

# RECONSTRUCTING RIEMANNIAN METRICS FROM RANDOM GEOMETRIC GRAPHS

HAN HUANG, PAKAWUT JIRADILOK, AND ELCHANAN MOSSEL

**ABSTRACT.** Random geometric graphs are random graph models defined on metric measure spaces. A random geometric graph is generated by first sampling points from a metric space and then connecting each pair of sampled points independently with a probability that depends on their distance.

In recent work of Huang, Jiradilok, and Mossel [HJM24], the authors study the problem of reconstructing an embedded manifold from a random geometric graph sampled from the manifold, where edge probabilities depend monotonically on the Euclidean distance between the embedded points. They show that, under mild regularity assumptions on the manifold, the sampling measure, and the connection probability function, it is possible to recover the pairwise Euclidean distances of the embedded sampled points up to a vanishing error as the number of vertices grows.

In this work we consider a similar and arguably more natural problem where the metric is the Riemannian metric on the manifold. Again points are sampled from the manifold and a random graph is generated where the connection probability is monotone in the Riemannian distance. Perhaps surprisingly we obtain stronger results in this setup. Unlike the previous work that only considered dense graph we provide reconstruction algorithms from sparse graphs with average degree  $n^{1/2} \text{polylog}(n)$ , where  $n$  denotes the number of vertices. Our algorithm is also a more efficient algorithm for distance reconstruction with improved error bounds. The running times of the algorithm is  $O(n^2 \text{polylog}(n))$  which up to polylog factor matches the size of the input graph. Our distance error also nearly matches the volumetric lower bounds for distance estimation.

## 1. INTRODUCTION

Random graphs provide a powerful framework for modeling complex networks. The classical Erdős–Rényi model  $G(n, p)$ , in which edges appear independently with probability  $p$ , remains a cornerstone of modern probabilistic combinatorics, offering analytic clarity for studying connectivity, typical distances, and phase transitions. Comprehensive references include [Bol11, JLR11] and the survey [FK15].

To model more complex and realistic graphs it is natural to consider edge probabilities that depend on spatial or geometric proximity. Random geometric graphs (RGGs) model this phenomenon by associating latent positions to vertices and defining adjacency as a function of distance. In the classical hard-disc model, two vertices  $v, w$  are connected whenever  $\text{dist}(X_v, X_w) \leq r$ . The soft or random-connection model generalizes this by assigning connection probabilities  $p(\text{dist}(X_v, X_w))$  that decay with distance, reflecting fading or noisy interactions as observed for example in wireless and spatial networks [Pen03, DG16].

An exciting new research direction in the study of RGGs concerns the extent to which one can infer latent geometric or probabilistic structure from the observed connectivity pattern. Earlier work established detectability thresholds distinguishing geometric graphs from Erdős–Rényi graphs in various regimes [BDER16, LMSY22, BBN20]. Other studies have explored nonparametric inference for translation-invariant RGGs on spheres, aiming to recover the connection kernel  $p(\langle x, y \rangle)$  from a single graph [ADC19, EMP22, DDC23]. A more recent work [HJM24] considered the ambitious task of reconstructing the Euclidean distance between sampled points in a manifold given a random geometric graph where the edge probabilities monotonically depend on the Euclidean

---

*Date:* November 10, 2025.

distance between points. There, it was shown that under mild regularity assumptions on a submanifold  $M \subset \mathbb{R}^N$ , its sampling measure  $\mu$ , and the distance–probability function  $p$ , both intrinsic and extrinsic distances between latent points can be recovered up to vanishing error. This result established that geometric information is, in principle, encoded in the combinatorial structure of random geometric graphs.

In the work of [HJM24], the authors consider  $M$  to be a smooth, compact, connected  $d$  dimensional submanifold of  $\mathbb{R}^N$  without boundary, equipped with a probability measure  $\mu$  satisfying the following lower Ahlfors regularity condition (see Assumption 1.4). The distance–probability function  $p : [0, \text{diam}(M)] \rightarrow [0, 1]$  is assumed to be non-increasing and bi-Lipschitz continuous. The random graph  $G$  is generated as follows. Sample  $X_1, \dots, X_n \in M$  independently from  $\mu$ . Each vertex  $i$  corresponds to the latent point  $X_i$ , and for each distinct pair  $i, j$ , include an edge with probability  $p(\text{dist}(X_i, X_j))$ , where  $\text{dist}(\cdot, \cdot)$  denotes the Euclidean distance in the ambient space  $\mathbb{R}^N$ .

The main algorithm of [HJM24] assumes that the distance–probability function  $p$ , the intrinsic dimension  $d$ , curvature bounds of  $M$ , and a quantitative non-self-intersection bound are known. Given only the observed graph  $G$  and this auxiliary information, the algorithm produces estimates  $d(i, j)$  for the Euclidean distances between latent points. Its theoretical guarantees are summarized below.

**Theorem 1.1** ([HJM24], Theorem 1.1). *The Algorithm HJM24 runs in time  $O(n^3)$  and, with probability at least  $1 - n^{-\omega(1)}$ , produces estimates satisfying*

$$|d(i, j) - \text{dist}(X_i, X_j)| \leq Cn^{-\varepsilon/d},$$

for all vertex pairs  $i, j$ , where  $\varepsilon \in (0, 1)$  is an absolute constant and  $C$  depends only on geometric properties of  $M$  (curvature and diameter bounds, separation parameter, and regularity parameters of  $\mu$ ) as well as on the bi-Lipschitz constants of  $p$ .

**1.1. Our Contributions and comparison to prior work.** In this paper, we consider a different and arguably more natural variant of the question studied in [HJM24]. Here, we consider the case where the latent space is a compact Riemannian manifold  $(M, g)$  without boundary, equipped with a probability measure  $\mu$  satisfying lower Ahlfors regularity as above. Crucially the probabilities of edges are now determined by a connection probability function  $p$  applied to the intrinsic distance between the points on the manifold. This is arguably a better model, as for example, on Earth it is more natural to measure distances by the distance needed to travel rather by the Euclidean distance in 3 dimensions. Our results also relax the condition on the connection function  $p$  to be Lipschitz continuous, but only bi-Lipschitz on a small interval  $[0, r_p]$  near zero (see Definition 1.3). We introduce a sparsity parameter  $s_n \in (0, 1]$  into the model, allowing each edge to appear independently with probability  $s_n p(|X_i X_j|)$ , where

$$|X_i X_j| := \text{the Riemannian distance between } X_i \text{ and } X_j \text{ on } M.$$

Informally, our main theorem is stated below; see Theorem 1.7 for a formal statement.

**Theorem A** (Informal). *There exists an  $O(n^2 \text{polylog}(n))$ -time algorithm that, given a random geometric graph  $G \sim G(n, M, g, \mu, p, s_n)$ , with sparsity parameter*

$$1 \geq s_n = \omega\left(\frac{\log n}{\sqrt{n}}\right),$$

$p$ , and some bounds on the geometric properties of  $(M, g)$ , the regularity of  $\mu$ , produces an estimator of all pairwise distances. With high probability, all vertex pairs  $v, w$ ,

$$|d(v, w) - \text{dist}(X_v, X_w)| \leq C\left(\frac{\log^2 n}{s_n n}\right)^{\frac{1}{d+1}},$$

where  $C$  is a constant depending only on geometric properties of  $M$  (sectional curvature bounds, injectivity radius, diameter), the regularity parameters of  $\mu$ , and the bi-Lipschitz properties of  $p$ .

To further complement our result, we also establish a lower bound showing that no algorithm can achieve a uniform distance error of order smaller than  $(1/n)^{6/(d+2)}/\log(n)$  with high probability; see Theorem 1.8 below. Compared with Theorem A, the lower bound is different from the upper bound only in the exponent of  $n$  by a factor of approximately  $6/(d+2)$  versus  $1/(d+2)$ , up to logarithmic factors.

In comparison to the previous work [HJM24] we would like to comment that:

- Riemannian distance vs. Euclidean embedded distance. The model we study here is more geometric as the distances depend only on the Riemannian distance rather than on the Euclidean distances which depend on the specifics of the embedding.
- Sparse connectivity regime. We introduce a sparsity parameter  $s_n \in (0, 1]$  so that each edge appears independently with probability  $s_n p(|X_i X_j|)$ . Our analysis covers the regime

$$s_n = \omega\left(\frac{\log n}{\sqrt{n}}\right),$$

which allows for substantially sparser graphs than previously studied.

- Sharper quantitative guarantees to near volumetric limits in dense settings. We derive explicit error bounds that nearly match the volumetric lower bounds for distance estimation with  $n$  samples: With  $n$  points from a  $d$ -dimensional manifold, it cannot form an  $\varepsilon$ -net for  $\varepsilon$  smaller than order  $n^{-\frac{1}{d}}$  via simple volume arguments. Though  $n^{-1/d}$  is not necessarily a lower bound for distance reconstruction error, it does provide a natural benchmark for achievable accuracy.
- Improved algorithmic efficiency to near running time of reading the input graph. We design an efficient reconstruction algorithm that runs in  $O(n^2 \text{polylog}(n))$  time, improving upon the earlier cubic-time procedure. The method is only polylogarithmically slower than reading the input graph itself.

**1.2. Formal Definitions and Statements.** We now present the formal definitions, standing assumptions, and the main results.

**Assumption 1.2** (Manifold). *We assume that  $(M, g)$  is a compact, connected, smooth Riemannian manifold of dimension  $d \geq 2$ . Let  $K(\sigma)$  denote the sectional curvature at a 2-plane  $\sigma \subset TM$ , and define*

$$\kappa := \sup_{\sigma} |K(\sigma)|,$$

*which is finite by compactness of  $M$ . Let  $r_{\text{inj}}(M)$  denote the injectivity radius of  $M$ , which is also positive due to compactness.*

**Assumption 1.3** (Distance-probability function). *Let  $p : [0, \infty) \rightarrow [0, 1]$  be a non-increasing function. We assume there exist positive constants  $L_p, \ell_p, r_p > 0$  such that*

- *for all  $a, b \in [0, r_p]$ ,  $|p(a) - p(b)| \geq \ell_p |a - b|$ , and*
- *for all  $a, b \in [0, \text{diam}(M)]$ ,  $|p(a) - p(b)| \leq L_p |a - b|$ .*

*That is,  $p$  is bi-Lipschitz on  $[0, r_p]$  and globally Lipschitz on  $[0, \text{diam}(M)]$ . We refer to  $p$  as the distance-probability function.*

**Assumption 1.4** (Sampling measure). *Let  $\mu$  be a probability measure on  $M$ , and define*

$$\mu_{\min}(r) := \min_{p \in M} \mu(B_r(p)).$$

We assume that  $\mu$  satisfies a lower Ahlfors regularity condition: there exist constants  $r_\mu > 0$  and  $c_\mu > 0$  such that

$$(1) \quad \mu_{\min}(r) \geq c_\mu r^d, \quad \forall r \in [0, r_\mu].$$

**Definition 1.5** (Random geometric graph). Given a triple  $(M, \mu, p)$ , a vertex set  $\mathbf{V}$  with  $|\mathbf{V}| = n$ , and a sparsity parameter  $s_n \in (0, 1]$ , define the random geometric graph  $G(\mathbf{V}, M, g, \mu, p, s_n)$  as follows. Sample i.i.d. points  $\{X_v\}_{v \in \mathbf{V}}$  from  $\mu$ , and let  $\{\mathcal{U}_{i,j}\}_{i \neq j}$  be i.i.d.  $\text{Uniform}(0, 1)$  random variables. The edge set of  $G$  is

$$E(G) = \left\{ \{i, j\} \in \binom{\mathbf{V}}{2} : \mathcal{U}_{i,j} < s_n p(|X_i X_j|) \right\}.$$

For brevity we write  $G(n, M, g, \mu, p, s_n)$  when  $|\mathbf{V}| = n$ .

**Definition 1.6** (Accessible radius). Define

$$r_G := \frac{1}{16} \min \left\{ r_p, r_{\text{inj}}(M), \frac{1}{\sqrt{\kappa}}, r_\mu \right\}.$$

This constant depends only on  $(M, g)$ ,  $p$ , and  $\mu$ . Within distance  $r_G$ , the manifold is approximately Euclidean,  $\mu$  satisfies lower Ahlfors regularity, and  $p$  behaves bi-Lipschitzly.

**Theorem 1.7.** Assume  $s_n^2 n = \omega(\log^2 n)$ . There exists an algorithm (Algorithm 5) running in time  $O(n^2 \text{polylog}(n))$  such that the following holds.

For any  $(M, g, \mu)$  satisfying Assumptions 1.2 and 1.4, and any distance–probability function  $p$  satisfying Assumption 1.3, the algorithm takes as input

$$G \sim G(\mathbf{V}, M, g, \mu, p, s_n), \quad p, \quad s_n, \quad \text{a lower bound on both } r_G \text{ and } c_\mu,$$

and with probability  $1 - n^{-\omega(1)}$  it outputs estimates  $d(w, v)$  for all  $w, v \in \mathbf{V}$  satisfying

$$| |X_w X_v| - d(w, v) | \leq C \left( \frac{\log^2 n}{s_n n} \right)^{1/(d+2)},$$

for all  $w, v \in \mathbf{V}$ , where  $C$  depends only on  $d, L_p, \ell_p, c_\mu, r_G$ , and  $\text{diam}(M)$ .

