\rho_{min}}{\rho_{max}} \widehat{c}_3, \tag{1.16}$$

where we use the fact that $v(n, -\Lambda^2, \rho) \ge \omega_n \rho^n$. Let

$$r_0 = \min(\frac{1}{2H}\phi_0, i_0, \frac{\pi}{2\Lambda}), \quad r_1 = \frac{1}{2} r_0, \quad \phi_1 = \frac{c_1}{2}\phi_0. \tag{1.17}$$

Then Definition 1.2 implies

$$\Phi(x, y) \ge c_1 \Phi^1(r_0) \ge \phi_1, \quad \text{for } d(x, y) \le r_0. \tag{1.18}$$

Acknowledgements. The authors express their gratitude to the Fields Institute, where parts of this work have been done. Ch.F. was partly supported by AFOSR, grant DMS-1265524, and NSF, grant FA9550-12-1-0425. S.I. was partly supported by RFBR, grants 14-01-00062 and 17-01-00128-A, M.L. was supported by AF, grants 284715, 312110. and H.N. was partly supported by NSF grant DMS-1620102 and a Ramanujan Fellowship.

2. Reformulation of the main result with several parameters

Our aim is to prove the following result that yields of Theorem 1 when it is combined with the results in Appendix A on manifold reconstruction with small deterministic errors.

Theorem 2. *Let $n \geq 2$ and $D, \Lambda, i_0, \rho_{min}, \rho_{max}, c_0, c_1, H, \sigma, \beta > 0$ be given. Then there are $C_2 > 1$, $\widehat{\varepsilon}_1 < 1$, and $\widehat{\delta}_1 < 1$ depending on $n, D, \Lambda, i_0, \rho_{min}, \rho_{max}, c_0, c_1, H, \sigma, \beta$, such that the following holds for $\varepsilon_1 \leq \widehat{\varepsilon}_1, \delta_1 \leq \widehat{\delta}_1$, and $\theta \in (0, 1/2)$: Let*

$$N_0 = \lfloor C_2 \delta_1^{-n} (\log(\frac{1}{\theta}) + \log(\frac{1}{\delta_1}))\rfloor, \quad N \geq N_0 + \lfloor C_2 \varepsilon_1^{-2n} \Big(\log^2(\frac{1}{\theta}) + \log^2(\frac{1}{\delta_1}) + \log^8(\frac{1}{\varepsilon_1})\Big) \rfloor. \tag{2.1}$$

Also, let X_j and $\overline{D}_{jk}$, $j, k = 1, 2, \dots, N$ be as in Definitions 1.1 and 1.2. Suppose that we are given samples of the random variables $\overline{D}_{jk}$ for $j, k = 1, 2, \dots, N$. Let r_1 be given in (1.17). Then with a probability larger than $1 - \theta$ the set $\{X_j : j = 1, 2, \dots, N_0\}$ is a δ_1-net in M and one can determine the approximate distances $d^{(a)}(X_j, X_{j'})$ so that the following holds:

For all $j, j' \in \{1, 2, \dots, N_0\}$,

$$|d^{(a)}(X_j, X_{j'}) - d_M(X_j, X_{j'})| \leq \varepsilon_1, \quad \textit{if } d_M(X_j, X_{j'}) < r_1, \tag{2.2}$$

$$d^{(a)}(X_j, X_{j'}) \geq r_1 - \varepsilon_1, \quad \textit{if } d_M(X_j, X_{j'}) \geq r_1. \tag{2.3}$$

In the case when the probability $\Phi(x, y)$, for that the information on the distance of x and y is not missing, is bounded from below by a positive constant, that is, $\Phi(x, y) \geq \phi_2$, for all $(x, y) \in M \times M$, the inequality (2.2) holds for all $X_j, X_{j'}$, $j, j' \in \{1, 2, \dots, N_0\}$.

We note when $\Phi(x, y)$ is bounded from below by a positive constant, $\Phi(x, y) \geq c_1 \phi_0$, we can choose the function Φ^1 to be equal to the constant ϕ_0. Without loss of generality, we can assume that in this case $\phi_2 = c_1 \phi_0$.

2.1. Probability that the sample points form a dense net.

First we estimate the probability that the set $S_0 = \{X_1, \dots, X_{N_0}\}$ is a δ_1-net, using standard methods based on the collectors problem.

Lemma 2.1. *There is $C_3 \geq 10$ such if $\theta \in (0, \frac{1}{2})$, $\delta_1 \in (0, \frac{D}{2})$, and*

$$N_0 \geq C_3 \delta_1^{-n} (\log(\delta_1^{-1}) + \log(\theta^{-1})), \tag{2.4}$$

then the probability that the set $\{X_j : j = 1, 2, \dots, N_0\}$ is a δ_1-net in M is larger than $1 - \frac{1}{2}\theta$.

Proof. Recall that by (1.15), we have $\mathrm{vol}_g(M) \leq V_0$. Also, by (1.16), the $\mu-$volume of a metric ball $B_M(x, \delta_1/6) \subset M$ is bounded from below by $c_3(\delta_1/6)^n$.

Let $\{z_1, \dots, z_m\}$ be a maximal $(\delta_1/3)$-separated subset in M. Then

$$m \leq m_0 = 1/(c_3(\delta_1/6)^n) = 6^n c_3^{-1} \delta_1^{-n} = C_5 \delta_1^{-n}.$$

Let $V_k = \{y \in M : \mathrm{dist}(y, z_k) < \mathrm{dist}(y, z_j) \text{ for all } j \neq k\}$ be the open Voronoi sets corresponding to points z_k and let W_k, $k = 1, 2, \dots, m$, be such disjoint sets that $V_k \subset W_k \subset \overline{V}_k$ and that the union of the sets W_k is M. Note that then the balls $B_M(z_j, \delta_1/6)$ are disjoint and thus there is $C_6 = C_6(n, \Lambda, D, i_0, \rho_{max}, \rho_{min})$ so that

$$\mu(W_k) \geq c_3(\delta_1/6)^{-n} \geq 1/(C_6 m)$$

for all $k = 1, 2, \dots, m$. Observe that $\mathrm{diam}\,(W_k) < \delta_1$, and that if for all $k = 1, 2, \dots, m$ there is X_j, $j \leq N_0$ such that $X_j \in W_k$, then the set $\{X_j : j = 1, 2, \dots, N_0\}$ is a δ_1-net in M.

We can use the classical collectors problem to estimate the probability of the event A_{m,N_0} that all sets W_k contain at least one point X_j, $j = 1, 2, \dots, N_0$. The tail estimates are used to give a solution for this problem, and for the convenience of the reader we give the details of this below (see also [12, 13, 41] for related results).

Let us choose an infinite sequence of i.i.d. random variables $X_1, X_2, \dots$ having distribution μ on M and let T be the smallest number such that all sets W_k, $k = 1, 2, \dots, m$ contain at least one point X_j, $j = 1, 2, \dots, T$. Let E_i^r denote the event that the i-th set W_i does not contain any of the first r points $X_1, \dots, X_r$. Let $m_1 = C_6 m_0 \geq C_6 m$ and $b > 1$. Then

$$\mathbb{P}\left[E_i^r\right] \leq \left(1 - \frac{1}{C_6 m}\right)^r \leq \left(1 - \frac{1}{m_1}\right)^r \leq e^{-r/m_1}.$$

For $r = \lfloor b m_1 \log m_1 \rfloor + 1$, we have $\mathbb{P}\left[E_i^r\right] \leq e^{(-b m_1 \log m_1)/m_1} = m_1^{-b}$. Then,

$$\mathbb{P}\left[T > b m_1 \log m_1\right] = \mathbb{P}\left[\bigcup_{i=1}^m E_i^r\right] \leq \frac{m_1}{C_6} \cdot \mathbb{P}[E_1^r] \leq \frac{1}{C_6} m_1^{-b+1}.$$

Hence, when $b = 1 + \frac{\log(2(C_6\theta)^{-1})}{\log m_1}$, we have

$$\mathbb{P}\left[T > b m_1 \log m_1\right] \leq \frac{1}{C_6} m_1^{-b+1} \leq \frac{\theta}{2}.$$

Observe that $m_1 = C_6 C_5 \delta_1^{-n}$ and

$$b m_1 \log m_1 = \left(1 + \frac{\log(2(C_6\theta)^{-1})}{\log(C_6 C_5 \delta_1^{-n})}\right) C_6 C_5 \delta_1^{-n} \log(C_6 C_5 \delta_1^{-n}) = (\log(C_6 C_5 \delta_1^{-n}) + \log(2(C_6\theta)^{-1})) C_6 C_5 \delta_1^{-n}.$$

Therefore, when (2.4) is valid with a suitable C_3, we have $N_0 > b m_1 \log m_1$. This implies $\mathbb{P}(A_{m,N_0}) = \mathbb{P}(T \leq N_0) \geq 1 - \frac{\theta}{2}$. □

3. The modified L^2-norm of the differences of the distance functions

Let $y, z \in M$ be (deterministic) points on M. Denote

$$\Phi_y(x) = \Phi(x, y), \quad \Phi_y^1(x) = \Phi^1(d_M(x, y)).$$

Definition 3.1. For $x \in M$, let $r_x \in C(M)$ we define the distance function

$$r_x(y) = d_M(x, y), \quad y \in M.$$

For $y, z \in M$, let

$$k_\Phi(y, z) = \|(r_y - r_z)\Phi_y^{1/2}\Phi_z^{1/2}\|^2_{L^2(M, d\mu)}, \quad A(y, z) = \int_M \Phi_y(x)\Phi_z(x) d\mu(x). \tag{3.1}$$

The map $R : M \to L^\infty(M)$, given by $R(x) = r_x$, defines an isometric embedding $R : M \to R(M) \subset L^\infty(M)$. Below we will consider this map in different functions spaces. The function $A(y, z)$ measures the relative density of the points $x \in M$ for which the both distances $d_M(x, y)$ and $d_M(x, z)$ are non-missing in the data that is given to us. In the next lemmas we analyze these functions. Recall that $r_1 = r_0/2$.

Lemma 3.2. *If $d_M(y, z) \leq r_0$ then*

$$A(y, z) = \int_M \Phi_y(x)\Phi_y(x) d\mu(x) \geq c_4 = c_3 \phi_1^2 r_1^n.$$

Proof. Let $x, y \in M$ be such that $\ell = d_M(y, z) \leq r_1$. Also, let $[yz] = \gamma_{y,\xi}([0, \ell])$, $\xi \in S_y M$ be a length minimizing geodesic from y to z. Let $q = \gamma_{y,\xi}(\ell/2)$. Using properties of Φ given in (1.18) (see also (1.5)-(1.6)), we see that for all $x \in B_M(q, r_1)$ we have $x \in B_M(y, r_0)$ and $x \in B_M(z, r_0)$, and so $\Phi_y(x)\Phi_z(x) \geq \phi_1^2$. Observe that

$$\mu(B_M(q, r_1)) \geq c_3 r_1^n. \tag{3.2}$$

Hence,

$$A(y, z) = \int_M \Phi_y(x)\Phi_y(x) d\mu(x) \geq \int_{B_M(q, r_1)} \Phi_y(x)\Phi_y(x) d\mu(x) \geq \phi_1^2 c_3 r_1^n.$$

□

Let X have distribution μ. Let $Y_{X,y}$ be a random variable, taking values in $\{0,1\}$, that is 1 with probability $\Phi(X,y)$ and $Y_{X,z}$ be a random variable, taking values in $\{0,1\}$, that is 1 with probability $\Phi(X,z)$. Also, let η,η' have the zero mean and variance σ^2, be such that all X, η, η' are independent random variables. We assume that under the condition that X is given, random variables $Y_{X,y}$, $Y_{X,z}$, η, and η' are independent.

Lemma 3.3. *Let $y,z\in M$. We have*

$$\mathbb{E}((|(d_M(y,X)+\eta)-(d_M(z,X)+\eta')|^2-2\sigma^2)Y_{X,y}Y_{X,z})=k_\Phi(y,z) \tag{3.3}$$

Proof. We denote $R_y(X)=d_M(y,X)+\eta$ and $R_z(X)=d_M(z,X)+\eta'$. Then

$$\begin{aligned}\mathbb{E}(|R_z(X)-R_y(X)|^2Y_{X,y}Y_{X,z}) &= \mathbb{E}_{\eta,\eta'}\int_M |(d_M(y,x)+\eta)-(d_M(z,x)+\eta')|^2\Phi_y(x)\Phi_z(x)d\mu(x)\\ &= \|(r_y-r_z)\Phi_y^{1/2}\Phi_z^{1/2}\|^2_{L^2(M,d\mu)}+2\sigma^2A(y,z).\end{aligned}$$

□

3.1. **Deterministic estimates for the rough distance function.** In this subsection, we consider the rough distance function $k_\Phi(y,z)$.