**Theorem 1.8** (Lower bound). For every  $n$ , there exist two  $d$ -dimensional Riemannian manifolds  $(M_1, g_1, \mu_1)$  and  $(M_2, g_2, \mu_2)$  together with distance–probability functions  $p$  so that the following holds. First, the corresponding  $r_G, c_\mu$ , and diameters bounded by universal constants (independent of  $n$ ). Second, there exists a coupling  $\pi$  of the random geometric graphs

$$G(n, M_1, g_1, \mu_1, p, 1) \quad \text{and} \quad G(n, M_2, g_2, \mu_2, p, 1),$$

such that with probability  $1 - o_n(1)$  for  $(G_1, G_2) \sim \pi$ , we have  $G_1 = G_2$  and

$$| |X_u^{(1)} X_v^{(1)}|_{M_1} - |X_u^{(2)} X_v^{(2)}|_{M_2} | \geq \frac{1}{\log n} n^{-6/(d+2)}$$

for some  $\{u, v\} \subseteq V(G_1) = V(G_2)$ . Here  $X_u^{(i)}$  denotes the latent point in  $M_i$  corresponding to vertex  $u$  in  $G_i$ .

**1.3. Ideas from the Proof.** We sketch some of the main ideas of the proof. For the discussion below, we assume that  $\mathbf{V}$  denotes the vertex set of the graph. We begin by outlining the main ideas and the overall construction of the algorithm, illustrating the comparison from the approach in [HJM24]. The emphasis in this first part is on the algorithmic aspects and we treat the embedded manifold and the Riemannian manifold settings uniformly by subsuming both under the umbrella term “locally Euclidean.” In the second part, we turn to the geometric aspects and discuss how to extend the approach to the Riemannian manifold setting.

### 1.3.1. Clusters and cluster edge density.

Cluster edge density approach. A natural idea that was already used in [HJM24] is that given a cluster of nearby points, we can use concentration to estimate the distance from (points in) this cluster to a point that is far away. We first elaborate on this idea and will then discuss how to find clusters.

For any vertex  $v$  and a small parameter  $\xi$ , suppose we are given the vertex subset  $U_v$  such that, for every  $v' \in U_v$ ,

$$|X_v X_{v'}| \leq \xi.$$

Then, for any other vertex  $w$ , we estimate  $|X_v X_w|$  by the edge density from  $U_v$  to  $w$ , as follows.



FIGURE 1. For  $|X_v X_w|$  in the bi-Lipschitz regime of  $p$ , edge density from cluster  $U_v$  to vertex  $w$  reveals  $p(|X_v X_w|)$ .

First, for any  $v' \in U_v$ , since  $|X_v X_{v'}| \leq \xi$ , the Lipschitz continuity of  $p$  implies

$$|p(|X_v X_w|) - p(|X_{v'} X_w|)| \leq L_p \xi.$$

Second, conditional on the latent locations  $X_{U_v} = \{X_{v'}\}_{v' \in U_v}$  and  $X_w$ , the edges between  $U_v$  and  $w$  appear independently with probability  $s_n p(|X_{v'} X_w|)$ . Hence the number of such edges is a sum of independent Bernoulli random variables with mean

$$s_n \sum_{v' \in U_v} p(|X_{v'} X_w|).$$

By a Chernoff bound, with probability at least  $1 - n^{-\omega(1)}$ ,

$$\left| (\# \text{ of edges from } U_v \text{ to } w) - s_n \sum_{v' \in U_v} p(|X_{v'} X_w|) \right| \leq O(\sqrt{s_n |U_v| \log n}).$$

Normalizing by  $s_n |U_v|$  yields the estimator

$$p(|X_v X_w|) = N_{U_v}(w) + O(\xi + \varepsilon_{U_v}),$$

where

$$N_{U_v}(w) := \frac{\# \text{ of edges from } U_v \text{ to } w}{s_n |U_v|} \quad \text{and} \quad \varepsilon_{U_v} := \sqrt{\frac{\log n}{s_n |U_v|}}.$$

If  $|X_v X_w|$  is in the bi-Lipschitz regime of  $p$ , then we can invert  $p$  to estimate  $|X_v X_w|$  up to error  $O(\xi + \varepsilon_{U_v})$ . Thus, the error in estimating  $|X_v X_w|$  has two components: the fluctuation error  $\varepsilon_{U_v}$  from edge randomness, and the radius error  $\xi$  from replacing  $p(|X_{v'} X_w|)$  by  $p(|X_v X_w|)$  for  $v' \in U_v$ .

Balancing the two errors. Suppose  $\xi$  is small relative to  $r_G$  defined in (1.6), and take the “optimal” set

$$U_v := \{v' \neq v : |X_v X_{v'}| \leq \xi\}.$$

If  $\mu$  is the uniform measure on  $M$ , then

$$|U_v| = \frac{\text{Vol}(B_\xi(X_v))}{\text{Vol}(M)} n \gtrsim \xi^d n.$$

For  $X_v$  and  $X_w$  not too far apart (in the Bi-Lipschitz regime of  $p$ ), we have

$$s_n \sum_{v' \in U_v} p(|X_{v'} X_w|) \simeq s_n |U_v|,$$

so the fluctuation scale is  $\sqrt{s_n|U_v|} \simeq \sqrt{s_n\xi^d n}$ . Balancing the two errors gives

$$\xi \simeq (s_n n)^{-\frac{1}{d+2}}.$$

Consequently, this cluster edge density method estimates  $p(|X_v X_w|)$  up to error  $(s_n n)^{-\frac{1}{d+2}}$ , which is the rate achieved in Theorem 1.7. In other words, if for each vertex  $v$  we can find a set  $U_v$  of size at least  $\xi^d n$  within radius  $\xi$ , with the choice of  $\xi$  above, then we can estimate all pairwise distances up to error  $O((s_n n)^{-\frac{1}{d+2}})$ . For vertices  $v, w$  that are far apart, we can rely on the fact that the geodesic distance is a path metric, meaning that we can estimate  $|X_v X_w|$  by summing up the estimated distances along a path of short hops between  $v$  and  $w$ .

**1.3.2. Finding clusters via common-neighbor counts in [HJM24].** For a subset  $W$  of vertices, call it  $(\varepsilon, m)$ -dense if, for every  $i \in W$ , the set

$$W_i := \{i' \in W : |X_i X_{i'}| \leq \varepsilon\}$$

has size at least  $m$ .

Consider the feature map

$$i \in W \mapsto p_i : M \rightarrow [0, 1], \quad p_i(x) := p(|X_i x|),$$

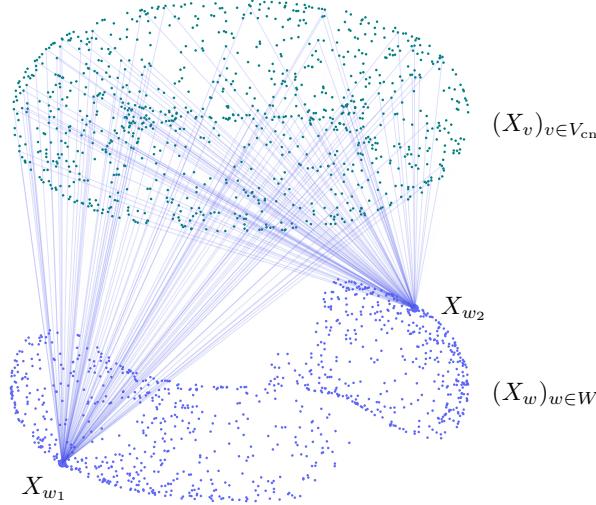
and define the inner product

$$\langle p_i, p_j \rangle_W := \mathbb{E}_{Z \sim \mu}[p_i(Z) p_j(Z)].$$

Let  $V_{cn} \subset V$  be another subset of vertices disjoint from  $W$ . For distinct  $i, j \in W$ , the number of common neighbors of  $i$  and  $j$  in  $V_{cn}$  is a sum of  $|V_{cn}|$  independent Bernoulli random variables with mean  $s_n^2 \langle p_i, p_j \rangle_W$ , conditional on  $X_i, X_j$ . Thus we estimate  $\langle p_i, p_j \rangle_W$  via the normalized common-neighbor count

$$N_{V_{cn}}(i, j) := \frac{\# \text{ of common neighbors between } i \text{ and } j}{s_n^2 |V_{cn}|},$$

except that the self inner product  $\langle p_i, p_i \rangle_W$  is not directly observable.



**Figure 2:** In the diagram, two sets of latent points  $X_W = (x_w)_{w \in W}$  and  $X_{V_{cn}} = (x_v)_{v \in V_{cn}}$  are shown (with  $M$  being a torus). The (normalized) common neighbor count between any two vertices  $w_1, w_2 \in W$  in  $V_{cn}$  estimates the inner product  $\langle p_i, p_j \rangle_W$ . When  $W$  is  $(\varepsilon, m)$ -dense with small  $\varepsilon$ , the closer the common neighbor count  $N_{V_{cn}}(w_1, w_2)$  is to  $\max_{w'_1, w'_2 \in \binom{W}{2}} N_{V_{cn}}(w'_1, w'_2)$ , the smaller the latent distance  $|X_{w_1} X_{w_2}|$  is.

The maximizer of  $\langle p_i, p_j \rangle_W$  over  $i, j \in W$  occurs at  $i = j$ ; this holds even without a manifold structure for  $M$  or regularity of  $\mu$ . To conclude the converse implication

$$\langle p_i, p_j \rangle_W \text{ near its maximum} \Rightarrow |X_i X_j| \text{ is small},$$

we do rely heavily on the manifold structure of  $M$ , essentially that  $M$  is locally Euclidean, together with the lower Ahlfors  $d$ -regularity of  $\mu$ . This idea goes back to [HJM24]; here we adapt it to the

Riemannian manifold setting. Building on it, we construct an algorithm (Algorithm 1) that, given any  $(\varepsilon, m)$ -dense set  $W$  with  $\varepsilon$  sufficiently small, extracts a cluster pair  $(U, u)$  with  $u \in U \subseteq W$  such that

$$(2) \quad |U| \gtrsim m \quad \text{and} \quad |X_u X_{u'}| \leq O(\sqrt{\varepsilon}) \quad \text{for all } u' \in U,$$

by seeking  $(U, u)$  where each pair  $(u, u')$  has a common-neighbor count close to the maximum over all pairs in  $W$ . We remark that there is a loss in (2). When  $\mu$  is uniform on  $M$ , the standard volumetric bound suggests any subset  $W \subseteq \mathbf{V}$  can at most be  $(\varepsilon, \Theta(\varepsilon^d n))$ -dense for  $\varepsilon \gtrsim n^{-1/d}$ . Thus the smallest non-empty cluster we can hope to extract from this approach has radius at least  $\sqrt{n^{-1/d}} = n^{-1/(2d)}$ , which is already suboptimal for the balanced choice of  $\xi$  in the cluster edge density method.

Iterative cluster extraction. The procedure above does not control which vertex is chosen as the cluster center  $u$ . To obtain clusters around different centers, we rerun the algorithm on suitably chosen subsets  $W$  that remain  $(\varepsilon, m)$ -dense.

Initialize with  $W_1 = \mathbf{V}$  to obtain  $(U_1, u_1)$ . Suppose we have  $(U_1, u_1), \dots, (U_k, u_k)$ . Using the cluster edge density method, we can estimate the distances from every vertex to each  $u_i$ ,  $i \in [k]$ . We then filter  $\mathbf{V}$  to a smaller set  $W_{k+1}$  consisting of vertices that satisfy certain approximate distance constraints to the existing centers. If  $W_{k+1}$  is still  $(\varepsilon, m)$ -dense, we run the cluster extractor again to obtain a new cluster  $(U_{k+1}, u_{k+1})$  whose center automatically satisfies the new constraints (since all vertices in  $W_{k+1}$  do). Iterating this yields enough clusters so that every  $v \in \mathbf{V}$  is close to some cluster center  $u$ .