In the study of metric spaces, Kuratowski observed that the map $R(x)=r_x$ defines an isometric embedding $R:M\to R(M)\subset C(M)$ of the manifold M into the vector space $C(M)$. When there are no missing data, that is, $\Phi=1$, the following proposition show that the map $\overline{R}:M\to L^2(M)$, given by $\overline{R}(x)=r_x$ defines a bi-Lipschitz embedding $\overline{R}:M\to\overline{R}(M)\subset L^2(M)$. Note that here $L^2(M)$ is the space $L^2(M,d\mu)$, where μ is a probability measure on M.

Proposition 3.4. *There is a constant $c_5\in(0,1)$ such that*

$$c_5 d_M(y,z) \le \|r_y-r_z\|_{L^2(M,d\mu)}\le d_M(y,z). \tag{3.4}$$

Due to this, we call the map $\overline{R}:M\to L^2(M)$ the L^2-Kuratowski embedding. The proof of the Proposition 3.4 is the special case of the Proposition 3.5 when $\Phi=1$, claims (i)-(ii), given below.

Proposition 3.5. *(i) We have for all $y,z\in M$ the inequality*

$$\|(r_y-r_z)\Phi_y^{1/2}\Phi_z^{1/2}\|_{L^2(M,d\mu)}\le A(y,z)d_M(y,z)\le d_M(y,z). \tag{3.5}$$

(ii) Let $\widehat{c}_4=\frac14\min(c_2Hr_1,c_4)$. There is $c_5\le1$ such that if $A(y,z)\ge\widehat{c}_4$ then

$$\|(r_y-r_z)\Phi_y^{1/2}\Phi_z^{1/2}\|_{L^2(M,d\mu)}=k_\Phi(y,z)^{1/2}\ge c_5d_M(y,z). \tag{3.6}$$

(iii) For all $y,z\in M$ satisfying $d_M(y,z)\le r_1$, where r_1 is defined in (1.17), we have $A(y,z)\ge c_4\ge\widehat{c}_4$, and so the inequality (3.6) is valid.

By this proposition, if $A(y,z)\ge\widehat{c}_4$ then $k_\Phi(y,z)^{1/2}$ approximates $d_M(y,z)$.

Proof. (i) We have by triangular inequality

$$\|(r_y-r_z)\Phi_y^{\frac12}\Phi_z^{\frac12}\|^2_{L^2} = \int_M |d_M(y,\cdot)-d_M(z,\cdot)|^2\Phi_y\Phi_z d\mu\le|d_M(y,z)|^2A(y,z).$$

As $A(y,z)\le1$, this proves the inequality (3.5).

(ii) To prove the inequality in (3.6), we use the following (well known) corollary of Toponogov's theorem. Similar kind of formulas are used in Section 4.5 of [5]. However, we present the results in the form needed later and give the proof for the convenience of the reader.

Lemma 3.6. *Let M be a Riemannian manifold with sectional curvature bounded below by $-\Lambda^2$. Let $x,y,z\in M$ and $\beta=\angle xyz$ be the angle of the length minimizing curves $[xy]$ and $[yz]$ at y. Assume that $d_M(y,z)\le\frac12d_M(x,y)$ and $d_M(x,y)\le\frac23\min(i_0,\pi/(2\Lambda))$. Then*

$$\Big|d_M(x,z)-(d_M(x,y)-d_M(y,z)\cos\beta)\Big|\le\frac{d_M(y,z)^2}{d_M(x,y)}. \tag{3.7}$$

Proof. Let $\gamma_{y,\xi}([0,\ell])$ be a distance minimizing geodesic from y to z, where $|\xi|=1$ and $\ell=d_M(y,z)$. Consider functions

$$F(p)=d_M(x,p),\quad f(s)=F(\gamma_{y,\xi}(s)).$$

Let $\ell_0=\min(i_0,\pi/(2\Lambda))$. Observe that then $d_M(\gamma_{y,\xi}(s),x)\le \ell_0$ for all $s\in[0,\ell]$.

The gradient of $F(p)$ at $p\in B(0,\ell_0)$ is equal to the normal vector ν of the sphere $\Sigma=\partial B(x,r)$, where $r=d_M(p,x)$, at the point p and the Hessian of F at p and the shape operator $S(p)$ of the sphere Σ have the relation $\mathrm{Hess}(F)(\xi,\eta)=g(S(p)\xi,\eta)$, $\xi,\eta\in T_pM$, where $g:T_pM\times T_pM\to\mathbb{R}$ is the quadratic form determined by the metric tensor g, see [45]. By the standard comparision estimates [45, Ch. 6, Thm. 2.1], in the space $T_p\Sigma$ we have

$$\frac{\Lambda\cosh(\Lambda r)}{\sinh(\Lambda r)}\le S(p)\le \frac{\Lambda\cos(\Lambda r)}{\sin(\Lambda r)}. \tag{3.8}$$

As $\frac{d}{dt}(\tan(t))=1/\cos^2(t)$, the mean value theorem implies that for $0<s<\pi/(2\Lambda)$ we have $\tan(\Lambda s)\ge \Lambda s$, so that $\|S(p)\|\le 1/F(p)$.

Since $\partial_s f(s)=g(\nabla F(\gamma_{y,\xi}(s)),\dot\gamma_{y,\xi}(s))=g(\nu(\gamma_{y,\xi}(s)),\dot\gamma_{y,\xi}(s))$, where $\nu(x)=\nabla F(x)$ is the normal of the sphere $\partial B(y,s)$ at the point $x=\gamma_{y,\xi}(s)$. Moreover, $f(0)=d_M(x,y)$ and $\partial_s f(0)=g(-\dot\gamma_{y,\xi}(s),\dot\gamma_{y,\xi}(s))=-\cos\beta$. Also, since

$$\partial_s^2 f(s)=(\mathrm{Hess}\,F)(\dot\gamma(s),\dot\gamma(s))+g(\nabla F(\gamma(s)),\nabla_{\gamma(s)}\dot\gamma(s))=g(S(\gamma(s))\dot\gamma(s),\dot\gamma(s)),$$

where $\gamma=\gamma_{y,\xi}$, we have

$$|\partial_s^2 f(s)|\le \frac{1}{f(s)}\le \frac{1}{d_M(x,y)-d_M(y,z)}.$$

Hence, using Taylor's series we see that

$$\Big|f(s)-(d_M(x,y)-s\cos\beta)\Big|\le \frac{s^2}{2(d_M(x,y)-d_M(y,z))}\le \frac{s^2}{d_M(x,y)}.$$

This proves the claim. □

Next we continue the proof of inequality (3.6) We consider the claim in two cases:

Case 1. Assume that $d_M(y,z)\ge r_1/16$. We show that there $c_5'>0$ such that

$$\|r_y-r_z\|_{L^2(M,d\mu)}\ge \|(r_y-r_z)\Phi_y^{1/2}\Phi_z^{1/2}\|_{L^2(M,d\mu)}\ge c_5' d_M(y,z). \tag{3.9}$$

Proof in Case 1. Assume that $0<d_M(y,z)<r_1/16$. Then, if $x\in M$ is such that $r_1/2<d_M(x,z)<r_1$, we have $d_M(x,y)\ge r_1/4$ and by (3.7),

$$d_M(x,z)\quad\le\quad d_M(x,y)-d_M(y,z)\cos\beta+\frac14 d_M(y,z),$$

where β is the angle $\angle xyz$. When $\beta<\pi/4$, this yields

$$d_M(x,y)-d_M(x,z)\quad\ge\quad d_M(y,z)\cos\beta-\frac14 d_M(y,z)\ge \frac14 d_M(y,z).$$

Thus, let

$$W=\{x\in M;\ r_1/2<d_M(x,z)<r_1,\ \angle xyz<\pi/4\}.$$

Then

$$\|(r_y-r_z)\Phi_y^{1/2}\Phi_z^{1/2}\|^2_{L^2(M,d\mu)}\quad\ge\quad \frac{1}{16}d(y,z)^2\phi_1^2\,\mathrm{vol}_\mu(W),$$

where by [45, Cor. 2.4 in Chapter 6.2], see also (1.16), (1.17), there is $c_3'=c_3'(n,\Lambda)>0$ such that

$$\mathrm{vol}_\mu(W)\ge \rho_{min}\mathrm{vol}_M(W)\ge \rho_{min}4^{-n}\omega_n\Big(\frac{\sin(\Lambda r_1)}{\Lambda r_1}\Big)^{n-1}(r_1^n-(r_1/2)^n)\ge 2^{-4n}\rho_{min}c_3' r_1^n.$$

Thus there exists c_5' such that (3.9) is valid.

Case 2. Assume that $d_M(y,z)\ge r_1/16$. Then we show that there c_5'' such that

$$\|(r_y-r_z)\Phi_y^{1/2}\Phi_z^{1/2}\|_{L^2(M,d\mu)}=k_\Phi(y,z)^{1/2}\ge c_5'' d_M(y,z). \tag{3.10}$$

Proof in Case 2. The assumption $d_M(y,z) \geq r_1/16$ and definition of $\widehat{c}_4$ imply that $4\widehat{c}_4 c_2^{-1} H^{-1} \leq r_1$ so that $d_M(y,z) \geq \frac{1}{4}\widehat{c}_4 c_2^{-1} H^{-1}$. Denote $d_M(y,z) = \ell$. Then $\ell \geq 2a$, where $a = \frac{1}{8}\widehat{c}_4 c_2^{-1} H^{-1}$.

Assume that $A(y,z) \geq \widehat{c}_4$. Since $d_M(y,x) + d_M(x,z) \geq d_M(y,z) = \ell$, we see that for all $x \in M$ we have either $d_M(x,y) \geq \ell/2$ or $d_M(x,z) \geq \ell/2$. Let us assume that the latter is true. As $s \mapsto \Phi^1(s)$ is non-increasing and Φ takes values in $[0,1]$, we have $\Phi_y(x)\Phi_z(x) \leq \Phi_z(x) \leq c_2\Phi^1(d_M(x,z)) \leq c_2\Phi^1(\ell/2)$. This yields that

$$\widehat{c}_4 \leq A(y,z) = \int_M \Phi_y(x)\Phi_z(x)d\mu(x) \leq c_2\Phi^1(\ell/2),$$

so that $\Phi^1(\ell/2) \geq \widehat{c}_4/c_2$.

Let $[yz] = \gamma_{y,\xi}([0,\ell])$, $\xi \in S_yM$ be a length minimizing geodesic from y to z. Let $q = \gamma_{y,\xi}(\ell/2)$, $p = \gamma_{y,\xi}(\ell/2 - a)$, and $r = a/2$. When $d_M(x,p) < r$, we have

$$d_M(y,x) \leq d_M(y,p) + d_M(p,x) \leq \ell - \frac{a}{2}, \quad d_M(z,x) \geq d_M(z,p) - d_M(p,x) \geq \ell + \frac{a}{2},$$

so that $d_M(z,x) - d_M(y,x) \geq a$. Recall that $\Phi(x,y) \geq c_1\Phi^1(d_M(x,y))$, $\Phi^1 : [0,\infty) \to [0,1]$ and $\|\Phi^1\|_{C^1(\mathbb{R})} \leq H$. As $\Phi^1_y(q) = \Phi^1_z(q) = \Phi^1(\ell/2) \geq \widehat{c}_4/c_2$ and $B_M(p,r) \subset B_M(q, \frac{3}{2}a)$, we have for all $x \in B_M(p,r)$

$$\Phi_y(x)\Phi_z(x) \geq c_1^2(\Phi^1_y(p) - H(a+r))\,(\Phi^1_z(p) - H(a+r)) \geq c_1^2(c_2^{-1}\widehat{c}_4 - \frac{3}{2}Ha)^2 \geq \frac{1}{4}c_1^2 c_2^{-2}\widehat{c}_4^{\,2}.$$

Then, as $\ell/D \leq 1$,

$$\begin{aligned} \|(r_y - r_z)\Phi_y^{1/2}\Phi_z^{1/2}\|^2_{L^2(M,d\mu)} &\geq \int_{B_M(p,r)} |d_M(y,x) - d_M(z,x)|^2 \Phi_y(x)\Phi_z(x)\, d\mu(x) \\ &\geq (c_5'')^2\ell^2 = (c_5'')^2 d_M(y,z)^2, \end{aligned}$$

where $c_5'' = (16^{-n-3}\rho_{min} c_1^2 c_2^{-n+4}\widehat{c}_3 H^{-(n+2)}(\widehat{c}_4)^{n+4} D^{-2})^{1/2}$. This yields (ii) with $c_5 = \min(c_5', c_5'')$.