Orthogonal cluster frame. The main technical effort in [HJM24] is to prove that each filtered set  $W_{k+1}$  remains  $(\varepsilon, m)$ -dense, which requires careful control of error accumulation in the distance estimates. An intuitive visualization is that  $W_{k+1}$  typically corresponds to the intersection of several fuzzy annuli centered at the existing cluster centers  $u_1, \dots, u_k$ , where “fuzzy” reflects the uncertainty caused by estimation errors. The “corners” and “boundaries” of this intersection are the regions most prone to losing density, making it challenging to ensure the  $(\varepsilon, m)$ -density property. A key intermediate step is that, for a given cluster  $(U_k, u_k)$ , one also extracts several orthogonal clusters around  $u_k$ , denoted  $(U_{k,1}, u_{k,1}), \dots, (U_{k,d}, u_{k,d})$ , such that the directions from  $X_{u_k}$  to  $X_{u_{k,s}}$  for  $s \in [d]$  are approximately orthogonal in the tangent space  $T_{X_{u_k}} M$ . With these  $d + 1$  clusters, for any other vertex  $v$  whose latent position  $X_v$  is not too far from  $X_{u_k}$ , the distances  $|X_v X_{u_{k,s}}|$  for  $s \in [d]$  together with  $|X_v X_{u_k}|$  — or, from observation of the graph — the edge densities from  $U_{k,s}$  to  $v$  — can be interpreted as defining a local coordinate system around  $X_{u_k}$ . This local coordinate system helps control the geometry of the filtered set  $W_{k+1}$  (all vertices satisfying certain distance constraints to  $u_k$  and  $u_{k,s}$  for  $s \in [d]$ ), preserve its density, and enable the extraction of a new cluster near  $v$  in the next iteration, which is not too far from  $u_k$  in the latent distance.

Limitations in [HJM24]. One of the main limitations of this approach is that the radius of each extracted cluster scales as  $O(\sqrt{\varepsilon})$ , rather than  $O(\varepsilon)$ , as mentioned in (2). From the earlier error-balancing analysis, this enlarged cluster radius leads to a suboptimal choice of  $\xi$  in the cluster edge density method.

Moreover, to preserve probabilistic independence during the iterative filtering and cluster extraction process, the method in [HJM24] partitions the vertex set into  $n^c$  disjoint subsets (for some constant  $c > 0$ ), each of size  $n^{1-c}$ , and runs the cluster extraction algorithm on each subset whenever a new cluster is sought. This partitioning results in a degradation of the final error bound, since each subset contains only  $n^{1-c}$  vertices instead of the full  $n$  is weakening its  $(\varepsilon, m)$ -density property, which leads to a looser final error bound.

Finally, the dominant computational cost arises from computing the number of common neighbors for every pair of vertices in each working subset, which leads to an overall running time of  $O(n^3)$ .

**1.3.3. Our approach.** Our approach builds upon the framework of [HJM24], but stems from a simple yet powerful observation that was not used in [HJM24]. Suppose we have a cluster  $(U, u)$  such that

$$|X_u X_{u'}| \leq \eta, \quad \forall u' \in U, \quad \text{and} \quad |U| \gtrsim \eta^d n,$$

for some radius parameter  $\eta > 0$ . Now consider two vertices  $v, w$  whose latent positions  $X_v, X_w$  are both not too far from  $X_u$ . Then the difference between the expected numbers of edges from  $U$  to  $v$  and from  $U$  to  $w$  is

$$\begin{aligned} & \mathbb{E}[\# \text{ of edges from } U \text{ to } v] - \mathbb{E}[\# \text{ of edges from } U \text{ to } w] \\ &= s_n \sum_{u' \in U} (p(|X_{u'} X_v|) - p(|X_{u'} X_w|)). \end{aligned}$$

By the Lipschitz continuity of  $p$  and the triangle inequality, we have

$$|p(|X_{u'} X_v|) - p(|X_{u'} X_w|)| = O(|X_v X_w|).$$

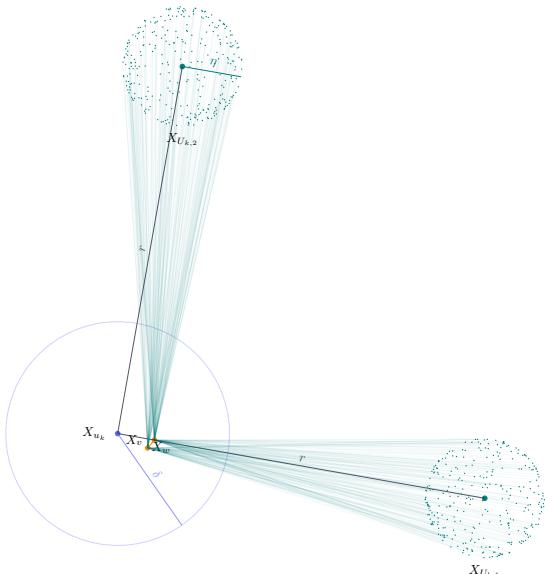
In fact, when  $X_v$  and  $X_w$  are relatively close but much farther from  $X_u$  than  $\eta$ , we can obtain a finer estimate:

$$c_\theta |X_v X_w| \leq (p(|X_{u'} X_v|) - p(|X_{u'} X_w|)) \leq C_\theta |X_v X_w|,$$

where  $\theta$  denotes the angle  $\angle X_u X_v X_w$  in the geodesic triangle formed by  $X_u, X_v, X_w$  centered at  $X_v$ , and  $c_\theta, C_\theta > 0$  are constants depending on  $\theta$  and  $p$ . Consequently,

$$c_\theta |X_v X_w| - O(\varepsilon_U) \leq |N_{U_u} v - N_{U_u} w| \leq C_\theta |X_v X_w| + O(\varepsilon_U).$$

Importantly, unlike the cluster edge density distance estimator, this method introduces no radius error. In summary, while clusters with larger radii may be suboptimal for estimating the absolute distance  $|X_u X_w|$  compared with the balanced-error cluster, they are significantly more effective at detecting changes in distance between two nearby points  $v$  and  $w$ . As an informal analogy, a coarse measurement scale may be unsuitable for measuring absolute lengths precisely, but can be more sensitive in detecting small relative changes.



**Figure 3:** In the diagram, when two vertices  $v, w$  both lie within distance at most  $\delta$  from the cluster center  $X_{u_k}$ —which can be estimated using the cluster edge density method—the difference in their common-neighbor counts  $N_{U_{k,s}}(v)$  and  $N_{U_{k,s}}(w)$ , with respect to the orthogonal cluster  $U_{k,s}$ , provides an estimate of  $|X_v X_w|$  up to a multiplicative error (through the mapping  $\phi$  defined in (3)). This holds whenever  $|X_v X_w| \gg \sqrt{\frac{\log n}{s_n n}}$ , the dimension-free fluctuation scale. In particular,  $|X_v X_w|$  can be negligible relative to the cluster radius  $\eta$ . (Here,  $\eta < \delta < r < r_G$  are some parameters will be of order constants in the paper.)

Multiplicative error distance estimation down to dimension-free fluctuation. Suppose in addition we have found  $d$  clusters  $(U_{u,s}, u_s)$ ,  $s \in [d]$ , around  $u$ , each satisfying the same size and radius conditions as  $(U, u)$ , and such that the directions from  $X_u$  to  $X_{u_s}$  for  $s \in [d]$  are approximately orthogonal and have roughly equal lengths in the tangent space  $T_{X_u} M$ . Then, for any vertices  $v$  and  $w$  whose latent positions  $X_v, X_w$  are both not too far from  $X_u$ , define the vector

$$(3) \quad \phi := \left( p^{-1}(N_{U_{u,s}} v) - p^{-1}(N_{U_{u,s}} w) \right)_{s \in [d]}.$$

This vector serves as an approximate distance estimator satisfying

$$c |X_v X_w| - O(\varepsilon_U) \leq \|\phi\|_2 \leq C |X_v X_w| + O(\varepsilon_U),$$

for some constants  $c, C > 0$ . In this setting, it is advantageous to choose  $\eta$  to be a fixed constant fraction of the Accessible radius  $r_G$  defined in (1.6), which is independent of  $n$ . Then the fluctuation error satisfies

$$\varepsilon_U = O\left(\sqrt{\frac{\log n}{s_n n}}\right),$$

which is dimension-free. In other words, by leveraging differences in the common-neighbor counts across multiple clusters, as long as

$$\sqrt{\frac{\log n}{s_n n}} \ll |X_v X_w| \lesssim r_G,$$

we can estimate  $|X_v X_w|$  up to a multiplicative error. Here, the upper bound ensures that all involved points remain within a locally Euclidean region of  $M$ , and the proper invertibility of  $p$  in this regime. Because the resulting error is multiplicative, this method outperforms the cluster edge density method whenever  $|X_v X_w|$  is small comparing to  $(\log n / (s_n n))^{1/(d+2)}$ .

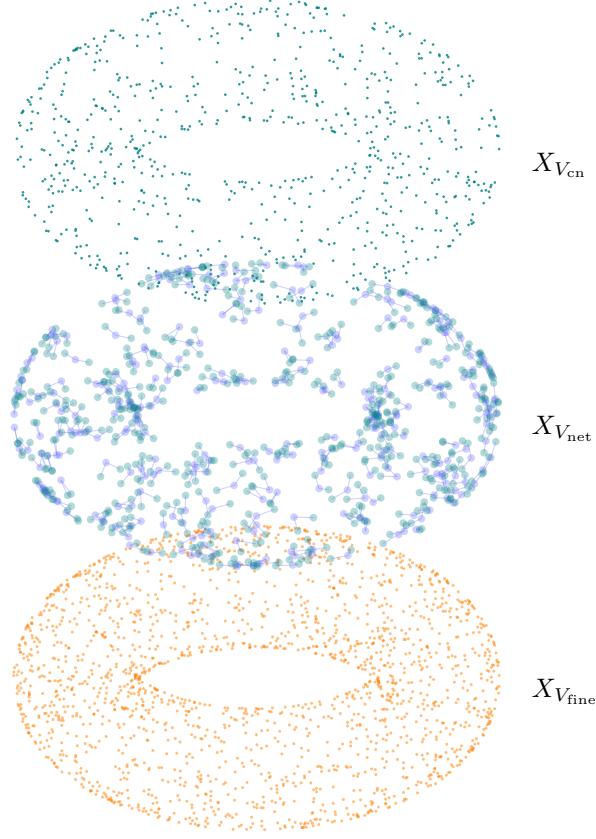
Two-stage approach. Our distance estimation algorithm (Algorithm 5) combines the two ideas above into a unified two-stage framework.

Stage 1: Cluster with constant radius construction. Following the general structure of [HJM24], we iteratively extract clusters  $(U_k, u_k)$  for  $k = 1, 2, \dots$  until they form a  $\delta$ -net of  $M$ , where  $\delta$  is a small constant fraction of  $r_G$ . Around each center  $u_k$ , we further obtain  $d$  orthogonal clusters  $(U_{k,s}, u_{k,s})$  for  $s \in [d]$ . The key differences from [HJM24] are threefold:

- Our cluster radius and  $\delta$  are constants independent of  $n$ , rather than vanishing quantities  $o_n(1)$  as in [HJM24].
- We carefully handle the probabilistic dependence among clusters extracted in different iterations, allowing the entire algorithm to operate using only four disjoint vertex partitions, instead of  $n^c$  partitions (for some constant  $c > 0$ ) as required in [HJM24]. This modification removes one of the two sources of error degradation [HJM24], namely the need to breaking the vertex set into  $n^c$  batches.
- Since we extract only constant-radius clusters, for each vertex pair  $(v, w)$  we need only compute their number of common neighbors over a subset of vertices  $V_{cn}$  with size  $|V_{cn}| = \Theta(s_n^{-2} \log^2(n))$ , which is of order  $\log^2(n)$  when  $s_n = 1$ , comparing with  $|V_{cn}| = n^{1-c}$  in [HJM24]. This yields a substantial computational improvement, reducing the overall runtime from  $O(n^3)$  in [HJM24] to  $O(n^2 \text{polylog}(n))$  in our algorithm.

Stage 2: Local refinement. In the second stage, we use a single new batch of vertices. We may assume that for every vertex  $v$ , there exists a cluster center  $u_k$  such that  $|X_v X_{u_k}| \leq \delta$ . We then employ the multiple-cluster-difference method to extract a refined cluster centered precisely at  $v$ , with radius

$$\eta = O\left(\left(\frac{\log n}{s_n n}\right)^{1/(d+2)}\right) \gg \sqrt{\frac{\log n}{s_n n}}.$$



**Figure 4:** In the diagram, the vertex set  $\mathbf{V}$  is partitioned into three disjoint subsets:  $V_{\text{cn}}$ ,  $V_{\text{net}}$ , and  $V_{\text{fine}}$ . The first two subsets are used in **Stage 1**, where a constant-radius cluster net is constructed over the latent positions of vertices in  $V_{\text{net}}$  based on common-neighbor counts within  $V_{\text{cn}}$ . In **Stage 2**, for each vertex  $v \in V_{\text{fine}}$ , a nearby cluster center  $u_k$  from Stage 1 is identified (via  $N_{U_k}(v)$ ). By leveraging the orthogonal clusters around  $u_k$ , one can estimate the distance between  $X_v$  and any other vertex's latent position  $X_w$  using the multiple-cluster-difference method. This allows us to collect all  $w \in V_{\text{fine}}$  such that  $|X_v X_w| \leq \xi$  to form a cluster around  $v$  with balanced-error radius  $\xi$ . Once all balanced-error clusters have been constructed via edge counting, the distance estimates are obtained through edge counting from fine clusters to vertices.

While smaller clusters can be extracted around each vertex, the key is to select  $\eta$  that balances the radius error and the fluctuation error in the cluster edge density method. Once such clusters are obtained, we can estimate  $|X_v X_w|$  for every pair of vertices  $v, w$  up to an additive error of order  $O((\log^2 n / (\varsigma_n n))^{1/(d+2)})$ , whenever  $|X_v X_w|$  lies within the bi-Lipschitz regime of  $p$ . Finally, we apply a shortest-path computation (e.g., Dijkstra's algorithm) to recover all pairwise distance estimates. With the second stage, we eliminate the other source of error degradation present in [HJM24], which is not capable of extracting optimal error-balanced clusters.

Our Algorithm in the Embedded case. We believe that our algorithms and analysis carry over also to the embedded case analyzed in [HJM24]. Thus we believe that the statement of our main theorem also holds for random graphs sampled according to the Euclidean distance between embedded points, where we can recover these Euclidean distances under the same assumptions and same error bounds as in our main theorem, with one caveat on the Bi-Lipschitz radius of  $p$ . We leave the verification of this to future work.

**1.3.4. Triangle comparison, a bridge between the local structure of  $M$  and the flat Euclidean space.** Although the Nash embedding theorem ensures that any Riemannian manifold can be embedded into a higher-dimensional Euclidean space, our setting differs fundamentally from that of [HJM24]. In their framework, edge probabilities are determined by Euclidean distances in the ambient space, whereas in ours they depend on geodesic distances within the manifold. Consequently, even after embedding, Euclidean and geodesic distances may diverge significantly, which necessitates working directly with the intrinsic Riemannian structure of  $M$ .

Consider a Riemannian manifold  $(M, g)$ . For each point  $p \in M$ , the exponential map at  $p$  is the smooth map

$$\exp_p : T_p M \longrightarrow M,$$

which sends a vector  $v \in T_p M$  to the point reached, at parameter time 1, by the geodesic starting at  $p$  with initial velocity  $v$ . The injectivity radius of  $M$  is defined by

$$(4) \quad r_{\text{inj}}(M) := \sup \left\{ r > 0 : \exp_p|_{B_r(\vec{0})} \text{ is a diffeomorphism for every } p \in M \right\}.$$

For convenience, whenever  $p \in M$  and  $x \in B_{r_{\text{inj}}(M)}(p) \subseteq M$ , we set

$$(5) \quad \hat{x} = (\hat{x})_p := \exp_p^{-1}(x) \in T_p M.$$

A standard consequence of the Gauss lemma (see Proposition A.1 and see e.g. [CE08]) is that, for every  $p \in M$ , the exponential map  $\exp_p$  maps the open ball  $B_{r_{\text{inj}}(M)}(0) \subseteq T_p M$  diffeomorphically onto the open ball  $B_{r_{\text{inj}}(M)}(p) \subseteq M$ . And for each  $x \in B_{r_{\text{inj}}(M)}(p)$ , the curve

$$t \mapsto \exp_p(t \hat{x}), \quad t \in [0, 1],$$

is the unique minimizing geodesic from  $p$  to  $x$ . In particular, the Riemannian (geodesic) distance between  $p$  and  $x$ , denoted by  $|p x|$ , satisfies

$$|p x| = \|\hat{x}\|.$$

Consider two points  $x, y \in B_{r_{\text{inj}}(M)}(p)$ . When  $M$  is flat, the geodesic distance satisfies

$$|x y| = \|\hat{x} - \hat{y}\|.$$

A basic recurring task in the paper is to estimate how much  $M$  around  $p$  deviates from  $T_p M$ , and one of them is to estimate how much the geodesic distance  $|x y|$  can differ from its Euclidean proxy under the sectional curvature bound; let us say that there exists  $\kappa > 0$  such that the sectional curvature  $K$  of  $M$  at any 2-plane  $\sigma$  in  $T_p M$  for every  $p \in M$  satisfies

$$-\kappa \leq K(\sigma) \leq \kappa.$$

(The existence of such a  $\kappa$  is guaranteed when  $M$  is compact.) Rather than expanding the exponential map and tracking higher-order error terms, we use triangle-comparison techniques. Specifically, we compare the geodesic triangle with its counterpart in the model space  $M_\kappa$  of constant curvature  $\kappa$ . The Rauch Comparison Theorem and Toponogov's Triangle Comparison Theorem (see e.g. [CE08]) allow us to translate the curvature bound into precise inequalities for side lengths and angles.

**Definition 1.9.** For any  $\kappa \in \mathbb{R}$ , let  $M_\kappa^d$  be the  $d$ -dimensional, simply-connected, complete Riemannian manifold whose sectional curvature is identically  $\kappa$ . The three standard cases are summarized below:

Curvature $\kappa$	Name of model space $M_\kappa^d$	Concrete realization	Diameter $\varpi_\kappa$
$\kappa > 0$	$d$ -sphere of radius $1/\sqrt{\kappa}$	$\frac{1}{\sqrt{\kappa}} \mathbb{S}^d \subseteq \mathbb{R}^{d+1}$	$\frac{\pi}{\sqrt{\kappa}}$
$\kappa = 0$	Euclidean $d$ -space	$\mathbb{R}^d$	$\infty$
$\kappa < 0$	Hyperbolic $d$ -space of curvature $\kappa$	$\mathbb{H}_\kappa^d$	$\infty$

In what follows, when the dimension  $d$  is clear from the context, we may write  $M_\kappa$  instead of  $M_\kappa^d$ , omitting the superscript  $d$ .

**Definition 1.10.** Let  $\kappa \in \mathbb{R}$ ,  $\theta \in [0, \pi]$ , and  $a, b > 0$  with  $a + b < \varpi_\kappa$ . We write

$$\gamma^\kappa(\theta; a, b)$$

for the length of the side opposite the angle  $\theta$  in a geodesic triangle whose two adjacent sides forming the angle  $\theta$  have lengths  $a$  and  $b$  in the model space  $M_\kappa^d$ .

We remark that the assumption  $|px| + |py| < \varpi_\kappa$  guarantees that the corresponding geodesic triangle is well-defined when  $\kappa > 0$  in the positively curved model space  $M_\kappa^d$ . Now suppose that  $\kappa > 0$ , and suppose that three points  $p, x, y \in M$  satisfy

$$|px| + |py| < \varpi_\kappa.$$

Let  $\theta$  denote the angle between  $\hat{x}$  and  $\hat{y}$  in  $T_p M$ . Then by the Rauch Comparison Theorem (see e.g. [CE08]), the length of the side opposite to  $\theta$  obeys

$$(6) \quad \gamma^\kappa(\theta; |px|, |py|) \leq |xy| \leq \gamma^{-\kappa}(\theta; |px|, |py|).$$

Because the opposite side length in each model space can be written explicitly by the spherical and hyperbolic laws of cosines, each bound in (6) may be compared with the flat case

$$\gamma^0(\theta; |px|, |py|) = \|\hat{x} - \hat{y}\|,$$

which yields estimates of the form

$$|xy| = \|\hat{x} - \hat{y}\| + \text{higher-order terms in } \max\{|px|, |py|\}.$$

In many arguments we must also compare the angle at  $p$  with its analogue in a model space. To avoid repeatedly referring to the angle between the tangent vectors  $\hat{x}, \hat{y} \in T_p M$ , we adopt a notation similar to that of [AKP24].

**Definition 1.11.** Let  $p, x, y \in M$  with  $x, y \in B_{r_{\text{inj}}(M)/2}(p) \setminus \{p\}$ . We define

$$\angle(p_y^x) := \angle(\hat{x}, \hat{y}) = \arccos\left(\frac{\langle \hat{x}, \hat{y} \rangle}{\|\hat{x}\| \|\hat{y}\|}\right),$$

the angle at  $p$  of the geodesic triangle  $\triangle pxy$  in  $M$ . Given  $\kappa > 0$  and assuming  $|px| + |py| < \varpi_\kappa$ , we further set

$$\angle^\kappa(p_y^x) \quad \text{and} \quad \angle^{-\kappa}(p_y^x)$$

to be, respectively, the angles at  $\tilde{p}$  of the comparison triangle  $\triangle \tilde{p}\tilde{x}\tilde{y} \subseteq M_\kappa$  and of the corresponding triangle in  $M_{-\kappa}$ —that is, the unique triangle in the model space whose side lengths equal  $|px|$ ,  $|py|$ , and  $|xy|$ .

The requirement  $x, y \in B_{r_{\text{inj}}(M)/2}(p)$  guarantees that  $|px|$ ,  $|py|$ , and  $|xy|$  are all less than  $r_{\text{inj}}(M)$ , so the distance-minimizing geodesics are unique; hence there is no ambiguity in the definition of the corresponding geodesic triangles in  $M$ . As the dual inequality to (6), the angles obey

$$(7) \quad \angle^{-\kappa}(p_y^x) \leq \angle(p_y^x) \leq \angle^\kappa(p_y^x).$$

Moreover, because the spherical and hyperbolic laws of cosines express  $\angle^{\pm\kappa}(p_y^x)$  explicitly in terms of the side lengths, one can compare them with the flat angle

$$\angle^0(p_y^x),$$

i.e., the ordinary Euclidean angle of the triangle with the same side lengths.

Taken together, the comparison bounds (6) and (7) form a quantitative bridge between the local geometry of  $M$  and that of its tangent plane of interest. They let us compare the two spaces on a genuinely local—rather than merely infinitesimal—scale. Equipped with these estimates we may treat  $M$  as almost flat in a sufficiently small neighborhood of  $p$ , which in turn allows us to use many arguments from linear algebra.

Outline of the paper. The section structure of the paper is as follows.

- Section 2 introduces some basic notions and events that will be used throughout the paper.
- Section 3 presents several geometric estimates on Riemannian manifolds that will be used in the analysis.
- Section 4 we state and prove the results on cluster extraction through common-neighbor counts, which form the basis of our distance estimation algorithm.

- Section 5 presents an intermediate algorithms for constructing constant-radius cluster nets and for local refinement, respectively.
- Section 6 presents the multi-clusters-difference distance estimator and analyzes its performance.
- Section 7-8 illustrate our two stages of distance estimation algorithm.
- Section 9 concludes the proof of our main theorem by combining the results from previous sections.
- Appendix A collects several standard results on Riemannian geometry used in the paper.
- Appendix B shows the proof of the lower bound result (Theorem 1.8).

#### ACKNOWLEDGMENTS

The authors were partially supported by Vannevar Bush Faculty Fellowship ONR-N00014-20-1-2826 and by Investigator award (622132). E.M. was also partially supported by ARO MURI W911NF1910217, NSF DMS-2031883, and NSF award CCF 1918421.

We would like to thank Tristan Collins, Tirasan Khandhawit, and Greg Parker for helpful discussions.

## 2. BASIC NOTIONS AND EVENTS

Without loss of generality, throughout the paper we assume  $\mathbf{V}$  is the vertex set of the random geometric graph  $G \sim G(n, M, \mu, p, s_n)$ . For simplicity, we assume  $|\mathbf{V}|$  is a constant multiple of  $n$ , instead of  $n$ , so that we can partition  $\mathbf{V}$  into several disjoint subsets of equal size when needed. In particular, we assume  $|\mathbf{V}| = 4n$  throughout the paper, which does not affect the statement of our main theorem (Theorem 1.7). Further, throughout the paper, we assume that the sparsity parameter  $s_n$  satisfies

$$s_n^2 n = \omega(\log^2 n).$$

Below we define some basic events that will be used in the proofs of our main results and show that they hold with high probability.