(iii) As $c_1 \leq 1$, Lemma 3.2 implies that if $d_M(y,z) \leq r_1$, then $A(y,z) \geq c_4 \geq \widehat{c}_4$. □

3.2. **Probabilistic estimates for the rough distance function.** In this subsection, we consider the rough distance function $k_\Phi(X_j, X_k)^{1/2}$.

Next we determine approximately $k_\Phi(X_j, X_k) = \|(r_{X_j} - r_{X_k})\Phi_{X_j}^{1/2}\Phi_{X_k}^{1/2}\|^2_{L^2(M,d\mu)}$ for $(j,k) \in \mathcal{I}^{(0)} \times \mathcal{I}^{(1)}$ using averaging over the data on the "densest net" $S_2 = \{X_j : \ j \in \mathcal{I}^{(2)}\}$.

For $(j,k) \in I^{(0)} \times I^{(1)}$, we consider the random variables

$$K_{jk} = \sum_{\ell \in I^{(2)}} \frac{1}{N_2}(|\overline{D}_{j,\ell} - \overline{D}_{k,\ell}|^2 - 2\sigma^2)Y_{j\ell}Y_{k\ell}, \tag{3.11}$$

$$K^L_{jk} = \sum_{\ell \in I^{(2)}} \frac{1}{N_2}(\min(|\overline{D}_{j,\ell} - \overline{D}_{k,\ell}|^2, L) - 2\sigma^2)Y_{j\ell}Y_{k\ell}. \tag{3.12}$$

We also consider the random variables

$$A_{jk} = A(X_j, X_k) = \int_M \Phi_{X_j}(x)\Phi_{X_k}(x)d\mu(x), \quad T_{jk} = \sum_{\ell \in I^{(2)}} Y_{j\ell}Y_{k\ell}. \tag{3.13}$$

Roughly speaking, below T_{jk} measures how well K^L_{jk} approximates $k_\Phi(X_j, X_k)$ that further approximates $d_M(X_j, X_k)^2$.

3.2.1. *Probabilistic notations.* To introduce some notations, let us assume for simplicity that the complete probability space $(\Omega, \Sigma, \mathbb{P})$ can be represented as a product $\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$ such that $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2 \times \mathbb{P}_3$ and $\omega = (\omega_1, \omega_2, \omega_3) \in \Omega_1 \times \Omega_2 \times \Omega_3$. We assume that $\eta = \eta(\omega_1)$, $\eta = \eta'(\omega_2)$ are random variables with variance σ, and $X = X(\omega_3)$. We assume that X is a random variables having distribution μ. Assume that X, η and η' are independent.

For an integrable function $F(\omega_1,\omega_2,\omega_3)=f(\eta(\omega_1),\eta'(\omega_2),X(\omega_3))$ we denote

$$\mathbb{E}_{\eta,\eta'}F=\mathbb{E}_{\eta,\eta'}f(\eta,\eta',X)=\int_{\Omega_1}\int_{\Omega_2}F(\omega_1,\omega_2,\omega_3)d\mathbb{P}_1(\omega_1)d\mathbb{P}_2(\omega_2). \tag{3.14}$$

As X is a random variable, we have that also $\mathbb{E}_{\eta,\eta'}f(\eta,\eta',X)$ is a random variable. The expectation $\mathbb{E}_{\eta,\eta'}f(\eta,\eta',X)$ over variables η,η' is function of X, and thus it can be considered as the expectation of $f(\eta,\eta',X)$ under the condition that X is known.

Below, we consider conditional expectations using σ-algebras. Let $\mathcal{B}_X\subset\Sigma$ be σ-algebra generated by the random variable $X:\Omega\to\mathbb{R}$, that is, the σ-algebra generated by sets $X^{-1}(S)\subset\Omega$, where $S\subset\mathbb{R}$ is an open set, see [31, Ch. 5]. We recall that $\mathbb{E}(F|\mathcal{B}_X)(\omega)$ is the $\mathcal{B}_X$-measurable random variable that satisfies

$$\int_S \mathbb{E}(F|\mathcal{B}_X)(\omega)\,d\mathbb{P}(\omega)=\int_S F(\omega)\,d\mathbb{P}(\omega)$$

for all sets $S\in\mathcal{B}_X$. In formula (3.14), $\mathbb{E}_{\eta,\eta'}F$ is in fact equal to the conditional expectation $\mathbb{E}(F|\mathcal{B}_X)=\mathbb{E}(F|\mathcal{B}_X)(\omega)$ of the random variable F with respect to the σ-algebra $\mathcal{B}_X\subset\Sigma$, that is, we have $(\mathbb{E}_{\eta,\eta'}F)(\omega_3)=\mathbb{E}(F|\mathcal{B}_X)(\omega)$ with $\omega=(\omega_1,\omega_2,\omega_3)$. As X is a random variable, $\mathbb{E}(Z\,|\mathcal{B}_X)$ is a random variable, too. Recall also the notation $\mathbb{P}(A\,|\mathcal{B}_X)=\mathbb{E}(1_A\,|\mathcal{B}_X)$ for an event $A\in\Sigma$, where $1_A(\omega)$ is the indicator function of the set $A\subset\Omega$. Below we use several times the fact that

$$\mathbb{E}(\mathbb{E}(Z\,|\mathcal{B}_X))=\mathbb{E}(Z),\quad \mathbb{E}(\mathbb{P}(A\,|\mathcal{B}_X))=\mathbb{P}(A). \tag{3.15}$$

Below, we will consider the σ-algebra $\mathcal{B}_j\subset\Sigma$ generated by the random variable $X_j:\Omega\to\mathbb{R}$. We also consider the σ-algebra $\mathcal{B}_{jk}$ generated by the random variables X_j and X_k.

By Lemma 3.3, the conditional expectation of K_{jk}, under the condition that X_j and X_k are known, satisfies $\mathbb{E}\Big(K_{jk}\,\Big|\mathcal{B}_{jk}\Big)=k_\Phi(X_j,X_k)$ where $k_\Phi(X_j,X_k)=\|(r_{X_j}-r_{X_k})\Phi_{X_j}^{1/2}\Phi_{X_k}^{1/2}\|^2_{L^2(M,d\mu)}$ is a random variable.

3.2.2. *Probabilistic estimates for rough distances K^L_{jk} and reliability values T_{jk}.* Below use the following form of Hoeffding's inequality.

Lemma 3.7 (Hoeffding's inequality [27]). *Let $Z_1,\dots,Z_N$ be N i.i.d. copies of the random variable Z satisfying $0\le Z\le L$, where $L>0$. Then, for $\varepsilon>0$, we have*

$$\mathbb{P}\left[\left|\frac{1}{N}\left(\sum_{i=1}^N Z_i\right)-\mathbb{E}[Z]\right|\le\varepsilon\right]\ge 1-2\exp(-2N\varepsilon^2L^{-2}).$$

Below, we will show that K^L_{jk}, defined in (3.12), can be considered to be an approximation of $k_\Phi(X_j,X_k)$ which further approximates $d_M(X_j,X_k)^2$ when $A(X_j,X_k)$ is larger than a suitable threshold value. Let

$$\varepsilon_3<\frac{1}{4}c_4,\quad b=\frac{1}{2}c_4. \tag{3.16}$$

For $j\in\mathcal{I}^{(0)}$, we consider the events $\mathcal{E}^{(1)}_j\subset\Omega$ and $\mathcal{E}^{(1)}\subset\Omega$, defined by

$$\mathcal{E}^{(1)}_j=\Big\{\omega\in\Omega\Big|\ \forall k\in I^{(1)}\ (\Big|\frac{T_{jk}}{N_2}-A_{jk}\Big|\le\varepsilon_3)\Big\},\quad \mathcal{E}^{(1)}=\bigcap_{j\in I^{(0)}}\mathcal{E}^{(1)}_j. \tag{3.17}$$

Below, we use a smooth cut-off functions

$$\psi_\rho(t)=\psi_1(t/\rho^2),\quad \beta_1(t)=1-\psi_1(t).$$

where $\psi_1\in C_0^\infty(\mathbb{R})$ satisfies supp $(\psi_1)\subset(-2,2)$ and $\psi_1(t)=1$ for $-1\le t\le 1$ and $0\le\psi_1(t)\le 1$ for all $t\in\mathbb{R}$, $\psi_1(-t)=\psi_1(t)$, $\|\psi_1\|_{C^1(\mathbb{R})}\le 2$, and the function $\psi_1|_{\mathbb{R}_+}$ is non-increasing.

Note that if $\beta_1(T_{jk}(bN_2)^{-1})>0$ then $\frac{T_{jk}}{bN_2}\ge 1$. Also, if $\mathcal{E}^{(1)}_j$ happens and $A_{jk}\ge c_4$ then

$$\frac{T_{jk}}{N_2}\ge A_{jk}-\varepsilon_3\ge c_4-\frac{1}{4}c_4\ge b. \tag{3.18}$$

Moreover, if $\frac{1}{b}\frac{T_{jk}}{N_2} \geq 1$, then (3.18) implies

$$A_{jk} \geq b - \varepsilon_3 \geq \frac{1}{4}c_4 \geq \widehat{c}_4. \tag{3.19}$$

For $y, z \in M$, let $Y(z, y)$ be a random variable that is 1 with probability $\Phi(y, z)$ and 0 with probability $1 - \Phi(y, z)$. Assume that for $y, z \in M$ the random variables $Y(y, z)$ are independent. Let

$$T(y, z) \quad = \sum_{\ell \in I^{(2)}} Y(y, X_\ell)Y(z, X_\ell). \tag{3.20}$$

As $\mathbb{E}\frac{1}{N_2}T(y, z) = A(y, z)$, Hoeffding's inequality implies

$$\mathbb{P}\left[\left|\frac{T(y, z)}{N_2} - A(y, z)\right| \leq \varepsilon_3\right] \geq 1 - 2\exp(-2N_2\varepsilon_3^2). \tag{3.21}$$

As T_{jk} and $T(X_j, X_k)$ have the same distributions and $A_{jk} = A(X_j, X_k)$ for $j \in \mathcal{I}^{(0)}$ and $k \in \mathcal{I}^{(1)}$, inequality (3.21) implies for the conditional probability, under the condition that X_j and X_k are known, that

$$\mathbb{P}\Big[\left|\frac{T_{jk}}{N_2} - A_{jk}\right| \leq \varepsilon_3 \,\Big|\, \mathcal{B}_{jk}\Big] \geq 1 - 2\exp(-2N_2\varepsilon_3^2). \tag{3.22}$$

Thus we have by (3.15) $\mathbb{P}(\mathcal{E}_j^{(1)}) \geq 1 - 2N_1\exp(-2N_2\varepsilon_3^2)$. Hence,

$$\mathbb{P}(\mathcal{E}^{(1)}) \geq 1 - p^{(1)}, \quad p^{(1)} = 2N_0N_1\exp(-2N_2\varepsilon_3^2). \tag{3.23}$$

We recall that by Lemma 3.2 and Proposition 3.5,

$$\Big(d_M(y, z) \leq r_1 \implies A(y, z) \geq c_4 \geq \widehat{c}_4\Big), \quad \text{and} \tag{3.24}$$

$$\Big(A(y, z) \geq \widehat{c}_4 \implies d_M(y, z) \geq \|(r_y - r_z)\Phi_y^{1/2}\Phi_z^{1/2}\|_{L^2(M,d\mu)} = k_\Phi(y, z)^{\frac{1}{2}} \geq c_5 d_M(y, z)\Big).$$

Lemma 3.8. *Let $L > 2\max(D^2, \sigma)$, $\varepsilon_2 > 0$, and $\varepsilon(L) := \beta e^{-(L^{1/2}-D)/2}(D^2 + 6\beta^2)$ and consider the events $\mathcal{E}_j^{(2)} \subset \Omega$, $j \in I^{(0)}$, and $\mathcal{E}^{(2)} \subset \Omega$,*

$$\mathcal{E}_j^{(2)} = \{\forall k \in I^{(1)} : \ |K_{jk}^L - k_\Phi(X_j, X_k)| \leq \varepsilon_2 + \varepsilon(L)\}, \quad \mathcal{E}^{(2)} = \bigcap_{j \in I^{(0)}} \mathcal{E}_j^{(2)}.$$

Then

$$\mathbb{P}(\mathcal{E}^{(2)}) \geq 1 - p^{(2)}, \quad p^{(2)} = 2N_0N_1\exp(-2N_2\varepsilon_2^2L^{-2}). \tag{3.25}$$