**Definition 2.1** (Normalized common neighbors). *Let  $U, V \subseteq \mathbf{V}$ . We define the normalized number of common neighbors of  $V$  in  $U$  by*

$$(8) \quad N_U(V) := \frac{|\{u \in U : v \sim u \text{ for all } v \in V\}|}{(s_n)^{|V|} |U|}.$$

For a singleton  $V = \{v\}$  we write  $N_U(v)$ , and for a pair  $V = \{v, w\}$  we write  $N_U(v, w)$ . Further, for a singleton  $v$ , we also define the (conditional) expectation

$$p_U(v) := \frac{1}{|U|} \sum_{u \in U} p(|X_v X_u|),$$

so that  $\mathbb{E}[N_U(v) | X_v, X_U] = p_U(v)$ . Accordingly, we also define a notion of fluctuation error  $\varepsilon_U$  that depends on  $|U|$ :

$$(9) \quad \varepsilon_U := \frac{\log n}{\sqrt{s_n |U|}}, \quad \text{and} \quad \varepsilon_m := \frac{\log n}{\sqrt{s_n m}}. \quad (m \geq 1).$$

**Lemma 2.2.** (Edge fluctuation event) *Let  $U, V \subseteq \mathbf{V}$  be disjoint with  $|V| \leq n$ . Define*

$$(10) \quad \mathcal{E}_{\text{navi}}(U, V) := \left\{ \forall v \in V : |N_U(v) - p_U(v)| \leq \varepsilon_U \right\}.$$

Then for all fixed realizations  $X_V = x_V$ ,  $X_U = x_U$ ,

$$\Pr(\mathcal{E}_{\text{navi}}(U, V)^c | X_V = x_V, X_U = x_U) \leq n^{-\omega(1)}.$$

The above bound follows simply from the union bound and Chernoff bound:

**Lemma 2.3** (Chernoff Bound). *Let  $Y = \sum_{i=1}^n Y_i$  be the sum of independent Bernoulli random variables with  $\mu = \mathbb{E}[Y]$ . Then for any  $0 < t < \mu$ ,*

$$\Pr[|Y - \mu| \geq t] \leq 2 \exp\left(-\frac{t^2}{3\mu}\right).$$

*Proof.* Fix  $x_V, x_U$ . For each  $v \in V$ ,

$$N_U(v) = \frac{1}{s_n|U|} \sum_{u \in U} \mathbf{1}\{\mathcal{U}_{v,u} < s_n p(|X_v X_u|)\},$$

is a sum of independent Bernoulli random variables conditioned on  $X_v, X_U$ , normalized by  $\Delta := s_n|U|$  and with mean  $p_U(v)$ . By Chernoff bound, we have

$$\begin{aligned} \Pr(|N_U(v) - p_U(v)| \geq \varepsilon_U | X_v, X_U) &\leq \Pr(|N_U(v) - p_U(v)| \geq \varepsilon_U p_U(v) | X_v, X_U) \\ &\leq 2 \exp\left(-\frac{\varepsilon_U^2 \Delta^2}{3\Delta p_U(v)}\right) = 2 \exp(-\log^2 n/3), \end{aligned}$$

where we used the fact that  $p_U(v) \leq 1$ . Next, applying the union bound over  $v \in V$  yields the claim.  $\square$

**Definition 2.4.** (*Cluster event*) For  $u \in U \subseteq \mathbf{V}$  and  $\eta > 0$ , we define the event

$$(11) \quad \mathcal{E}_{\text{clu}}(\eta, U, u) := \left\{ \forall u' \in U, |X_{u'} X_u| \leq \eta \right\}.$$

So, if  $\mathcal{E}_{\text{clu}}(\eta, U, u)$  and  $\mathcal{E}_{\text{navi}}(U, v)$  hold, then the normalized number of common neighbors  $N_U(v)$  can be used to estimate the latent distance  $|X_v X_u|$ , with error coming from both the cluster radius  $\eta$  and the fluctuation error  $\varepsilon_U$ .

**Definition 2.5.** (*Size Calibration margin*) For a set  $U \subseteq \mathbf{V}$ , we say that the size of  $U$  satisfies the calibration margin with parameter  $\eta > 0$  if

$$(12) \quad \varepsilon_U \leq \ell_p \eta.$$

**Lemma 2.6** (Calibration toolkit: single, annulus). *Let  $\rho \in U \subseteq \mathbf{V}$ . Suppose the size of  $U$  satisfies the calibration margin (12) with parameter  $0 < \eta < r_G$ , and the events  $\mathcal{E}_{\text{clu}}(\eta, U, \rho)$  and  $\mathcal{E}_{\text{navi}}(U, w)$  hold. Then:*

(i) *Single-cluster calibration.*

$$(13) \quad |N_U(w) - p(|X_w X_\rho|)| \leq L_p \eta + \varepsilon_U \leq 2L_p \eta.$$

If moreover either  $N_U(w) \geq p(r_G)$  or  $|X_w X_\rho| \leq r_G$ , then

$$||X_w X_\rho| - p^{-1}(N_U(w))| \leq \eta + \ell_p^{-1} \varepsilon_U \leq 2\eta.$$

If  $N_U(w) < p(r_G)$ , then  $|X_w X_\rho| > r_G - 2\eta$ .

(ii) *Annulus-threshold equivalence.* Fix  $r \leq r_G$  and  $\alpha \in [0, r]$ .

$$\begin{aligned} (14) \quad N_U(w) \leq p(r - \alpha) &\Rightarrow |X_w X_\rho| \geq r - \alpha - 2\eta \\ N_U(w) \geq p(r + \alpha) &\Rightarrow |X_w X_\rho| \leq r + \alpha + 2\eta \\ ||X_w X_\rho| - r| \leq \alpha &\Rightarrow p(r + \alpha + 2\eta) \leq N_U(w) \leq p(r - \alpha - 2\eta). \end{aligned}$$

*Proof. Preparation.* Throughout we use that  $p$  is nonincreasing,  $L_p$ -Lipschitz on  $[0, \text{diam}(M)]$ , and bi-Lipschitz on  $[0, 16r_G]$  with lower constant  $\ell_p$ . If  $\mathcal{E}_{\text{clu}}(\eta, U, \rho)$  holds, then for any  $w$ ,

$$(15) \quad p(|X_w X_\rho| + \eta) \leq p_U(w) \leq p(|X_w X_\rho| - \eta).$$

This is immediate from the triangle inequality and the monotonicity of  $p$ . If  $\mathcal{E}_{\text{navi}}(U, w)$  holds, then by definition

$$(16) \quad |N_U(w) - p_U(w)| \leq \varepsilon_U.$$

Single-cluster calibration. Combining (15) and (16) and using the  $L_p$ -Lipschitz property, let  $D = \frac{N_U(w)}{|X_w X_\rho|}$ . Then

$$|N_U(w) - p(|X_w X_\rho|)| \leq \varepsilon_U + \max \left\{ |p(D + \eta) - p(D)|, |p(D - \eta) - p(D)| \right\} \leq \varepsilon_U + L_p \eta.$$

This proves the first inequality in (13). Since  $\varepsilon_U \leq \ell_p \eta \leq L_p \eta$ , we also get  $|N_U(w) - p(D)| \leq 2L_p \eta$ . Now consider the inversion bound. If  $N_U(w) \geq p(r_G)$ , then

$$p(D - \eta) \geq p_U(w) \geq N_U(w) - \varepsilon_U \geq p(r_G) - \varepsilon_U \Rightarrow D \leq r_G + \eta + \ell_p^{-1} \varepsilon_U < 16r_G.$$

On the other hand, if  $D \leq r_G$ , then

$$N_U(w) \geq p_U(w) - \varepsilon_U \geq p(D + \eta) - \varepsilon_U \geq p(r_G + \eta) - \varepsilon_U \Rightarrow p(N_U(w)) \leq r_G + \eta + \ell_p^{-1} \varepsilon_U < 16r_G.$$

In either case, we have  $D, N_U(w) \in [0, 16r_G]$ , so the low-Lipschitz property applies:

$$|D - p^{-1}(N_U(w))| \leq \ell_p^{-1} |p(D) - N_U(w)| \leq \ell_p^{-1} \varepsilon_U + \ell_p^{-1} L_p \eta \leq \ell_p^{-1} \varepsilon_U + \eta \leq 2\eta.$$

Finally, if  $N_U(w) < p(r_G)$ , a similar argument shows  $p(D + \eta) \leq p(r_G) + \varepsilon_U$  implying  $D > r_G - \eta - \ell_p^{-1} \varepsilon_U > r_G - 2\eta$ .

Annulus-threshold equivalence. Fix  $r \leq r_G$  and  $\alpha \in [0, r]$ . Assume  $N_U(w) \leq p(r - \alpha)$ . Then,

$$p(D + \eta) \leq p_U(w) \leq N_U(w) + \varepsilon_U \leq p(r - \alpha) + \varepsilon_U \leq p(r - \alpha - \ell_p^{-1} \varepsilon_U) \Rightarrow D \geq r - \alpha - 2\eta.$$

Similarly, if  $N_U(w) \geq p(r + \alpha)$ , then invoking the inequalities in the other direction leads to

$$D \leq r + \alpha + 2\eta.$$

This proves the first implication in (14). Now we assume  $|D - r| \leq \alpha$ . Then,

$$N_U(w) \leq p_U(w) + \varepsilon_U \leq p(D - \eta) + \varepsilon_U \leq p(r - \alpha - \eta) + \varepsilon_U \leq p(r - \alpha - 2\eta),$$

and similarly in the other direction. From (15),

$$p(D + \eta) \leq \frac{1}{|U|} \sum_{u \in U} p(D X_w X_u) \leq p(D - \eta).$$

Adding/subtracting  $\varepsilon_U$  via (16) and using monotonicity,

$$p(r + \alpha + \eta) - \varepsilon_U \leq N_U(w) \leq p(r - \alpha - \eta) + \varepsilon_U.$$

This matches the claimed bounds.  $\square$

**Definition 2.7** (Common neighbor as inner product). *For points  $x, y \in M$ , we define the following inner product-like quantity that measures the common neighbor probability between  $x$  and  $y$*

$$(17) \quad \langle x, y \rangle := \mathbb{E}_{Z \sim \mu} p(|x Z|) p(|y Z|) = \int_M p(|x z|) p(|y z|) d\mu(z),$$

where in particular, it is the inner product of the functions  $p(|x \cdot|)$  and  $p(|y \cdot|)$  in the Hilbert space  $L^2(M, \mu)$ . Moreover,

$$\langle X_i, X_j \rangle$$

is precisely the probability that a third vertex  $k$  is in the common neighbor of  $i$  and  $j$  (Definition 1.5) conditioned on the latent positions  $X_i$  and  $X_j$ .

**Lemma 2.8.** *Suppose  $V, U$  are two disjoint set of vertices in our random geometric graph model (Definition 1.5) with  $|V| \leq n$ . Consider the event*

$$(18) \quad \mathcal{E}_{cn}(U, V) := \left\{ \forall v, w \in V, \left| N_U(v, w) - \langle X_v, X_w \rangle_{cn} \right| \leq \frac{\log(n)}{\sqrt{s_n^2 |U|}} \right\},$$

which is an event of  $\mathcal{U}_{U,V}$  and the latent points  $X_V, X_U$ . If

$$s_n^2 |U| = \omega(\log(n)),$$

then

$$\Pr \left( \mathcal{E}_{\text{cn}}(U, V)^c \mid X_V = x_V \right) = n^{-\omega(1)} \quad \forall x_V \in M^{|V|}.$$

*Proof.* The proof is similar to that of Lemma 2.2. It is an application of the Chernoff bound together with the union bound over  $\binom{|V|}{2}$  pairs of vertices in  $V$ . Let us fix any realization of the latent points  $X_V = x_V$ . Fix  $\{v, w\} \in \binom{|V|}{2}$ . Observe

$$N_U(v, w) = \frac{1}{s_n^2 |U|} \sum_{u \in U} \underbrace{\mathbf{1}(\mathcal{U}_{v,u} < s_n p(|X_v X_u|)) \mathbf{1}(\mathcal{U}_{v,u} < s_n p(|X_v X_u|))}_{:= Z_u}.$$

Notice that  $\{Z_u\}_{u \in U}$  are independent Bernoulli random variables with mean

$$\mathbb{E}_{X_u, \mathcal{U}_{u, \{v,w\}}} [Z_u \mid X_v, X_w] = \mathbb{E}_{X_u} \left[ s_n^2 p(|X_v X_u|) p(|X_w X_u|) \mid X_v, X_w \right] = s_n^2 \langle X_v, X_w \rangle_{\text{cn}},$$

which implies that

$$\mathbb{E}_{X_u, \mathcal{U}_{u, \{v,w\}}} [N_U(v, w) \mid X_v, X_w] = \langle X_v, X_w \rangle_{\text{cn}}.$$

The standard Chernoff bound gives

$$\begin{aligned} & \Pr \left( \left| N_U(v, w) - \langle X_v, X_w \rangle_{\text{cn}} \right| \geq \frac{\log(n)}{\sqrt{s_n^2 |U|}} \mid X_V = x_V \right) \\ &= \Pr \left( \left| \sum_{u \in U} Z_u - |U| s_n^2 \langle X_v, X_w \rangle_{\text{cn}} \right| \geq \log(n) \sqrt{s_n^2 |U|} \mid X_V = x_V \right) \leq 2 \exp(-c \log^2(n)) = n^{-\omega(1)}, \end{aligned}$$

where we used the fact that  $\langle X_v, X_w \rangle_{\text{cn}} \leq 1$ . Within the complement of the event described above, we have, after normalization, the event  $\mathcal{E}_{\text{cn}}(U, V)$  holds.  $\square$

**Lemma 2.9.** Suppose  $\{X_v\}_{v \in V}$  are i.i.d. random points sampled from a distribution  $\mu$  on a manifold  $M$  and  $|V| = n$ . The event

$$(19) \quad \mathcal{E}_{\text{pt}}(V) := \left\{ \left| \{X_i\}_{i \in [n]} \cap B_\varepsilon(p) \right| \geq \mu_{\min}(\varepsilon/3) \frac{n}{2} \quad \forall p \in M, \left( \frac{\log(n)}{n} \right)^{1/d} \leq \varepsilon \leq r_\mu \right\},$$

holds with probability  $1 - n^{-\omega(1)}$ .

The above lemma follows from standard volumetric arguments and Chernoff bound; see e.g. [HJM24] for a similar proof.