Proof. Denote $\eta_{jk\ell} = \eta_{j\ell} - \eta_{k\ell}$ and $D_{jk\ell} = d_M(X_j, X_\ell) - d_M(X_k, X_\ell)$. Then

$$\mathbb{E}\eta_{jk\ell} = 0, \quad \mathbb{E}\eta_{jk\ell}^2 = 2\sigma^2, \quad \mathbb{E}e^{|\eta_{jk\ell}|} \leq (\mathbb{E}e^{|\eta_{j\ell}|})(\mathbb{E}e^{|\eta_{k\ell}|}) \leq \beta^2. \tag{3.26}$$

Let $r = L^{1/2}$. Observe that $|D_{jk\ell}| \leq D$, so that if $|D_{jk\ell} + \eta_{jk\ell}| \geq r$ then $|\eta_{jk\ell}| > r - D$. By (3.26), $Y = e^{|\eta_{jk\ell}|}$ satisfies

$$\mathbb{P}(|\eta_{jk\ell}| > r - D) \leq \frac{\mathbb{E}(e^{|\eta_{jk\ell}|})}{e^{r-D}} = \beta^2 e^{-(r-D)}.$$

Thus, using the fact that $\eta_{jk\ell}$ and $D_{jk\ell}$ are independent, we see using Schwartz inequality that

$$\begin{aligned}\mathbb{E}\Big(|(\min((D_{jk\ell} + \eta_{jk\ell})^2, L) - (D_{jk\ell} + \eta_{jk\ell})^2)Y_{j\ell}Y_{k\ell}| \,\Big|\, \mathcal{B}_{jk}\Big) \quad &\leq \quad \mathbb{E}(\chi_{|\eta_{jk\ell}|>r-D}(D_{jk\ell} + \eta_{jk\ell})^2|\, \mathcal{B}_{jk}) \\ &\leq (\mathbb{P}(|\eta_{jk\ell}| > r - D))^{\frac{1}{2}} \,(\mathbb{E}((D_{jk\ell} + \eta_{jk\ell})^4|\, \mathcal{B}_{jk}))^{\frac{1}{2}} \\ &\leq \beta e^{-(r-D)/2}(D^2 + 6\beta^2) \leq \varepsilon(L).\end{aligned}$$

By Lemma 3.3, the above shows that

$$k_\Phi^L(X_j, X_k) := \mathbb{E}((\min((D_{jk\ell} + \eta_{jk\ell})^2, L) - 2\sigma^2)Y_{j\ell}Y_{k\ell}|\, \mathcal{B}_{jk})$$

satisfies

$$\left| k_\Phi^L(X_j, X_k) - k_\Phi(X_j, X_k) \right| \le \varepsilon(L). \tag{3.27}$$

By arguing as in (3.21)-(3.22), we see that Hoeffding's inequality implies

$$\mathbb{P}\left[|K_{jk}^L - k_\Phi^L(X_j, X_k)| \le \varepsilon_2 \middle| \mathcal{B}_{jk} \right] \ge 1 - 2\exp(-2N_2 \varepsilon_2^2 L^{-2}). \tag{3.28}$$

Using this, (3.15) and (3.27), and summing over $j \in I^{(0)}$, we obtain (3.25). □

4. Determination of the approximate distances in the coarse net

Next we assume that

$$\rho \le r_1 = r_0/2 \le \phi_0/(4H), \quad u_0 = \phi_1 c_3 (\rho/4)^n, \quad u_1 = u_0/2, \quad u_2 = u_0/4. \tag{4.1}$$

Next we define $Q_{j,j'}$ that will turn out to the approximate distances $d_M(X_j, X_{j'})$ for points X_j and $X_{j'}$, where $j, j' \in I^{(0)}$, that are sufficiently close to each other.

Definition 4.1. Let $\rho \in (0,1)$ satisfy (4.1). For $j, j' \in I^{(0)}$, let

$$Q_{j,j'} = \frac{V_{j,j'}}{W_{j,j'}}, \tag{4.2}$$

$$V_{j,j'} = \frac{1}{N_1} \sum_{k \in I^{(1)}} \beta_1(T_{jk}(bN_2)^{-1}) \psi_\rho(K_{jk}^L) Y_{j'k} \overline{D}_{k,j'}, \tag{4.3}$$

$$W_{j,j'} = \frac{1}{N_1} \sum_{k \in I^{(1)}} \beta_1(T_{jk}(bN_2)^{-1}) \psi_\rho(K_{jk}^L) Y_{j'k}. \tag{4.4}$$

In the case when $W_{j,j'}$ is zero, we define $Q_{j,j'}$ to be D. We define for $j, j' \in I^{(0)}$

$$d^{app}(X_j, X_{j'}) = \begin{cases} Q_{j,j'}, & \text{if } W_{j,j'} > u_2, \\ D, & \text{otherwise.} \end{cases} \tag{4.5}$$

Roughly speaking, above the function $\beta_1(T_{jk}(bN_2)^{-1})$ measures the reliability of the terms $\psi_\rho(K_{jk}^L)$ in formulas (4.3) and (4.4), and $W_{j,j'}$ measures reliability of $Q_{j,j'}$ in formula (4.5). The numbers $d^{app}(X_j, X_{j'})$ will be the final approximation for the distances $d_M(X_j, X_{j'})$ for all pairs $(X_j, X_{j'})$ of points that are close to each other. Observe that $Q_{j,j'}$ and $W_{j,j'}$ can be computed from the given data.

For technical purposes, we define deterministic (indexed with (d)) and random (indexed with (r)) functions

$$W^{(d),-}(y,z) = \int_M \beta_1(A(y,x)b^{-1}) \Phi(z,x)\, \psi_{\rho/2}(k_\Phi(y,x)) d\mu(x), \tag{4.6}$$

$$W^{(r),-}(y,z) = \frac{1}{N_1} \sum_{k \in I^{(1)}} \beta_1(A(y, X_k)b^{-1}) \Phi(z, X_k)\, \psi_{\rho/2}(k_\Phi(y, X_k)). \tag{4.7}$$

The motivation behind defining functions $W^{(d),-}$ and $W^{(r),-}$ is that we can use Hoeffding's inequality to estimate how close $W^{(d),-}(X_j, X_{j'})$ and $W^{(r),-}(X_j, X_{j'})$ are when X_j and $X_{j'}$ are known. Also, we show that we have $W_{j,j'} \ge W^{(r),-}(X_j, X_{j'})$ with a large probability.

Lemma 4.2. *If $d_M(x,z) < r_1$, then*

$$W^{(d),-}(y,z) \ge u_0. \tag{4.8}$$

Moreover, when we have $\Phi(x,y) \ge c_1 \phi_0$ for all $x, y \in M$, the inequality (4.8) holds for all $x, y \in M$.

Proof. Let

$$w^-(x,y,z) = \beta_1(A(y,x)b^{-1})\Phi(z,x)\,\psi_{\rho/2}(k_\Phi(y,x))$$

and recall that by (4.1), $\rho \le r_1$.

When $d_M(x,y) < \rho/4$, by Lemma 3.5 (iii), we have that $A(x,y) \ge c_4$. Also, in the case when $\Phi(x,y) \ge c_1\phi_0 \ge \phi_1$ for all $(x,y) \in M\times M$, we have $A(x,y) \ge c_4$. Thus in both cases, $\beta_1(A(y,x)b^{-1}) = 1$. Moreover, by 3.5 (i), we have $k_\Phi(y,x) \le (d_M(y,x))^2 < (\rho/2)^2$, so that $\psi_{\rho/2}(k_\Phi(y,x)) = 1$.

Also, by (1.18), when $d_M(x,z) < r_1 = r_0/2$, we have $\Phi(z,x) \ge \phi_1$.

Thus, when $d_M(x,z) < r_1$ or $\Phi(x,y) \ge c_1\phi_0 \ge \phi_1$, and we have $d_M(x,y) < \rho/4$, it holds that $w_-(x,y,z) \ge \phi_1$. Hence we see that

$$W^{(d),-}(y,z) = \int_M w_-(x,y,z)d\mu(x) \ge \phi_1\cdot \mu(B_M(y,\frac{\rho}{4})) \ge \phi_1\cdot c_3(\rho/4)^n = u_0.$$

□

Let us write for $j,j' \in I^{(0)}$

$$Q_{j,j'} = Q^1_{j,j'} + Q^2_{j,j'},\ \text{where}\ Q^1_{j,j'} = \frac{V^1_{j,j'}}{W_{j,j'}},\quad Q^2_{j,j'} = \frac{V^2_{j,j'}}{W_{j,j'}}, \tag{4.9}$$
$$V^1_{j,j'} = \frac{1}{N_1}\sum_{k\in I^{(1)}} \beta_1(T_{jk}(bN_2)^{-1})\psi_\rho(K^L_{jk})Y_{j'k}d_M(X_k,X_{j'}),$$
$$V^2_{j,j'} = \frac{1}{N_1}\sum_{k\in I^{(1)}} \beta_1(T_{jk}(bN_2)^{-1})\psi_\rho(K^L_{jk})Y_{j'k}\eta_{k,j'},$$

and consider the terms $Q^1_{j,j'}$ and $Q^2_{j,j'}$ separately.

First we will show that $Q^2_{j,j'}$ is small with a large probability when $W_{j,j'} \ge u_2$. To that end, let $h_0 > 0$ and $\mathcal{E}^{(3)}_{j,j'} \subset \Omega$, $(j,j') \in I^{(0)}\times I^{(0)}$, and $\mathcal{E}^{(3)} \subset \Omega$ be the events

$$\mathcal{E}^{(3)}_{j,j'} = \{(W_{j,j'} \ge u_2) \implies (|Q^2_{j,j'}| \le h_0)\},\quad \mathcal{E}^{(3)} = \bigcap_{(j,j')\in I^{(0)}\times I^{(0)}} \mathcal{E}^{(3)}_{j,j'}. \tag{4.10}$$

Lemma 4.3. *For any $h_0 \in (0,1)$ we have*

$$\mathbb{P}(\mathcal{E}^{(3)}) \le 1 - p^{(3)},\quad p^{(3)} = 2N_0^2\exp(-e^{-2\beta}N_1u_2h_0^2/4). \tag{4.11}$$

Proof. Let us next recall some basic facts: Let $a = (a_k)_{k=1}^{N_1}$ satisfy $0 \le a_k \le 1$ and

$$S_{N_1} = (\sum_{k=1}^{N_1} a_k)^2/(\sum_{k=1}^{N_1} a_k^2),\quad Z_{N_1} = \sum_{k=1}^{N_1} a_k,\quad V_{N_1} = \frac{1}{Z_{N_1}}\sum_{k=1}^{N_1} a_k\eta_k, \tag{4.12}$$

where η_k are i.i.d. variables, $\mathbb{E}\eta_k = 0$ and $\mathbb{E}e^{|\eta_k|} \le \beta$. Since $0 \le a_k \le 1$, we have $a_k^2 \le a_k$, so that $\sum_{k=1}^{N_1} a_k^2 \le \sum_{k=1}^{N_1} a_k$, and

$$S_{N_1} = (\sum_{k=1}^{N_1} a_k)^2/(\sum_{k=1}^{N_1} a_k^2) \ge (\sum_{k=1}^{N_1} a_k)^2/(\sum_{k=1}^{N_1} a_k) \ge \sum_{k=1}^{N_1} a_k = Z_{N_1}. \tag{4.13}$$

Then using Jensen's inequality for the random variable $R = e^{\eta_k}$ with concave function $h : [0,\infty) \to \mathbb{R}$, $h(s) = s^t$ with $t \in [0,1]$ (for which the standard Jensen's inequality reverses), we obtain $\mathbb{E}(R^t) \le (\mathbb{E}R)^t$. This yields

$$\mathbb{E}(\exp(t\eta_k)) \le \beta^t \le e^{\beta t}$$

for $t \in [0,1]$. Hence, as $\mathbb{E}\eta_k = 0$, the moment generating function of η_k, $M(t) = \mathbb{E}(\exp(t\eta_k))$ satisfies, by considerations involving Taylor series and the mean value theorem, $|M_k(t) - 1| \le c't^2,\quad t \in [0,1]$, where $c' \le e^{2\beta}$. Then, by using the independency of random variables η_k, we see that the moment generating function of V_{N_1} satisfies for $s \in [0, Z_{N_1}]$

$$\mathbb{E}\exp(sV_{N_1}) = \prod_{k=1}^{N_1} M(s\frac{a_k}{Z_{N_1}}) \le \exp(c'\sum_{k=1}^{N_1}(s\frac{a_k}{Z_{N_1}})^2) \le \exp(e^{2\beta}\frac{s^2}{S_{N_1}}) \le \exp(e^{2\beta}\frac{s^2}{Z_{N_1}}),$$

where we use the fact that $\sum_{k=1}^{N_1}(a_k/Z_{N_1})^2 = 1/S_{N_1}$. Similarly, by considering $-\eta_k$ instead of η_k, we see that $\mathbb{E}\exp(-sV_{N_1}) \le \exp(e^{2\beta}\frac{s^2}{S_{N_1}}) \le \exp(e^{2\beta}\frac{s^2}{Z_{N_1}})$.