## 2.1. Additional Notation.

**Assumption 2.10** (Parameter Configurations). *In the algorithm, we will choose several distance parameters:  $\eta, \delta, r$  corresponding to cluster radius, net radius, and orthonormal frame radius respectively. Further, we also have a cluster size parameter  $m$  (a desired lower bound on the size of an extracted cluster). For any gap parameter  $c \in [0, 1]$  which might depends on  $\ell_p, L_p, \mu_{\min}(\cdot)$ , and  $d$ , but not on  $n$ , we say the tuples  $(r, \delta, \eta, m)$  satisfy the parameter configurations with gap  $c$  if*

$$(20) \quad r \leq c \cdot r_G, \quad \delta = c \cdot r, \quad \eta = c \cdot \delta \quad \text{and} \quad \varepsilon_m \leq c \eta^2.$$

## 3. TRIANGLE COMPARISON

Here we will collect the main tools from Riemannian geometry that allow us to compare distances and angles in a Riemannian manifold with those in model spaces of constant curvature. The main result is the following Triangle Comparison Lemma, which is a direct consequence of the Rauch Comparison Theorem (see e.g. [CE08]). As most of the statements in this section are standard in Riemannian geometry, we will postpone them to the appendix. Further, in the end of this section,

we also include a regularity lemma (Lemma 3.6) for measuring the change of edges count from a cluster to a single vertex when perturbing the latent position of the vertex slightly.

**Lemma 3.1.** *Let  $\kappa > 0$  be a positive constant. Let  $(M, g)$  be a compact  $d$ -dimensional Riemannian manifold whose sectional curvatures satisfy*

$$-\kappa \leq K(\sigma) \leq \kappa,$$

for every 2-plane  $\sigma$  in the tangent bundle  $TM$ . Fix  $p \in M$  and set

$$r := \min\{r_{\text{inj}}(M), \pi/\sqrt{\kappa}\}.$$

Consider  $x, y \in B_{r/2}(p)$  and write  $\theta := \angle(p_y^x)$ . Then

$$(21) \quad \gamma^\kappa(\theta; |px|, |py|) \leq |xy| \leq \gamma^{-\kappa}(\theta; |px|, |py|)$$

and

$$(22) \quad \angle^{-\kappa}(p_y^x) \leq \angle(p_y^x) \leq \angle^\kappa(p_y^x).$$

It is also worth recalling the laws of cosines for model spaces of constant curvature  $\kappa$ , which relate the side lengths and an angle of a geodesic triangle.

**Fact 3.2** (The laws of cosines for model spaces, see e.g. [AKP24]). *Let  $\kappa > 0$  and consider a geodesic triangle in the spherical model space  $M_\kappa$  with side lengths  $a, b, c$  satisfying*

$$a, b, c < \frac{1}{2}\varpi_\kappa = \frac{\pi}{2\sqrt{\kappa}}.$$

If  $\theta$  denotes the angle opposite the side of length  $c$ , then

$$(23) \quad \cos(\sqrt{\kappa}c) = \cos(\sqrt{\kappa}a)\cos(\sqrt{\kappa}b) + \sin(\sqrt{\kappa}a)\sin(\sqrt{\kappa}b)\cos(\theta).$$

For a geodesic triangle in the hyperbolic model space  $M_{-\kappa}$  ( $\kappa > 0$ ) with side lengths  $a, b, c$  and  $\theta$  again the angle opposite the side of length  $c$ , one has

$$(24) \quad \cosh(\sqrt{\kappa}c) = \cosh(\sqrt{\kappa}a)\cosh(\sqrt{\kappa}b) - \sinh(\sqrt{\kappa}a)\sinh(\sqrt{\kappa}b)\cos(\theta).$$

Following from the Triangle Comparison Lemma 3.1 and the laws of cosines above, we have a few lemmas regarding angles and distances which essentially arrive from expanding the laws of cosines in Taylor series.

**Lemma 3.3.** *Let  $(M, g)$  be a Riemannian manifold whose sectional curvatures are uniformly bounded in absolute value by a constant  $\kappa > 0$ . Suppose  $p, q, x$  be points in  $M$  such that*

$$0 \leq |px|, |pq| < \frac{1}{4} \min\{\varpi_\kappa, r_{\text{inj}}(M)\} \quad \text{and} \quad |pq| \leq \frac{1}{4}|px|.$$

Expressing

$$a := |px|, \quad b := |pq|, \quad c := |qx|, \quad \text{and} \quad \theta := \angle(p_q^x).$$

Then

$$-\frac{7}{6}\frac{b^2}{a} - \frac{\kappa b^2}{3} \cdot b|\cos(\theta)| \leq a - c - b\cos(\theta) \leq \pi\frac{b}{a} \cdot b|\cos(\theta)|.$$

**Remark 3.4.** In the flat Euclidean plane, the ordinary law of cosines gives

$$c^2 = a^2 + b^2 - 2ab\cos\theta \implies c = a - b\cos\theta + O(b^2) \quad (b \rightarrow 0, a \text{ fixed}).$$

Hence  $a - c - b\cos\theta$  is the second-order remainder in the Taylor expansion of  $c$  with respect to  $b$  at  $b = 0$ . Our goal is to show that, with the curvature correction, how this second-order remainder is controlled.

**Lemma 3.5.** *Fix  $\kappa > 0$ . Consider a geodesic triangle in  $M_\kappa$  with adjacent sides  $a, b$  and angle  $\theta$  between them. Assume*

$$0 < a, b < \frac{1}{4}\varpi_\kappa \quad \text{and} \quad c := \gamma^\kappa(\theta; a, b) \leq \frac{1}{2}a.$$

Then

$$\theta \leq 2 \frac{c}{a}.$$

Given these lemmas are standard in Riemannian geometry and not to detract from the main flow of the paper, we postpone their proofs to Appendix A.

**Lemma 3.6** (Edge Counting Regularity). *Let  $(M, g)$  be a Riemannian manifold whose sectional curvatures are uniformly bounded in absolute value by a constant  $\kappa > 0$ . Let  $p, x, q \in M$  be three points in a Riemannian manifold  $M$  with*

$$|px|, |pq| < \frac{1}{8} \min\{r_{\text{inj}}(M), \varpi_\kappa\} \quad \text{and} \quad 16 \frac{|pq|}{|px|} \leq |\cos(\angle(p_x^q))|.$$

Suppose  $x' \in M$  is a point satisfying

$$16 \frac{|xx'|}{|px|} \leq |\cos(\angle(p_x^q))|.$$

Then, the following holds:

$$(25) \quad \frac{1}{2} \leq \frac{|px'| - |qx'|}{\cos(\angle(p_x^q))|pq|} \leq 2.$$

Furthermore, with the additional assumption that

$$|px| \leq \frac{1}{2} r_p,$$

we have

$$\frac{1}{2} \ell_p \leq \frac{p(|qx'|) - p(|px'|)}{\cos(\angle(p_x^q))|pq|} \leq 2L_p.$$

*Proof.* Cosine value comparison. We first invoke Lemma 3.1 to compare the angle  $\angle(p_{x'}^x)$  with the model angle  $\angle^\kappa(p_{x'}^x)$ , and then Lemma 3.5 to obtain the following estimate on the angle  $\angle(p_{x'}^x)$ :

$$\angle(p_{x'}^x) \leq \angle^\kappa(p_{x'}^x) \leq 2 \frac{|xx'|}{|px|},$$

which in turn implies

$$|\angle(p_x^q) - \angle(p_{x'}^q)| \leq \angle(p_{x'}^x) \leq 2 \frac{|xx'|}{|px|}.$$

Relying on the fact that  $t \mapsto \cos(t)$  is a Lipschitz function with constant 1,

$$|\cos(\angle(p_x^q)) - \cos(\angle(p_{x'}^q))| \leq 2 \frac{|xx'|}{|px|} \leq \frac{1}{8} |\cos(\angle(p_x^q))|,$$

where the last inequality follows from the assumption  $16 \frac{|xx'|}{|px|} \leq |\cos(\angle(p_x^q))|$ . Therefore, we conclude that  $\cos(\angle(p_x^q))$  and  $\cos(\angle(p_{x'}^q))$  have the same sign and

$$(26) \quad \frac{7}{8} |\cos(\angle(p_x^q))| \leq |\cos(\angle(p_{x'}^q))| \leq \frac{9}{8} |\cos(\angle(p_x^q))|.$$

Invoking Lemma 3.3. We apply Lemma 3.3 to the triple  $(p, q, x')$ . To do that, we still need to show

$$(27) \quad |px'| < \frac{1}{4} \min\{r_{\text{inj}}(M), \varpi_\kappa\} \quad \text{and} \quad |pq| < \frac{1}{4} |px'|.$$

To drive this, we first note that the  $16 \frac{|xx'|}{|px|} \leq |\cos(\angle(p_x^q))| \leq 1$  implies

$$(28) \quad \frac{15}{16} |px| \leq |px'| \leq \frac{17}{16} |px|.$$

Combining this with the assumption  $|px| < \frac{1}{8} \min\{r_{\text{inj}}(M), \varpi_\kappa\}$  and  $|pq| \leq \frac{1}{8}|px|$  leads to (27) and a slightly stronger estimate

$$(29) \quad |pq| \leq \frac{2}{15}|px'|.$$

Now expressing

$$a := |px'|, \quad b := |pq|, \quad c := |qx'|, \quad \text{and} \quad \theta := \angle(p_q^{x'})$$

and invoking Lemma 3.3, we have

$$(30) \quad -\frac{7}{6}\frac{b^2}{a} - \frac{\kappa b^2}{3} \cdot b|\cos(\theta)| \leq a - c - b\cos(\theta) \leq \pi\frac{b}{a} \cdot b|\cos(\theta)|.$$

Simplifying the estimates. We first claim that

$$\pi\frac{b}{a} \cdot b|\cos(\theta)| \leq \frac{b}{2}|\cos(\theta)| \quad \text{and} \quad \frac{7}{6}\frac{b^2}{a} + \frac{\kappa b^2}{3} \cdot b|\cos(\theta)| \leq \frac{b}{2}|\cos(\theta)|.$$

The first inequality is a straightforward consequence of (29) and the fact that  $\frac{2}{15}\pi \leq 0.42$ . By (27),

$$\frac{\kappa b^2}{3} = \frac{\kappa a^2}{3}\frac{b^2}{a^2} \leq \frac{1}{16 \cdot 3}\frac{1}{16}\frac{b}{a}.$$

Also, our assumption  $16\frac{|xx'|}{|px|} \leq |\cos(\angle(p_x^q))|$ , (28), and (26) imply

$$\frac{b}{a} = \frac{|pq|}{|px'|} = \frac{|pq|}{|px|} \cdot \frac{|px|}{|px'|} \leq \frac{1}{16}|\cos(\angle(p_x^q))| \cdot \frac{16}{15} \leq \frac{1}{16}\frac{9}{8}|\cos(\theta)| \cdot \frac{16}{15} = \frac{3}{40}|\cos(\theta)|.$$

Hence,

$$\frac{7}{6}\frac{b^2}{a} \leq \frac{7}{80}b|\cos(\theta)|.$$

Combining these two estimates for the summands leads to the claim. Finally, the claim together with (30) implies

$$1 - 0.42 \leq \frac{a - c}{b\cos(\theta)} \leq 1 + 0.42,$$

regardless of the sign of  $\cos(\theta)$ . Unwrapping the definitions of  $a, b, c, \theta$ , together with (26), we obtain the first statement of the lemma.

Final estimate on p. Notice that

$$\max\{|px'|, |qx'|\} \leq \max\{|px| + |xx'|, |qp| + |px| + |xx'|\} \leq \left(\frac{1}{16} + 1 + \frac{1}{16}\right)|px| \leq r_p,$$

where the last inequality follows from the additional assumption  $|px| \leq \frac{1}{8}r_p$ . Within  $[0, r_p]$ , we can rely on the bi-Lipschitz continuity of  $p$  to obtain

$$\ell_p \leq \frac{p(|qx'|) - p(|px'|)}{|px'| - |qx'|} \leq L_p,$$

where we emphasize that due to  $p$  is monotone decreasing,  $p(|qx'|) - p(|px'|)$  and  $|px'| - |qx'|$  always have the same sign. Substituting this ratio into the first statement of the lemma, we obtain second statement of the lemma.  $\square$

#### 4. COMMON NEIGHBOR PROBABILITY AND CLUSTER FINDING ALGORITHM

Using the algebraic identity  $(a - b)^2 = a^2 - 2ab + b^2$ , we obtain the following proposition.

**Proposition 4.1.** *For any points  $x, y \in M$ , we have*

$$(31) \quad \langle x, y \rangle_{\text{cn}} = \frac{\langle x, x \rangle_{\text{cn}} + \langle y, y \rangle_{\text{cn}}}{2} - \frac{1}{2} \int_M (\mathbf{p}(|xz|) - \mathbf{p}(|yz|))^2 d\mu(z).$$

The following lemma derives an quantitative estimate on the integral of the above identity.