Then we see that for $0 < \lambda \le 1$ and $s \in [0, Z_{N_1}]$,

$$\mathbb{P}(|V_{N_1}| > \lambda) \quad \le \quad \mathbb{P}(sV_{N_1} > s\lambda) + \mathbb{P}(sV_{N_1} < -s\lambda) \le 2\exp(e^{2\beta}s^2 Z_{N_1}^{-1} - s\lambda).$$

When $s = \lambda Z_{N_1} e^{-2\beta}/2$, the above implies

$$\mathbb{P}(|V_{N_1}| > \lambda) \le 2\exp(e^{2\beta}\cdot 2^{-2}\lambda^2 Z_{N_1}^2 e^{-4\beta}\cdot Z_{N_1}^{-1} - 2^{-1}\lambda^2 Z_{N_1} e^{-2\beta}) = 2\exp(-\lambda^2 Z_{N_1} e^{-2\beta}/4). \tag{4.14}$$

Next we consider a fixed $j, j' \in I^{(0)}$ and let

$$a_k = \beta_1(T_{jk}(bN_2)^{-1})\psi_\rho(K_{jk}^L)Y_{j'k},$$

for $k \in I^{(1)}$, $\lambda = h_0$, and let $Z_{N_1} = N_1 W_{j,j'}$ be defined analogously to (4.12). Then we see that $Q^2_{j,j'} = V_{N_1}$, where V_{N_1} is be defined analogously to (4.12) with $\eta_k = \eta_{k,j'}$. Note that $\eta_{k,j'}$ are independent of variables T_{jk}, K^L_{jk}, and $Y_{j',k}$. Also, V_{N_1} is a function of these variables. Let $\mathcal{B}^*$ be the $\sigma-$algebra generated by all random variables T_{jk}, K^L_{jk}, and $Y_{j',k}$, $j, j' \in I^{(0)}$, $k \in I^{(2)}$. As Z_{N_1} is measurable with respect to the $\sigma-$algebra $\mathcal{B}^*$, by applying (3.15) and (4.14), we see

$$\mathbb{P}\Big((W_{j,j'} \ge u_2) \implies (|Q^2_{j,j'}| \le h_0)\Big) \quad \ge \quad 1 - 2\exp(-e^{-2\beta}N_1 u_2 h_0^2/4).$$

By doing this analysis for all $j, j' \in I^{(0)}$ and summing up the results, we obtain the claim. □

Next we analyse $Q^1_{j,j'}$. We assume that L is so large and $\varepsilon_2, \varepsilon_3$ are so small that

$$\varepsilon_2 + \varepsilon(L) \le \frac{1}{100}\rho^2, \quad \varepsilon_3 \le \frac{b}{4}u_1. \tag{4.15}$$

We denote, see (4.6),

$$W^{(d),-}_{j,j'} = W^{(d),-}(X_j, X_{j'}), \quad W^{(r),-}_{j,j'} = W^{(r),-}(X_j, X_{j'}), \quad W_{j,j'} = W(X_j, X_{j'}).$$

Let us consider the events $\mathcal{E}^{(4)}_{j,j'} \subset \Omega$ and $\mathcal{E}^{(4)} \subset \Omega$,

$$\mathcal{E}^{(4)}_{j,j'} = \Big\{\omega \in \Omega :\ (W^{(d),-}_{j,j'} \ge u_0) \implies (W^{(r),-}_{j,j'} \ge \frac{1}{2}u_0)\Big\}, \quad \mathcal{E}^{(4)} = \bigcap_{(j,j')\in I^{(0)}\times I^{(0)}} \mathcal{E}^{(4)}_{j,j'}.$$

Observe that

$$\Big\{\omega \in \Omega :\ |W^{(d),-}(X_j, X_{j'}) - W^{(r),-}(X_j, X_{j'})| \le \frac{1}{2}u_0\Big\} \subset \mathcal{E}^{(4)}_{j,j'}. \tag{4.16}$$

We see using Hoeffding's inequality that for $y, z \in M$ that

$$\mathbb{P}\Big(|W^{(d),-}(y,z) - W^{(r),-}(y,z)| \le \frac{1}{2}u_0\Big) \ge 1 - 2\exp(-2N_1(u_1/2)^2).$$

Thus we see using (4.16) for all $j, j' \in I^{(0)}$ and (3.15), and summing the results,

$$\mathbb{P}(\mathcal{E}^{(4)}) \ge 1 - p^{(4)}, \quad p^{(4)} = 2N_0^2\exp(-2N_1(u_1/2)^2).$$

Below, to shorten notations, let us denote $\mathcal{E}^{(5)} = \bigcap_{k\in\{1,2,3,4\}} \mathcal{E}^{(k)}$.

Lemma 4.4. *Assume the event $\mathcal{E}^{(5)}$ happens. Then, if (4.15) is valid, we have for all $(j, j') \in I^{(0)} \times I^{(0)}$ the implication $(W^{(d),-}_{j,j'} \ge u_0) \implies (W_{j,j'} \ge u_2)$.*

Proof. As $\mathcal{E}^{(5)}$ happens, also the event $\mathcal{E}^{(2)}$ happens. If $0 \le s \le \frac{1}{20}\rho^2$, then $\psi_\rho(s - \frac{1}{10^2}\rho^2) = 1$ and $\psi_{\rho/2}(s) = 1$. On the other hand, if $s > \frac{1}{20}\rho^2$, we have $4(s - \frac{1}{10^2}\rho^2) > s$. Thus, as the function $\psi_1|_{\mathbb{R}_+}$ is non-increasing,

$$\psi_\rho(s - \frac{1}{10^2}\rho^2) \ge \psi_{\rho/2}(4(s - \frac{1}{10^2}\rho^2)) \ge \psi_{\rho/2}(s)$$

for all $s \ge 0$. Hence, we have

$$\psi_\rho(s - (\varepsilon_2 + \varepsilon(L))) \quad \ge \quad \psi_\rho(s - \frac{1}{10^2}\rho^2) \ge \psi_{\rho/2}(s).$$

Thus, as $\mathcal{E}^{(2)}$ happens,

$$\psi_\rho(K^L_{jk}) \quad \ge \quad \psi_\rho(k_\Phi(X_j, X_k) - (\varepsilon_2 + \varepsilon(L))) \ge \psi_{\rho/2}(k_\Phi(X_j, X_k)). \tag{4.17}$$

When $\mathcal{E}^{(5)}$ happens, also $\mathcal{E}^{(4)}$ happens, and we have the implication $(W^{(d),-}_{j,j'} \ge u_0) \implies (W^{(r),-}_{j,j'} \ge u_1)$.

Recall that $\|\beta_1\|_{C^1(\mathbb{R})} \le 2$. Thus, when $\mathcal{E}^{(1)}$ happens,

$$|\beta_1(T_{jk}(bN_2)^{-1}) - \beta_1(A(X_j, X_k)b^{-1})| \le 2b^{-1}\varepsilon_3. \tag{4.18}$$

Then

$$\begin{aligned} & |\beta_1(T_{jk}(bN_2)^{-1})\psi_{\rho/2}(k_\Phi(X_j, X_k)) - \beta_1(A(X_j, X_k)b^{-1})\psi_{\rho/2}(k_\Phi(X_j, X_k))| \\ \le \; & 2b^{-1}\varepsilon_3 \cdot \psi_{\rho/2}(k_\Phi(X_j, X_k)) \le 2b^{-1}\varepsilon_3. \end{aligned}$$

This and (4.17) imply

$$\begin{aligned} \beta_1(T_{jk}(bN_2)^{-1})\psi_\rho(K^L_{jk}) \quad & \ge \quad \beta_1(T_{jk}(bN_2)^{-1})\psi_{\rho/2}(k_\Phi(X_j, X_k)) \\ & \ge \quad \beta_1(A(X_j, X_k)b^{-1})\psi_{\rho/2}(k_\Phi(X_j, X_k)) - 2b^{-1}\varepsilon_3. \end{aligned} \tag{4.19}$$

Recall that by the assumption of the claim, $\varepsilon_3 < \frac{b}{4}u_1$. Computing the average of the inequalities (4.19) times $Y_{j',k}$ over $k \in I^{(1)}$, we obtain

$$W_{j,j'} \ge W^{(r),-}_{j,j'} - 2b^{-1}\varepsilon_3 \ge W^{(r),-}_{j,j'} - \frac{1}{2}u_1.$$

Thus when $\mathcal{E}^{(5)}$ and thus $\mathcal{E}^{(4)}$ happen, we have

$$(W^{(d),-}_{j,j'} \ge u_0) \implies (W^{(r),-}_{j,j'} \ge u_1) \implies (W_{j,j'} \ge u_1 - \frac{1}{2}u_1 = u_2).$$

□

Lemma 4.5. *When the event $\mathcal{E}^{(5)}$ happens and (4.15) is valid, it holds for all $j, j' \in I^{(0)}$ that if $W(X_j, X_{j'}) \ge u_2$, then*

$$|Q_{j,j'} - d_M(X_j, X_{j'})| \le \frac{2}{c_5}\rho + h_0. \tag{4.20}$$

Proof. Below, in this proof assume that the event $\mathcal{E}^{(5)}$ happens. We will first show using Lemma 3.8 that if $W(X_j, X_{j'}) \ge u_2$, then

$$|Q^1_{j,j'} - d_M(X_j, X_{j'})| \le \frac{2}{c_5}\rho. \tag{4.21}$$

Consider next the indexes $(j,k) \in I^{(0)} \times I^{(1)}$ for which $\psi_\rho(K^L_{jk}) > 0$ and $\beta_1(T_{jk}(bN_2)^{-1})) > 0$. Then $\psi_\rho(K^L_{jk}) > 0$ implies $K^L_{jk} < 2\rho^2$, and hence, as the event $\mathcal{E}^{(2)}$ happens,

$$k_\Phi(X_j, X_k) < 2\rho^2 + \varepsilon_2 + \varepsilon(L). \tag{4.22}$$

Moreover, if $\beta_1(T_{jk}(bN_2)^{-1})) > 0$ then $T_{jk}N_2^{-1} > b$. As the event $\mathcal{E}^{(5)}$ happens,we have

$$A(X_j, X_k) = A_{jk} \ge T_{jk}N_2^{-1} - \varepsilon_3 \ge b - \varepsilon_3 \ge \frac{1}{4}c_4 \ge \widehat{c}_4,$$

see Proposition 3.5 and (3.16). Then by Proposition 3.5,

$$k_\Phi(X_j,X_k)^{1/2} \geq c_5 d_M(X_j,X_k), \tag{4.23}$$

that implies with (4.22) that

$$d_M(X_j,X_k) \leq \frac{1}{c_5} k_\Phi(X_j,X_k)^{1/2} \leq \frac{1}{c_5}\sqrt{2\rho^2+\varepsilon_2+\varepsilon(L)}. \tag{4.24}$$

We denote $v_{j,j',k} = \beta_1(T_{jk}(bN_2)^{-1})\psi_\rho(K^L_{jk})Y_{j'k}$ and observe that if $v_{j,j',k} > 0$ then the distance $d_M(X_j,X_k)$ satisfies (4.24).

Assume next that $W_{j,j'} \geq u_2$. Then, we have $\sum_{k\in I^{(1)}} v_{j,j',k} = W(X_j,X_{j'}) \geq u_2 > 0$ and

$$Q^1(X_j,X_{j'}) \quad = \quad V^1(X_j,X_{j'})/W(X_j,X_{j'}) = \frac{\sum_{k\in I^{(1)}} d_M(X_k,X_{j'})\, v_{j,j',k}}{\sum_{k\in I^{(1)}} v_{j,j',k}}.$$

This means that we can consider $Q^1(X_j,X_{j'})$ as a weighted average of distances $d_M(X_k,X_{j'})$. Hence, when the event $\mathcal{E}^{(5)}$ happens, we see that for all $(j,j') \in I^{(0)}\times I^{(0)}$ such that $W_{j,j'} \geq u_2$ we have

$$\begin{aligned} |Q^1_{j,j'} - d_M(X_j,X_{j'})| \quad &= \quad \Big|\frac{\sum_{k\in I^{(1)}} (d_M(X_k,X_{j'}) - d_M(X_j,X_{j'}))\, v_{j,j',k}}{\sum_{k\in I^{(1)}} v_{j,j',k}}\Big| \\ &\leq \quad \frac{\sum_{k\in I^{(1)}} d_M(X_k,X_j)\, v_{j,j',k}}{\sum_{k\in I^{(1)}} v_{j,j',k}} \leq \frac{1}{c_5}\sqrt{2\rho^2+\varepsilon_2+\varepsilon(L)}, \end{aligned}$$

where we have used (4.24) in the last inequality. By (4.15), here $\varepsilon_2+\varepsilon(L) < \frac{1}{100}\rho^2$ and the inequality (4.21) follows.

As we have assumed that the event $\mathcal{E}^{(5)}$ happens, also the event $\mathcal{E}^{(3)}$ happens, see (4.10). Thus if $W_{j,j'} \geq u_2$, then $Q^2_{j,j'} < h_0$. Combining the above, we see that

$$|Q_{j,j'} - d_M(X_j,X_{j'})| \leq |Q^1_{j,j'} - d_M(X_j,X_{j'})| + |Q^2_{j,j'}| \leq \frac{2}{c_5}\rho + h_0.$$

□

The above proposition means that if $W_{j,j'} \geq u_2$, then the number $Q_{j,j'}$ approximates $d_M(X_j,X_{j'})$ with a large probability when all parameters are suitably chosen.

Next we consider the proof of Theorem 2. We use below

$$\rho = \frac{2\varepsilon_1}{c_5},\ h_0 = \frac{\varepsilon_1}{2},\ \varepsilon_2 = \frac{\rho^2}{200},\ \varepsilon_3 = \frac{b}{4}u_1,\ L = 4\log^2\Big(e^{D/2}\frac{200\beta(D^2+6\beta^2)}{\rho^2}\Big). \tag{4.25}$$

Note that $\rho \leq 1$. Then $\varepsilon(L) \leq \rho^2/200$, implying that $\varepsilon_2+\varepsilon(L) \leq \rho^2/100$. Below we assume that $\varepsilon_1 < \widehat{\varepsilon}_1$, where $\widehat{\varepsilon}_1 = \min(1, 8c_5(\phi_1 c_3)^{-1/n})$. Then we have $\varepsilon_3 < \frac{1}{4}c_4$, see (3.16).

Below, see Lemma 2.1, let N_0 be

$$N_0 = \lfloor 2C_3\delta_1^{-n}(\log(\delta_1^{-1}) + \log(\theta^{-1}))\rfloor. \tag{4.26}$$

Next we consider the probability of $\mathcal{E}^{(5)}$. We see that $\mathbb{P}(\mathcal{E}^{(5)}) \geq 1-p^{(5)}$, where $p^{(5)} = p^{(1)}+p^{(2)}+p^{(3)}+p^{(4)}$. Next we consider there probabilities one by one.

Lemma 4.6. *There is $C_7 > 1$ such that we have $p^{(3)} < \theta/8$, when*

$$N_1 \geq C_7\varepsilon_1^{-n-2}(\log(\frac{1}{\delta_1}) + \log(\frac{1}{\theta})). \tag{4.27}$$

Proof. Below, we use that $t \leq -\log(1-t)$ for $0<t<1$. We see that $p^{(3)} < \theta/8$ if

$$N_1 \geq R_1 = \frac{2^{2n+4}e^{2\beta}\log(\frac{16N_0^2}{\theta})}{\phi_1 c_3\rho^n h_0^2} = \frac{2^{2n+8}e^{2\beta}\log(\frac{16N_0^2}{\theta})}{\phi_1 c_3 c_5^2\rho^n\rho^2}. \tag{4.28}$$

Next we use that $t\log t \le t^2$, so that for $t > e$ we have $\log(t\log t) \le 2\log t$. Also, recall that N_0 is given in (4.26). Then we see that

$$R_1 = \frac{2^{2n+8}e^{2\beta}\log(\frac{16N_0^2}{\theta})}{\phi_1 c_3 c_5^2 \rho^n \rho^2} \le C_9\varepsilon_1^{-n-2}(\log(\frac{1}{\delta_1}) + \log(\frac{1}{\theta})) = P_1$$

where C_9 is suitable. Thus (4.28) is valid when $N_1 \ge P_1$. This yields that claim. □

Lemma 4.7. *There is $C_{10} > 2C_7$ such that we have $p^{(4)} < \theta/8$ when*

$$N_1 = \lfloor C_{10}\varepsilon_1^{-2n}(\log(\frac{1}{\delta_1}) + \log(\frac{1}{\theta}))\rfloor. \tag{4.29}$$

Note that when N_1 is chosen as in (4.29), also the inequality (4.27) is valid.

Proof. Next we estimate $p^{(4)} = 2N_0^2\exp(-2N_1(u_1/2)^2)$. We see that $p^{(4)} < \theta/8$ if

$$N_1 \ge R_2 = \frac{\log(\frac{16N_0^2}{\theta})}{2(u_1/2)^2} = \frac{2\log(\frac{16N_0^2}{\theta})}{\phi_1^2 c_3^2(\rho/4)^{2n}}. \tag{4.30}$$

Also, we see that

$$R_2 \le \lfloor C_{10}\varepsilon_1^{-2n}(\log(\frac{1}{\delta_1}) + \log(\frac{1}{\theta}))\rfloor = P_2,$$

where C_{10} is suitably chosen. Thus (4.30) is valid when $N_1 \ge P_2$. □

Lemma 4.8. *There is $C_{12} > 0$ such that $p^{(2)} < \theta/8$ when*

$$N_2 \ge C_{12}\varepsilon_1^{-4}(\log(\frac{1}{\varepsilon_1}))^4\Big(\log(\frac{1}{\theta}) + \log(\frac{1}{\delta_1}) + \log(\frac{1}{\varepsilon_1})\Big). \tag{4.31}$$

Proof. Let N_0 and N_1 are given in (4.26) and (4.29). We see that $p^{(2)} \le \theta/8$ if

$$N_2 \ge R_3 = \frac{1}{2}\varepsilon_2^{-2}L^2\log(\frac{16N_0N_1}{\theta}). \tag{4.32}$$

We have

$$\begin{aligned} R_3 &\le C_{13}\varepsilon_1^{-4}L^2\log\Big(\frac{16}{\theta}\cdot 2C_3\delta_1^{-n}(\log(\frac{1}{\delta_1}) + \log(\frac{1}{\theta}))\cdot C_{10}\varepsilon_1^{-2n}(\log(\frac{1}{\delta_1}) + \log(\frac{1}{\theta}))\Big) \\ &\le C_{12}\varepsilon_1^{-4}(\log(\frac{1}{\varepsilon_1}))^4\Big(\log(\frac{1}{\theta}) + \log(\frac{1}{\delta_1}) + \log(\frac{1}{\varepsilon_1})\Big) = P_3, \end{aligned}$$

where C_{12} and C_{13} are suitable. Thus (4.32) is valid when $N_2 \ge P_3$. This yields the claim. □

Lemma 4.9. *There is $C_{14} > 0$ such that we have $p^{(1)} < \theta/8$ when*

$$N_2 \ge C_{14}\varepsilon_1^{-2n}\Big(\log(\frac{1}{\theta}) + \log(\frac{1}{\delta_1}) + \log(\frac{1}{\varepsilon_1})\Big). \tag{4.33}$$

Proof. Using (4.25), we see that the inequality

$$p^{(1)} = 2N_0N_1\exp(-2N_2\varepsilon_3^2) = 2N_0N_1\exp(-2N_2\frac{\phi_1^2}{2^{4n+8}}c_3^2c_4^2\rho^{2n}) < \frac{1}{8}\theta \tag{4.34}$$

is valid when $N_2 \ge R_4 = 2^{4n+7}\phi_1^{-2}c_3^{-2}c_4^{-2}\rho^{-2n}\ \log(\frac{16N_0N_1}{\theta})$. We see that

$$\begin{aligned} R_4 &\le \frac{2^{4n+7}}{\phi_1^2c_3^2c_4^2\rho^{2n}}\log\Big(\frac{16}{\theta}\cdot 2C_3\delta_1^{-n}(\log(\frac{1}{\delta_1}) + \log(\frac{1}{\theta}))\cdot C_{10}\varepsilon_1^{-2n}(\log(\frac{1}{\delta_1}) + \log(\frac{1}{\theta}))\Big) \\ &\le C_{14}\varepsilon_1^{-2n}\Big(\log(\frac{1}{\theta}) + \log(\frac{1}{\delta_1}) + \log(\frac{1}{\varepsilon_1})\Big) = P_4, \end{aligned}$$

where C_{14} is suitable. Thus (4.34) is valid when $N_2 \ge P_4$. This yields the claim. □

Next we prove Theorems 1 and 2.

Proof (of Theorem 2). We observe that when $\mathcal{E}^{(5)}$ happens, by Lemma 4.2 and Lemma 4.4, for all X_j and $X_{j'}$ such that $d_M(X_j, X_{j'}) < r_1$ we have $W_{j,j'} \geq u_2$.

Let N_0 and N_1 be given in (4.26) and (4.29). The conditions (4.33) and (4.31) are valid when we choose a suitable $C_{15} > 1$ and

$$N_2 \geq \lfloor C_{15}\varepsilon_1^{-2n}\Big(\log^2(\frac{1}{\theta}) + \log^2(\frac{1}{\delta_1}) + \log^8(\frac{1}{\varepsilon_1})\Big)\rfloor \tag{4.35}$$

Then $p^{(5)} = p^{(1)} + p^{(2)} + p^{(3)} + p^{(4)} \leq \frac{1}{2}\theta$. This and Lemma 2.1 prove that with probability $1-\theta$ we have for all $(j,j') \in I^{(0)} \times I^{(0)}$ that inequality (2.2) holds when $d_M(X_j, X_{j'}) < r_1$, and the inequality (2.3) holds when $d_M(X_j, X_{j'}) \geq r_1$.

In the case when $\Phi(x,y) \geq c_1\phi_0$, for $(x,y) \in M \times M$, we see that when the events $\mathcal{E}^{(5)}$ happens, Lemmas 4.2 and 4.4 yield that we will have with probability $1-\theta$ that $W_{j,j'} > u_2$ for all (j,j'). This implies that the inequality (2.2) holds for all pairs $(X_j, X_{j'})$ with $j,j' \in \{1,2,\dots,N_0\}$. □

Proof (of Thm. 1) Let $\varepsilon_1 = \delta^{3/2}$ and $\delta_1 = \Lambda^{2/3}\delta^{1/2}/20$. For N_0, N_1 given in (4.26) and (4.29),

$$N_0 \leq \lfloor C_{16}\delta^{-n/2}\Big(\log(\frac{1}{\theta}) + \log(\frac{1}{\delta})\Big)\rfloor, \quad N_1 \leq \lfloor C_{17}\delta^{-3n}\Big(\log^2(\frac{1}{\theta}) + \log^8(\frac{1}{\delta})\Big)\rfloor, \tag{4.36}$$

with suitable C_{16} and C_{17}. Moreover, N_2 satisfies (4.35) when we choose a suitable C_{18} and

$$N_2 = \lfloor C_{18}\delta^{-3n}\Big(\log^2(\frac{1}{\theta}) + \log^8(\frac{1}{\delta})\Big)\rfloor. \tag{4.37}$$

Let $\widehat{\delta} = \varepsilon_1 = \delta^{3/2}$ and $\widehat{r} = (\widehat{\delta}/\Lambda^2)^{1/3} = \Lambda^{2/3}\delta^{1/2}$. Then by Theorem 2, with probability $1-\theta$ the set $\mathcal{X} = \{X_j : \ j = 1,2,\dots,N_0\}$ is a δ_1-dense subset of M and the approximate distances $\widetilde{d}(X_j, X_{j'}) = d^{app}(X_j, X_{j'})$, $j,j' \in \{1,2,\dots N_0\}$, see (4.5), satisfy the conditions given in Proposition 4.10 in Appendix A. Thus with probability $1-\theta$ can apply Proposition 4.10 (see also [24, Corollary 1.10]) with $\widehat{\delta} = \delta^{3/2}$ and $\widehat{r} = (\widehat{\delta}/\Lambda^2)^{1/3}$ to construct a Riemannian manifold (M^*, g^*) that approximates the original manifold (M,g) so that the claims (1)-(3) in Theorem 1 are satisfied. □

Appendix A: Reconstruction of a manifold with a small deterministic errors

Here, we give results on the reconstruction of a Riemannian manifold when one is given distances with small deterministic errors. The following result is an improvement of Corollary 1.10 in [24].