**Lemma 4.2.** *Recall from (1.6) that*

$$r_G = \frac{1}{16} \min \{ r_p, r_{\text{inj}}(M), \varpi_\kappa, r_\mu \}.$$

For any  $p, q \in M$ , we have

$$\langle p, q \rangle_{\text{cn}} \leq \frac{\langle p, p \rangle_{\text{cn}} + \langle q, q \rangle_{\text{cn}}}{2} - c \cdot \min \{ r_G^2, |pq|^2 \},$$

where  $c = c_{4.2} := \mu_{\min}(r_G) \cdot \frac{\ell_p^2}{2\pi^2}$ .

Let us postpone the proof to the end of the section. With the quantitative estimate, we can collect pairs  $(X_i, X_j)$  whose  $\langle X_i, X_j \rangle_{\text{cn}}$  is close to  $\frac{\langle X_i, X_i \rangle_{\text{cn}} + \langle X_j, X_j \rangle_{\text{cn}}}{2}$ , which allows us to deduce that the distance  $|X_i X_j|$  is small. Building from this we have our first algorithm.

---

**Algorithm 1:** GenerateCluster

---

**Input :**  $U \subseteq \mathbf{V}$ .  $V \subseteq \mathbf{V}$ , a subset of size between 2 and  $n$ .

$$\eta > 0.$$

**Output:**  $W_{\text{gc}} \subseteq V$ , a subset of vertices.

$$i_{\text{gc}} \in W_{\text{gc}}, \text{ a vertex.}$$

- 1 Step 1. Sort  $\{i, j\} \in \binom{V}{2}$  according to  $N_U(i, j)$  from the largest to the smallest, and return a list  $\{i_1, j_1\}, \{i_2, j_2\}, \dots, \{i_L, j_L\}$ , where  $L = \binom{|V|}{2} = |\binom{V}{2}|$ .
- 2 Step 2. Let  $m$  be the largest positive integer such that

$$N_U(i_m, j_m) \geq N_U(i_1, j_1) - \frac{1}{2} c_{4.2} \eta^2,$$

where  $c_{4.2}$  is the constant from Lemma 4.2.

- 3 Step 3. Consider a graph  $G_{\text{gc}}$  with vertex set  $V_{\text{gc}} := (\bigcup_{k \in [m]} \{i_k\}) \cup (\bigcup_{k \in [m]} \{j_k\})$  and edge set  $E_{\text{gc}} := \{\{i_k, j_k\} : k \in [m]\}$ .
  - 4 Step 4. Take a pair  $(W_{\text{gc}}, i_{\text{gc}})$  where  $i_{\text{gc}} \in V_{\text{gc}}$  maximizes the size of neighbors in  $G_{\text{gc}}$  and  $W_{\text{gc}} = \{j \in V_{\text{gc}} : \{i_{\text{gc}}, j\} \in E_{\text{gc}}\} \cup \{i_{\text{gc}}\}$ .
  - 5 **return**  $(W_{\text{gc}}, i_{\text{gc}})$
- 

**Remark 4.3.** The common neighbor counts  $N_U(i, j)$  for all pairs  $\{i, j\} \in \binom{V}{2}$  are sub-entries of the matrix  $A^\top A$ , where  $A$  is the adjacency matrix restricted to the block  $U \times V$ . Using fast sparse matrix–matrix multiplication algorithms, the computation of  $A^\top A$  can be performed in time  $O(n^2 s_n^2 |U|)$  with probability  $1 - o(1)$  assuming the concentration of degree on the graph, and converting  $A$  to a sparse representation requires  $O(n|U|)$  time. Therefore, the overall running time of the algorithm is

$$O(n^2 (s_n^2 |U| + \log n)),$$

which is dominated by either the sorting step or by determining the common neighbor counts  $N_U(i, j)$  for all pairs  $\{i, j\} \in \binom{V}{2}$ .

**Definition 4.4** (( $\eta, m$ )-dense set). *For  $\eta > 0$  and a positive integer  $m$ , a subset of vertices  $W \subseteq \mathbf{V}$  is called  $(\eta, m)$ -dense, if for every  $w \in W$ , the set*

$$\left| \left\{ w' \in W : |X_w X_{w'}| \leq \eta \right\} \right| \geq m.$$

**Proposition 4.5.** *There exists a universal constant  $C_{4.5} > 1$  such that the following holds. Suppose  $V, U$  are two disjoint subset of vertices in our random geometric graph model (Definition 1.5) with  $|V| \leq n$  such that the events  $\mathcal{E}_{\text{cn}}(U, V)$  holds. Let  $\eta$  be a positive parameter satisfying*

$$(32) \quad C_{4.5} \left( \frac{\log(n)}{\sqrt{s_n^2 |U|}} \right)^{1/2} \leq \eta \leq r_G .$$

and set

$$(33) \quad m := \min_{i \in V} \left| \left\{ j \in V : |X_i X_j| \leq c\eta^2 \right\} \right|$$

where

$$c = c_{4.5} = \frac{1}{4} \frac{c_{4.2}}{\max\{L_p, 1\}}$$

and  $c_{4.2}$  is the constant from Lemma 4.2. In other words,  $V$  is a  $(c\eta^2, m)$ -dense set by the definition above. Then, the output of Algorithm 1 returns a pair  $(W_{\text{gc}}, i_{\text{gc}})$  such that the following holds:

$$\forall i \in W_{\text{gc}}, \quad |X_{i_{\text{gc}}} X_i| \leq \eta \text{ and } |W_{\text{gc}}| \geq m .$$

*Proof.* First, from Cauchy–Schwarz inequality, we have

$$\max_{i,j \in V} \langle X_i, X_j \rangle_{\text{cn}} = \max_{i \in V} \langle X_i, X_i \rangle_{\text{cn}} .$$

Let  $i_0 \in V$  be a vertex such that  $\langle X_{i_0}, X_{i_0} \rangle_{\text{cn}} = \max_{i \in V} \langle X_i, X_i \rangle_{\text{cn}}$ . For any  $(i, j) \in \binom{V}{2}$ , by the event  $\mathcal{E}_{\text{cn}}(U, V)$  we have

$$\max_{i,j \in V} N_U(i, j) \leq \max_{i,j \in V} \langle X_i, X_j \rangle_{\text{cn}} + \frac{\log(n)}{\sqrt{s_n^2 |U|}} \leq \langle X_{i_0}, X_{i_0} \rangle_{\text{cn}} + \frac{\log(n)}{\sqrt{s_n^2 |U|}} .$$

Lower bound on  $|W_{\text{gc}}|$ . Consider the set

$$S := \{j \in V \setminus \{i_0\} : |X_{i_0} X_j| \leq c\eta^2\} .$$

Therefore, for any  $j \in S$ , we have

$$(34) \quad \begin{aligned} N_U(i_0, j) &\geq \langle X_{i_0}, X_j \rangle_{\text{cn}} - \frac{\log(n)}{\sqrt{s_n^2 |U|}} \geq \langle X_{i_0}, X_{i_0} \rangle_{\text{cn}} - L_p c\eta^2 - \frac{\log(n)}{\sqrt{s_n^2 |U|}} \\ &\geq \max_{(i,j) \in \binom{V}{2}} N_U(i, j) - L_p c\eta^2 - 2 \frac{\log(n)}{\sqrt{s_n^2 |U|}} \\ &\geq \max_{(i,j) \in \binom{V}{2}} N_U(i, j) - \frac{1}{2} c_{4.2} \eta^2 , \end{aligned}$$

where the last inequality follows from our choice of  $c$  and setting  $C_{4.5}$  sufficiently large.

Therefore, the collection of pairs  $\{(i_0, j) : j \in S\}$  will be edges in the graph  $G_{\text{gc}}$  constructed in Algorithm 1. This implies that the number of neighbors of  $i_{\text{gc}}$  is at least the number of  $i_0$  in  $G_{\text{gc}}$ , which is at least  $|S| \geq m$  from the  $(c\eta^2, m)$ -density assumption on  $V$ .

Upper bound on  $|X_{i_{\text{gc}}} X_j|$  for  $j \in W_{\text{gc}}$ . Any pair  $(i', j') \in \binom{V}{2}$  appears as an edge in the graph  $G_{\text{gc}}$  if and only if

$$N_U(i', j') \geq \max_{(i,j) \in \binom{V}{2}} N_U(i, j) - \frac{1}{2} c_{4.2} \eta^2 .$$

First, replacing  $N_U(i', j')$  by its expected value we have

$$\langle X_{i'}, X_{j'} \rangle_{\text{cn}} \geq \max_{(i,j) \in \binom{V}{2}} N_U(i, j) - \frac{1}{2} c_{4.2} \eta^2 - \frac{\log(n)}{\sqrt{s_n^2 n}} .$$

From the estimate (34), take any  $j'' \in S$ , we have

$$\begin{aligned} \max_{(i,j) \in \binom{V}{2}} N_U(i,j) - \frac{\log(n)}{\sqrt{s_n^2 n}} &\geq N_U(i_0, j'') - \frac{\log(n)}{\sqrt{s_n^2 n}} \\ &\stackrel{(34)}{\geq} \langle X_{i_0}, X_{i_0} \rangle_{\text{cn}} - L_p c \eta^2 - \frac{\log(n)}{\sqrt{s_n^2 n}} - \frac{\log(n)}{\sqrt{s_n^2 n}} \\ &\geq \langle X_{i_0}, X_{i_0} \rangle_{\text{cn}} - \frac{1}{2} c_{4.2} \eta^2, \end{aligned}$$

which implies that

$$\langle X_{i'}, X_{j'} \rangle_{\text{cn}} \geq \langle X_{i_0}, X_{i_0} \rangle_{\text{cn}} - c_{4.2} \eta^2 \geq \frac{\langle X_{i'}, X_{i'} \rangle_{\text{cn}} + \langle X_{j'}, X_{j'} \rangle_{\text{cn}}}{2} - c_{4.2} \eta^2.$$

By Lemma 4.2, the above condition gives

$$|X_{i'} X_{j'}| \leq \eta.$$

□

For the proof of Lemma 4.2, we need a purely geometric statement which we now present as a separate lemma.

**Lemma 4.6.** *Let  $(M, g)$  be a Riemannian manifold whose sectional curvatures are uniformly bounded in absolute value by a constant  $\kappa > 0$ . Let  $r$  be any positive real number which satisfies*

$$0 < r \leq \min \left\{ \frac{1}{4} r_{\text{inj}}(M), \frac{1}{4} \varpi_\kappa \right\}.$$

*Then for any points  $p, q \in M$  with  $|pq| < r$  there exists a geodesic ball  $B = B_{r/4}(z) \subseteq M$  for some  $z \in M$  satisfying  $|pz| = r/2$  such that*

$$|qx| \geq |px| + \frac{1}{\pi} |pq| \quad \text{for every } x \in B.$$

*Proof of Lemma 4.6.* Consider the geodesic moving from  $p$  in the opposite direction from  $p$  to  $q$  and reaching a point in distance  $r/2$ . Let us denote this point by  $z$ . In our notations, we write

$$z = \exp_p \left( -\frac{r}{2} \cdot \frac{\hat{q}}{\|\hat{q}\|} \right).$$

Here we recall that  $\hat{q} \in T_p M$  is the unique tangent vector with  $\|\hat{q}\| < r$  such that  $\exp_p(\hat{q}) = q \in M$ . With  $r/2 < r_{\text{inj}}(M)$ , we know that this geodesic segment between  $p$  and  $z$  is minimizing:  $|pz| = r/2$ . Now consider the geodesic ball of radius  $r/4$  centered at  $z$  and an arbitrary point

$$x \in B_{r/4}(z) \subseteq M.$$

Lower bound on the angle  $\angle(p_x^q)$ . The comparison inequality (7) from Lemma 3.1 applied with the upper curvature bound gives

$$\angle(p_x^z) \leq \angle^\kappa(p_x^z).$$

Since we have

- $|pz| = r/2$ ,
- $|zx| < r/4$ ,
- $|px| \geq |pz| - |zx| > r/2 - r/4 = r/4$ ,
- $|px| \leq |pz| + |zx| < r/2 + r/4 = 3r/4$ , and
- $r \leq \varpi_\kappa/4 = \pi/(4\sqrt{\kappa})$ ,

a direct model-space computation (we leave the explicit computation to Lemma 3.5) yields

$$\angle^\kappa(p_x^z) \leq \frac{\pi}{3}.$$

Consequently

$$\angle(p_x^q) = \angle(p_z^q) - \angle(p_x^z) \geq \pi - \frac{\pi}{3} = \frac{2\pi}{3}.$$

Opposite-side comparison for the triangle  $\triangle pzx$ . By the triangle inequality, we have

$$|px| \leq |pz| + |zx| < \frac{r}{2} + \frac{r}{4} = \frac{3r}{4}.$$

Hence,

$$|px| + |pq| < \frac{3r}{4} + r \leq \varpi_\kappa,$$

where the last inequality follows from the definition of  $r$ . Consequently, the model angle  $\angle^\kappa(p_q^x)$  is well-defined and Lemma 3.1 gives

$$|qx| \geq \gamma^\kappa(\angle(p_x^q); |pq|, |px|) \geq \gamma^\kappa\left(\frac{2}{3}\pi; |pq|, |px|\right),$$

where the last inequality follows from the fact that  $\theta \mapsto \gamma^\kappa(\theta; |pq|, r/2)$  is a monotone increasing on  $[0, \pi]$  (see e.g. [AKP24, Ch. 1]), and we have shown above that  $\angle(p_x^q) \geq \frac{2}{3}\pi$ . Finally, a direct estimate in  $M_\kappa$  (see Lemma A.5) shows that

$$\gamma^\kappa\left(\frac{2}{3}\pi; |pq|, |px|\right) \geq |px| + \frac{1}{2} \left| \cos\left(\frac{2}{3}\pi\right) \right| |pq| = |px| + \frac{1}{4}|pq|.$$

□

*Proof of Lemma 4.2.* The proof is similar to the proof of Lemma 3.1 in [HJM24], which breaks into two cases depending on  $|pq|$ . Recall from Proposition 4.1 that

$$\frac{\langle x, x \rangle_{\text{cn}} + \langle y, y \rangle_{\text{cn}}}{2} - \langle x, y \rangle = \frac{1}{2} \int_M (p(|xz|) - p(|yz|))^2 d\mu(z),$$

so our goal is to estimate the integral from below.