Proposition 4.10. *There are $C_n' > 0$, depending on n, and $c_1'(n,K) > 0$, depending on n, K, such that the following holds: Let $0 < \widehat{\delta} < c_1'(n,K)$, $\widehat{r} = (\widehat{\delta}/K)^{1/3}$ and M be a compact n-dimensional manifold with $|\mathrm{Sec}(M)| \leq K$ and $\mathrm{inj}(M) > 2\widehat{r}$. Let $\mathcal{X} = \{x_j\}_{j=1}^N$ be an $\widehat{r}/20$-dense subset of M. Moreover, let $\widetilde{d}\colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+ \cup \{0\}$ be an approximate local distance function that satisfies*

$$|\widetilde{d}(x,y) - d_M(x,y)| \leq \widehat{\delta}, \quad \text{if } d_M(x,y) < \widehat{r}, \tag{4.38}$$

$$\widetilde{d}(x,y) > \widehat{r} - \widehat{\delta}, \quad \text{if } d_M(x,y) \geq \widehat{r}. \tag{4.39}$$

Then, given the values $\widetilde{d}(x_j, x_k)$, $j,k = 1,2,\dots,N$, one can construct a compact n-dimensional Riemannian manifold (M^*, g^*) such that:

(1) *There is a diffeomorphism $F : M^* \to M$ satisfying*

$$\frac{1}{L} \leq \frac{d_M(F(x), F(y))}{d_{M^*}(x,y)} \leq L, \quad \text{for } x,y \in M^*, \ L = 1 + C_n' K^{1/3}\widehat{\delta}^{\,2/3}.$$

(2) $|\mathrm{Sec}(M^*)| \leq C_n' K$.

(3) *The injectivity radius $\mathrm{inj}(M^*)$ of M^* satisfies*

$$\mathrm{inj}(M^*) \geq \min\{(C_n'K)^{-1/2}, (1 - C_n'K^{1/3}\widehat{\delta}^{\,2/3})\,\mathrm{inj}(M)\}.$$

Proof. A result similar to the claim is proven in [24, Corollary 1.10] under the assumption that the set $\mathcal{X}$ is a $\widehat{\delta}$-dense subset of M, instead of $\widehat{r}/20$-dense as it is assumed in the claim. Moreover, by [24, Corollary 1.10], it is enough to construct numbers $\widetilde{D}_{j,k}$, $j,k=1,2,\dots,\widetilde{N}$, such that the following is true: There a $\widehat{\delta}$-net $\mathcal{Y}=\{y_j:\ j=1,2,\dots,\widetilde{N}\}\subset M$ such that the conditions (4.38) and (4.39) are valid for the function $\widetilde{d}'\colon \mathcal{Y}\times\mathcal{Y}\to\mathbb{R}_+\cup\{0\}$ defined by $\widetilde{d}'(y_j,y_k)=\widetilde{D}_{j,k}$.

Next we construct the required $\widehat{\delta}$-net $\mathcal{Y}\subset M$ and an approximate distance function $\widetilde{d}'$ on $\mathcal{Y}\times\mathcal{Y}$. We assume that $c_1'(n,K)$ is chosen so small that $\widehat{\delta}/\widehat{r}=K\widehat{r}^2<\frac{1}{150}$.

For $p\in M$ we denote by E_p the restriction of the Riemannian exponential map $\exp_p$ to the $\widehat{r}$-ball in T_pM centered at the origin. This restriction is a diffeomorphism onto the $\widehat{r}$-ball centered at p in M. It distorts distances by at most $\frac{1}{2}\widehat{\delta}$, namely for all $u,v\in T_pM$ such that $|u|,|v|<\widehat{r}$ we have

$$|d_M(E_p(u),E_p(v))-|u-v||<\tfrac{1}{2}K\widehat{r}^3=\tfrac{1}{2}\widehat{\delta}. \tag{4.40}$$

This inequality holds as long as $\operatorname{inj}(M)>2\widehat{r}$ and $K\widehat{r}^2<\pi/2$, see [24, Section 4] for a proof.

For every $p\in\mathcal{X}$, define $X_p=\{x\in\mathcal{X}:\widetilde{d}(p,x)<\widehat{r}/6-\widehat{\delta}\}$. By (4.38) and (4.39), X_p is contained in the $\widehat{r}/6$-neighborhood of p. Define $\widetilde{X}_p=E_p^{-1}(X_p)$, let $\widetilde{V}_p$ be the convex hull of $\widetilde{X}_p$ in T_pM, and $V_p=E_p(\widetilde{V}_p)$. Since $\mathcal{X}$ is $\widehat{r}/20$-dense in M, (4.38) and (4.40) imply that for every $u\in T_pM$ such that $|u|<\widehat{r}/6-\widehat{r}/20-2\widehat{\delta}$ there exists $v\in\widetilde{X}_p$ such that $|u-v|<\widehat{r}/20+\widehat{\delta}/2$. This implies that $\widetilde{V}_p$ contains the ball of radius $\widehat{r}/6-2\widehat{r}/20-2\widehat{\delta}-\widehat{\delta}/2>\widehat{r}/20$ centered at the origin. (Here we use the assumption that $\widehat{\delta}/\widehat{r}<\frac{1}{150}$). Hence V_p contains the $\widehat{r}/20$-ball centered at p and therefore $\bigcup_{p\in\mathcal{X}}V_p=M$.

We represent points of $\widetilde{V}_p$ as linear combinations of points of $\widetilde{X}_p$ as follows. Let X_p^n be the set of all n-tuples of points of X_p and Δ^n the standard coordinate simplex in $\mathbb{R}^n$:

$$\Delta^n=\{(t_1,\dots,t_n)\in\mathbb{R}^n:t_1,\dots,t_n\ge 0,\sum t_i\le 1\}.$$

For $\alpha=(a_1,\dots,a_n)\in X_p^n$ and $\tau=(t_1,\dots,t_n)\in\Delta^n$, let

$$S_p(\alpha,\tau)=\sum t_iE_p^{-1}(a_i).$$

This defines a map $S_p\colon X_p^n\times\Delta^n\to T_pM$. For a fixed $\alpha=(a_1,\dots,a_n)\in X_p^n$, the range of $S_p(\alpha,\cdot)$ is a (possibly degenerate) affine simplex in T_pM with vertices $0,E_p^{-1}(a_1),\dots,E_p^{-1}(a_n)$. Since $0\in\widetilde{X}_p$, the union of all such affine simplices is precisely the convex hull of $\widetilde{X}_p$. Thus $S_p(X_p^n\times\Delta^n)=\widetilde{V}_p$.

Fix an ε'-dense finite set $\Sigma\subset\Delta^n$, where $\varepsilon'=\widehat{\delta}/(3\widehat{r}\sqrt{n})$, and define $Y_p=E_p(S_p(X_p^n\times\Sigma))\subset M$. Since $\widetilde{V}_p$ is contained in the $\widehat{r}/6$-ball, $S_p(\alpha,\cdot)$ is Lipschitz with Lipschitz constant $\widehat{r}\sqrt{n}/6$. Therefore $S_p(X_p^n\times\Sigma)$ is $\widehat{\delta}/2$-dense in $\widetilde{V}_p$. Hence, by (4.40), Y_p is $\widehat{\delta}$-dense in V_p.

Now define $\mathcal{Y}\subset M$ by $\mathcal{Y}=\bigcup_{p\in\mathcal{X}}Y_p$. Since the sets V_p cover M and Y_p is $\widehat{\delta}$-dense in V_p for each p, $\mathcal{Y}$ is a $\widehat{\delta}$-net in M. The points of $\mathcal{Y}$ are indexed by triples (p,α,τ) where $p\in\mathcal{X}$, $\alpha\in X_p^n$, $\tau\in\Sigma$. This index set can be enumerated algorithmically using the known data.

Our first goal is to compute approximate *squared* distances $Q(x,y)$ between sufficiently close pairs of points $x,y\in\mathcal{Y}$. Fix $p,q\in\mathcal{X}$ such that $\widetilde{d}(p,q)<2\widehat{r}/3-2\widehat{\delta}$ (the case $p=q$ is not excluded). By (4.38) and (4.39) we have $d_M(p,q)<2\widehat{r}/3-\widehat{\delta}$. Hence, by the triangle inequality, $d_M(x,y)<\widehat{r}-\widehat{\delta}$ for all $x\in V_p$ and $y\in V_q$, In particular V_q is contained in the range of E_p. By (4.40),

$$\left|d_M(x,y)^2-|E_p^{-1}(x)-E_p^{-1}(y)|^2\right|<\widehat{\delta}\widehat{r}\qquad\text{for all } x\in V_p \text{ and } y\in V_q. \tag{4.41}$$

We compute the values $Q(x,y)$ for all $x\in X_p\cup Y_p$ and $y\in X_q\cup Y_q$ in several steps. First consider $x\in X_p$ and $y\in X_q$. In this case we simply define $Q(x,y)=\widetilde{d}(x,y)^2$. Then, by (4.38),

$$|Q(x,y)-d_M(x,y)^2|<2\widehat{\delta}\widehat{r}. \tag{4.42}$$

Hence, by (4.41),

$$\left|Q(x,y)-|E_p^{-1}(x)-E_p^{-1}(y)|^2\right|<3\widehat{\delta}\widehat{r},\qquad x\in X_p,\ y\in X_q. \tag{4.43}$$

Now consider $x \in Y_p$ and $y \in X_q$. By the definition of Y_p we have $x = E_p(S_p(\alpha, \tau))$ for some $\alpha = (a_1, \dots, a_n) \in X_p^n$ and $\tau = (t_1, \dots, t_n) \in \Sigma$. We define $Q(x, y)$ using the values of Q that we have from the previous step. Introduce the following notation: $a_0 = p$, $t_0 = 1 - \sum_{i=1}^n t_i$, $v_i = E_p^{-1}(a_i)$ for $i = 0, \dots, n$ (in particular $v_0 = 0$), $v = E_p^{-1}(x) = \sum t_i v_i$, and $w = E_p^{-1}(y)$. In this notation, $v - w = \sum_{i=0}^n t_i(v_i - w)$, hence

$$|v - w|^2 = \sum_{0 \le i,j \le n} t_i t_j \langle v_i - w, v_j - w \rangle = \frac{1}{2} \sum_{0 \le i,j \le n} t_i t_j (|v_i - w|^2 + |v_j - w|^2 - |v_i - v_j|^2)$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in T_pM. With this identity in mind, we define

$$Q(x, y) = \frac{1}{2} \sum_{0 \le i,j \le n} t_i t_j (Q(a_i, y) + Q(a_j, y) - Q(a_i, a_j)). \tag{4.44}$$

Since $y \in X_q$ and $a_i, a_j \in X_p$, the values $Q(a_i, y)$, $Q(a_k, y)$ and $Q(a_i, a_j)$ are defined in the previous step (in the case of $Q(a_i, a_j)$, this is the previous step with $q = p$). Since $\sum_{i=0}^n t_i = 1$, applying (4.43) to the terms in the right-hand side of (4.44) yields that

$$|Q(x, y) - |E_p^{-1}(x) - E_p^{-1}(y)|^2| = |Q(x, y) - |v - w|^2| < \tfrac{9}{2}\widehat{\delta}\widehat{r} < 5\widehat{\delta}\widehat{r}. \tag{4.45}$$

Therefore, by (4.41),

$$|Q(x, y) - d_M(x, y)^2| < 6\widehat{\delta}\widehat{r}. \tag{4.46}$$

We now have the values $Q(x, y)$ satisfying (4.46) for all $x \in Y_p$ and $y \in X_q$. Exchanging the roles of p and q we similarly find $Q(x, y)$ for all $x \in X_p$ and $y \in Y_q$.

Finally, consider $x \in Y_p$ and $y \in Y_q$. Again, let $x = E_p(S_p(\alpha, \tau))$ where $\alpha = (a_1, \dots, a_n) \in X_p^n$ and $\tau = (t_1, \dots, t_n) \in \Sigma$, and define $a_0 = p$ and $t_0 = 1 - \sum_{i=1}^n t_i$. From the previous steps we already have values $Q(a_i, y)$ and $Q(a_i, a_j)$. Therefore we can define $Q(x, y)$ by the same formula (4.44). Then, starting from (4.46) instead of (4.42), we obtain the same estimates as above but with different constants: (4.43) with $7\widehat{\delta}\widehat{r}$ in the right-hand side, (4.45) with $11\widehat{\delta}\widehat{r}$ in the right-hand side, and finally (4.46) with $12\widehat{\delta}\widehat{r}$ in the right-hand side:

$$|Q(x, y) - d_M(x, y)^2| < 12\widehat{\delta}\widehat{r}, \qquad x \in Y_p,\ y \in Y_q. \tag{4.47}$$

Now one might take the square root of $Q(x, y)$ as an approximate distance between x and y; however this approximation is not good enough. For a better one, we use an algorithm described in [24, §2.4] to construct a map $F \colon Y_p \cup Y_q \to \mathbb{R}^n$ that preserves distances up to an error $O(\widehat{\delta})$. Let us outline how the algorithm works in the present set-up.

First define an approximate scalar product $P(x, y)$ for all pairs $x, y \in Y_p \cup Y_q$ by

$$P(x, y) = \tfrac{1}{2}(Q(p, x) + Q(p, y) - Q(x, y)).$$

By (4.41), (4.47) and the Euclidean identity $\langle u, v \rangle = \frac{1}{2}(|u|^2 + |v|^2 - |u - v|^2)$ for $u, v \in T_pM$, this approximates the scalar product of $E_p^{-1}(x)$ and $E_p^{-1}(y)$ in T_pM:

$$|P(x, y) - \langle E_p^{-1}(x), E_p^{-1}(y) \rangle| < 20\widehat{\delta}\widehat{r}. \tag{4.48}$$

Then, since $E_p^{-1}(X_p)$ is a $\widehat{\delta}/2$-net in $E_p^{-1}(V_p)$ and the latter contains the $\widehat{r}/6$-ball centered at the origin, we can find points $a_1, \dots, a_n \in Y_p$ such that the vectors $v_i := E_p^{-1}(a_i)$, $i = 1, \dots, n$, approximate an orthonormal basis of T_pM rescaled by the factor $\widehat{r}/6$:

$$|(\widehat{r}/6)^{-2} \langle v_i, v_j \rangle - \delta_{ij}| < C_1 \widehat{\delta}/\widehat{r}, \qquad 1 \le i, j \le n, \tag{4.49}$$

where δ_{ij} is the Kronecker delta and $C_1 = C_1(n) > 0$ is a suitable constant. (A straightforward modification of the algorithm from [24, §2.4] can be used to find such points efficiently).

The inequalities (4.49) imply that the linear map $L \colon T_pM \to \mathbb{R}^n$ defined by

$$L(v) = (\widehat{r}/6)^{-1}(\langle v, v_1 \rangle, \dots, \langle v, v_n \rangle)$$

is $(C_2\widehat{\delta}/\widehat{r})$-close in the operator norm to a linear isometry from T_pM to $\mathbb{R}^n$ for some constant $C_2 = C_2(n) > 1$, see [24, Lemma 2.6]. Hence L distorts distances within the $\widehat{r}$-ball by at most $2C_2\widehat{\delta}$. We approximate $L \circ E_p^{-1}$ by a map $F\colon Y_p \cup Y_q \to \mathbb{R}^n$ defined by

$$F(x) = (\widehat{r}/6)^{-1}(P(x, a_1), \dots, P(x, a_n)), \qquad x \in Y_p \cup Y_q,$$

and compute $\widetilde{d}'(x, y) = |F(x) - F(y)|$ for all $x, y \in Y_p \cap Y_q$. By (4.48) we have $|F(x) - L(E_p^{-1}(x))| < 120\sqrt{n}\,\widehat{\delta}$ for all $x, y \in Y_p \cap Y_q$. Hence, by (4.40) and the above mentioned property of L,

$$|\widetilde{d}'(x, y) - d_M(x, y)| < C_4\widehat{\delta} \tag{4.50}$$

where $C_4 = 2C_2 + 120\sqrt{n} + 1$.

The domain of the function $\widetilde{d}'$ defined by the above procedure includes all pairs $x, y \in \mathcal{Y}$ such that $d_M(x, y) < \widehat{r}/4$. Indeed, if $d_M(x, y) < \widehat{r}/4$ and $p, q \in X$ are such that $x \in Y_p$, $y \in Y_q$, then by the triangle inequality we have $d_M(p, q) < \widehat{r}/4 + 2\widehat{r}/6 < 2\widehat{r}/3 - 3\widehat{\delta}$, and hence, by (4.38), $\widetilde{d}(p, q) < 2\widehat{r}/3 - 2\widehat{\delta}$. Thus for any such pair x, y the value $\widetilde{d}'(x, y)$ is defined and satisfies (4.50).

To finish the construction, set $\widetilde{d}'(x, y) = \widehat{r}$ for all remaining pairs $x, y \in \mathcal{Y}$. Now the function $\widetilde{d}'$ is defined on $\mathcal{Y} \times \mathcal{Y}$ and it satisfies the assumptions [24, Corollary 1.10] for $\widehat{r}' = \widehat{r}/4$ in place of $\widehat{r}$, $\widehat{\delta}' = C_4\widehat{\delta}$ in place of $\widehat{\delta}$, and $K' = \widehat{\delta}'/(\widehat{r}')^3 = 2^6 C_4 K$ in place of K. Applying [24, Corollary 1.10] with these modified parameters finishes the proof of Proposition 4.10. □

The constructions in Prop. 4.10 and [24, Corollary 1.10] are algorithmic, for the details, see [24].

References

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

[2] M. Belkin, P. Niyogi, Laplacian eigenmaps and spectral techniques for embedding and clustering, *Adv. in Neural Inform. Process. Systems*, 14 (2001), 586–691.

[3] M. Bernstien, V. de Silva, J. Langford, J. Tenenbaum, *Graph approximations to geodesics on embedded manifolds.* Technical Report, Stanford University, 2000.

[4] M. Brand, Charting a manifold, *NIPS 15* (2002), 985–992.

[5] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry.* AMS, 2001. xiv+415 pp.

[6] M. Carreira-Perpinan, Z. Lu, Manifold Learning and Missing Data Recovery through Unsupervised Regression *11th IEEE International Conference on Data Mining, ICDM 2011,* Vancouver, 2011

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

[8] D. Chigirev and W. Bialek, Optimal Manifold Representation of Data: An Information Theoretic Approach. In: *Advances in Neural Information Processing Systems* 16, Ed. S. Thrun et al, The MIT press, 2004, pp. 164–168.

[9] R. Coifman, et al. Geometric diffusions as a tool for harmonic analysis and structure definition of data Part II: Multiscale methods. *Proc. of Nat. Acad. Sci.* 102 (2005), 7432–7438.

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

[11] T. Cox, M. Cox, *Multidimensional Scaling.* Chapman & Hall, London, (1994).

[12] P. Erdos, A. Renyi, On a classical problem of probability theory. *Magyar Tud. Akad. Mat.* 6 1961 215–220.

[13] L. Flatto, D. Newman, Random coverings. *Acta Math.* 138 (1977), 241–264.

[14] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.* **57** (2004), 11.

[15] M. de Hoop, P. Kepley, L. Oksanen, On the construction of virtual interior point source travel time distances from the hyperbolic Neumann-to-Dirichlet map. *SIAM J. Appl. Math.* 76 (2016), 805–825.

[16] D. Donoho, C. Grimes, When does geodesic distance recover the true hidden parametrization of families of articulated images? *Proceedings of ESANN 2002*, Bruges, Belgium, 2002.

[17] D. Donoho, D. Grimes, Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data, *Proceedings of the National Academy of Sciences, 100*, 5591–5596.

[18] D. Donoho, C. Grimes, Image Manifolds which are Isometric to Euclidean Space. *J. Math. Im. Vis* 23 (2005), 5.

[19] A. Gilbert and R. Sonthalia, Unrolling Swiss Cheese: Metric repair on manifolds with holes, *Proceedings of the 56th Annual Allerton Conference on Communication, Control, and Computing*, 2018.

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

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

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

[23] Ch. Fefferman, S. Ivanov, Y. Kurylev, M. Lassas, H. Narayanan, *Fitting a Putative Manifold to Noisy Data. Proceedings of the 31st Conference On Learning Theory, in PMLR.* 75688–720, 2018

[24] C. Fefferman, S. Ivanov, Y. Kurylev, M. Lassas, H. Narayanan, Reconstruction and interpolation of manifolds I: The geometric Whitney problem. arXiv:1508.00674.
[25] Ch. Fefferman, B. Klartag, Fitting C^m-smooth function to data I, *Ann. of Math.* 169 (2009), 315–346.
[26] Ch. Fefferman, S. Mitter, H. Narayanan, Testing the manifold hypothesis, *J. Amer. Math. Soc.* 29 (2016), 983–1049.
[27] W. Hoeffding, Probability inequalities for sums of bounded random variables. *JASA* 58 (1963), 13–30.
[28] P. Hoskins: Principles of ultrasound elastography, *Ultrasound* (20) 2012, 8–15.
[29] D. Joncas, M. Meila J. McQueen, Improved Graph Laplacian via Geometric Self-Consistency. NIPS 30 (2017)
[30] P. Jones, M., Maggioni, R. Schul, Universal local parametrizations via heat kernels and eigenfunctions of the laplacian, *Ann. Acad. Scient. Fen. 35* (2010), 1–44.
[31] O. Kallenberg, *Foundations of modern probability theory.* Springer (1997).
[32] A. Katchalov, Y. Kurylev, and M. Lassas, *Inverse Boundary Spectral Problems.* CRC, 2001, 290pp.
[33] Y. Kurylev, L. Oksanen, G. Paternain: Inverse problems for the connection Laplacian. *J. Diff. Geom.* 110 (2018), 457-494.
[34] M. Lassas, E. Saksman, S. Siltanen, Discretization-invariant Bayesian inversion and Besov space priors, *Inverse Problems and Imaging* 3 (2009), 87-122.
[35] J. Lee, G. Uhlmann, Determining anisotropic real-analytic conductivities by boundary measurements, *Comm. Pure Appl. Math.* 42 (1989), 1097–1112.
[36] T. Lin, H. Zha, Riemannian manifold learning. *IEEE Trans Pattern Anal Mach Intell.* 5 (2008), 796-809.
[37] T. Lin, H. Zha, S. Lee, Riemannian Manifold Learning for Nonlinear Dimensionality Reduction. In: Leonardis A., Bischof H., Pinz A. (eds) Computer Vision. ECCV 2006. Lecture Notes in CS, vol 3951. Springer.
[38] J. Mueller, S. Siltanen, *Linear and nonlinear inverse problems with practical applications.* SIAM, Philadelphia, 2012. xiv+351 pp.
[39] J. Nash, C^1-isometric imbeddings, *Ann. of Math.* 60 (1954), 383–396.
[40] J. Nash, The imbedding problem for Riemannian manifolds, *Ann. of Math.* 63 (1956), 20–63.
[41] D. Newman, L. Shepp, The double dixie cup problem. *Amer. Math. Monthly* 67 1960 58–61.
[42] D. Perraul-Joncas, M. Meila, Non-linear dimensionality reduction: Riemannian metric estimation and the problem of geometric discovery, arXiv:1305-7255, 2013.
[43] L. Pestov and G. Uhlmann, Two dimensional compact simple Riemannian manifolds are boundary distance rigid. *Ann. of Math.* 161 (2005), 1093–1110.
[44] K. Pearson, On lines and planes of closest fit to systems of points in space, *Phil. Mag.* 2 (1901), 559–572.
[45] P. Petersen, *Riemannian geometry.* 1st Ed. Springer, 1998. xvi+432 pp.
[46] X. Qu, et al., Bent ray ultrasound tomography reconstruction using virtual receivers for reducing time cost, Proc. SPIE 9419, Medical Imaging 2015: Ultrasonic Imaging and Tomography, 94190F.
[47] S. Roweis, L. Saul, Nonlinear dimensionality reduction by locally linear embedding *Science*, 290 (2000), 2323.
[48] S. Roweis, L. Saul, G. Hinton, Global coordination of local linear models, *Advances in Neural Information Processing Systems 14* (2001) 889–896.
[49] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms. *Phys. D* 60 (1992), 259.
[50] J. Shawe-Taylor, N. Christianini, *Kernel Methods for Pattern Analysis,* Cambridge University Press, (2004).
[51] J. Sylvester, G. Uhlmann, A global uniqueness theorem for an inverse boundary value problem, *Ann. of Math.* 125 (1987), 153–169.
[52] P. Stefanov, G. Uhlmann, A. Vasy, Boundary Rigidity with Partial Data. Journal of the AMS 29 (2016), 299–332.
[53] J. Tenenbaum, V. de Silva, J. Langford, *A global geometric framework for nonlinear dimensionality reduction*, Science, 290 (2000), 2319–2323.
[54] 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.
[55] N. Verma, Distance preserving embeddings for general n-dimensional manifolds. (aka An algorithmic realization of Nash's embedding theorem), *Journal of Machine Learning Research* 23 (2012) 32.1–32.28.
[56] H. Zha and Z. Zhang, Continuum Isomap for manifold learnings. Comp. Stat. Data Anal. 52 (2007), 184-200.
[57] Z. Zhang and H. Zha, Principal manifolds and nonlinear dimension reduction via local tangent space alignment, *SIAM J. Sci. Computing*, **26** (2005), 313–338.

Charles Fefferman, Princeton University, Mathematics Department, USA.
Sergei Ivanov, St. Petersburg Department of Steklov Institute of Mathematics, Russia.
Matti Lassas, University of Helsinki, Finland.
Hariharan Narayanan, Tata Institute for Fundamental Research, India.