Case 1:  $|pq| < 4r_G$ . Here we can invoke Lemma 4.6 to show that there exists a ball  $B = B_{r_G}(z)$  for some  $z \in M$  with  $|pz| = 2r_G$  such that for every  $x \in B$ , we have

$$||px| - |qx|| \geq \frac{1}{\pi}|pq|.$$

Following triangle inequalities, we have

$$\max\{|px|, |qx|\} < (4+2+1)r_G < r_p.$$

Hence, we have

$$|p(|px|) - p(|qx|)| \geq \frac{\ell_p}{\pi}|pq|.$$

Now we use this bound to estimate the integral

$$\int_M (p(|px|) - p(|qx|))^2 d\mu(x) \geq \int_B (p(|px|) - p(|qx|))^2 d\mu(x) \geq \mu_{\min}(r_G) \frac{\ell_p^2}{\pi^2} |pq|^2.$$

Case 2:  $|pq| \geq 4r_G$ . Now suppose  $|pq| \geq 4r_G$ . Here we consider the ball  $B = B_{r_G}(p)$ . For any  $x \in B$ , simply from triangle inequalities, we have

$$|qx| \geq |pq| - |px| \geq 4r_G - r_G = 3r_G.$$

Together with  $|px| \leq r_G$ , we have

$$|p(|px|) - p(|qx|)| \geq p(3r_G) - p(r_G) \geq \ell_p \cdot 2r_G.$$

Then, applying the same argument as in Case 1, we have

$$\int_M (p(|px|) - p(|qx|))^2 d\mu(x) \geq \int_B (p(|px|) - p(|qx|))^2 d\mu(x) \geq \mu_{\min}(r_G) \cdot \ell_p^2 \cdot 4r_G^2.$$

□

## 5. SCENARIO 1: FORMING AN APPROXIMATE ORTHOGONAL BASIS

We introduce the first scenario in an intermediate step of the overall algorithm.

**Definition 5.1.** For  $k \in [0, d]$ , and consider  $k + 1$  pairs  $(\rho_i)_{i \in [0, k]} \in \mathbf{V}$ . For parameters  $0 < \eta < \delta < r$ , define

$$\mathcal{E}_\perp(r, \delta, \eta, (\rho_i)_{i \in [0, k]})$$

to be the event of  $(X_{U_{\rho_i}})_{i \in [0, k]}$  that the following two conditions hold:

$$(35) \quad |X_{\rho_0} X_{\rho_i}| - r | \leq \delta, \quad \text{for } i \in [k], \quad \text{and} \quad |X_{\rho_i} X_{\rho_j}| - \sqrt{2}r | \leq \delta, \quad \text{for } i, j \in [k],$$

The main statement of this scenario is as follows:

**Proposition 5.2.** There exists a constant  $c_{5.2} = c(d, L_p, \ell_p) > 0$  and  $c'_{5.2} = c'(d, L_p, \ell_p) > 0$  such that the following holds. For any  $c < c_{5.2}$ , let  $0 < \eta < \delta < r$  and a positive integer  $m$  satisfies the parameter configuration (20):

$$(36) \quad \varepsilon_m < c^2 \eta, \quad \eta = c\delta, \quad \delta = cr, \quad r \leq c r_G.$$

Suppose we are given  $k + 1$  pairs  $(\rho_i, U_{\rho_i})_{i \in [0, k]}$  with  $\rho_i \subseteq U_{\rho_i} \subset \mathbf{V}$  and  $|U_{\rho_i}| \geq m$  (with  $k \leq d$ ).

Let  $W \subset \mathbf{V}$  be disjoint from all  $U_{\rho_i}$  and let

$$m' := \inf_q \left| \{X_w\}_{w \in W} \cap B_{c'_{5.2} c_{4.5} \eta^2}(q) \right|$$

where the infimum is taken over all points  $q \in M$  satisfying

$$|X_{\rho_0} q| - r | < 0.92\delta \quad \text{and} \quad |X_{\rho_i} q| - \sqrt{2}r | < 0.92\delta \quad (\forall i \in [k]).$$

Then, on the event

$$\mathcal{E} := \mathcal{E}_\perp(r, \delta, \eta, (\rho_i)_{i \in [0, k]}) \cap \bigcap_{i \in [0, k]} \mathcal{E}_{\text{clu}}(\eta, U_{\rho_i}, \rho_i) \cap \bigcap_{i \in [0, k]} \mathcal{E}_{\text{navi}}(U_{\rho_i}, W),$$

the set

$$(37) \quad \begin{aligned} W' &:= W'(r, \delta, W, (U_{\rho_i}, \rho_i)_{i \in [k]}) \\ &= \left\{ w \in W : \begin{array}{l} p(r + 0.91\delta) \leq N_{U_{\rho_0}}(w) \leq p(r - 0.91\delta), \\ p(\sqrt{2}r + 0.91\delta) \leq N_{U_{\rho_i}}(w) \leq p(\sqrt{2}r - 0.91\delta) \quad (\forall i \in [k]) \end{array} \right\} \end{aligned}$$

is non-empty, and for every  $w \in W'$ , the following holds:

$$|B_{c_{4.5}\eta^2}(X_{w'}) \cap W'| \geq m'.$$

**Remark 5.3.** Within the event  $\mathcal{E}$ , by (14) from Calibration Lemma 2.6, for every  $w' \in W'$  we have

$$|X_{\rho_0} X_{w'}| - r | \leq 0.92\delta \quad \text{and} \quad |X_{\rho_i} X_{w'}| - \sqrt{2}r | \leq 0.92\delta,$$

provided that  $c$  is chosen sufficiently small.

**Remark 5.4.** In view of Algorithm 1, the set  $W'$  defined in (37) is a  $(c_{4.5}\eta^2, m')$ -dense set, which is capable of extracting a cluster in  $W'$  with radius  $\eta$  and size at least  $m'$  by Proposition 4.5 by measuring common neighbor sizes.

**5.1.  $W'$  is not empty.** Our goal in this subsection is to show that the set  $W'$ , defined in Proposition 5.2, is non-empty. The key idea is to construct a point  $y \in M$  that is approximately orthogonal to the points  $X_{\rho_i} \in M$ , where each  $X_{\rho_i}$  belongs to one of the clusters  $U_{\rho_i}$ . This construction is possible when  $r$  is sufficiently smaller than  $\min\{r_{\text{inj}}(M), \varpi_\kappa\}$ , which minimizes the effect of curvature, and when  $\delta/r$  is small enough to ensure approximate orthogonality. Both conditions can be achieved by choosing a sufficiently small constant  $c$  in the parameter assumptions of (36) from Proposition 5.2. We now formalize this in the following geometric lemma:

**Lemma 5.5.** *Let  $p, x_1, \dots, x_k \in M$ , where  $k < d$ , be points in  $M$  such that*

$$|p x_i| - r \leq \delta, \quad \text{for } i \in [k],$$

*and assume  $r > \delta > 0$  satisfies the parameter assumption from Proposition 5.2, with a sufficiently small constant  $c$ . Then, there exists a point  $y \in M$  such that*

$$|p y| = r, \quad \text{and} \quad |y x_i| - \sqrt{2}r \leq 0.9\delta \quad \text{for } i \in [k].$$

We now apply this geometric construction to prove the non-emptiness of  $W'$  under the assumptions of Proposition 5.2. The proof of Lemma 5.5 is postponed to the end of this subsection.

**Corollary 5.6.** *Under the assumptions of Proposition 5.2,  $W'$  is not empty.*

*Proof.* Let  $p = X_{\rho_0}$  and  $x_i = X_{\rho_i}$  for  $i \in [k]$ . These satisfy the assumption of Lemma 5.5 by construction. Hence, there exists a point  $q \in M$  such that

$$|p q| = r, \quad \text{and} \quad |q x_i| - \sqrt{2}r \leq 0.9\delta \quad \text{for } i \in [k].$$

This point  $q$  satisfies the geometric constraints of Proposition 5.2. By assumption, this implies the existence of some  $w \in W$  such that

$$|p X_w| \leq c_{4.5} \eta^2.$$

Using the triangle inequality:

$$|X_{\rho_0} X_w| - r \leq c_{4.5} \eta^2 \quad \text{and} \quad |X_{\rho_i} X_w| - \sqrt{2}r \leq 0.9\delta + c_{4.5} \eta^2.$$

Under the event  $\mathcal{E}_{\text{clu}}(\eta, U_{\rho_0}, \rho_0)$  and the event  $\mathcal{E}_{\text{navi}}(U_{\rho_0}, W)$ , the Calibration Lemma 2.6 applies, yielding

$$p(r + 0.9\delta + 2\eta) \leq N_{U_{\rho_0}}(w) \leq p(r - 0.9\delta - 2\eta),$$

and when  $c$  is sufficiently small, we have  $2\eta < 0.01\delta$ . Further, an identical argument applies to all  $i \in [k]$ , showing that  $w \in W'$ . Therefore,  $W'$  is non-empty.  $\square$

*Proof of Lemma 5.5.* Recall

$$\hat{x}_i := (\hat{x}_i)_p \in T_p M \quad \text{for } i \in [d]$$

is the unique vector (thanks to  $|x_i p| < c \min\{r_{\text{inj}}(M), \varpi_\kappa\}$ ) such that  $\exp_p(\hat{x}_i) = x_i$  and  $\|\hat{x}_i\| = |p x_i|$ .

Since  $k < d$ , there exists a vector  $w \in T_p M$  orthogonal to all  $\hat{x}_i$ , i.e.,

$$\langle \hat{x}_i, w \rangle = 0 \quad \text{for all } i \in [k] \quad \text{and} \quad \|w\| = r.$$

We claim that

$$y = \exp_p(w)$$

is the desired point. Clearly,

$$|p y| = \|w\| = r.$$

To estimate  $|y x_i|$ , we note that

$$\angle(p_y^{x_i}) = \arccos\left(\frac{\langle \hat{x}_i, w \rangle}{\|\hat{x}_i\| \|w\|}\right) = \arccos(0) = \frac{\pi}{2}.$$

We now apply Lemma A.7 with  $\theta = \frac{\pi}{2}$ ,  $|p x_i| = \|\hat{x}_i\|$ , and  $|p y| = r$  to obtain

$$\left| |x_i y|^2 - 2r^2 \right| \leq 2\delta r + \delta^2 + 2\kappa(r + \delta)^4.$$

With the assumption that  $r \leq c \min\{r_{\text{inj}}(M), \varpi_{\kappa_1}\}$  and  $\delta = cr$ , we have

$$2\kappa(r + \delta)^4 = 2\kappa(1 + c)^4 r^4 \leq 2(1 + c)^4 c^2 r^2 \leq 2(1 + c)^4 \delta^2 \leq 3\delta^2.$$

Thus,

$$2\delta r + \delta^2 + 2\kappa(r + \delta)^4 \leq \delta r(2 + 3c) \leq 3cr^2.$$

We can first conclude that

$$|x_i y| \geq \sqrt{2r^2 - 3cr^2} = \sqrt{2 - 3c}r,$$

and consequently,

$$\left| |x_i y| - \sqrt{2}r \right| = \frac{| |x_i y|^2 - 2r^2 |}{| |x_i y| + \sqrt{2}r |} \leq \frac{\delta r(2 + 3c)}{\sqrt{2 - 3c}r + \sqrt{2}r} \leq 0.9\delta,$$

where the last inequality holds when  $c$  is sufficiently small.  $\square$

**5.2. Proof of Proposition 5.2.** Having established that  $W'$  is non-empty, we now aim to show that for every  $w' \in W'$ , there exists a ball of radius  $\eta$  centered at  $X_{w'}$  such that

$$|B_\eta(X_{w'}) \cap W| \geq m'.$$

Intuitively,  $W'$  corresponds to the subset of latent points in  $W$  that lie near the intersection of  $k+1$  fuzzy annuli — “fuzzy” in the sense that membership is defined via edge constraints to cluster cores rather than precise metric distances. The main difficulty arises when a point  $X_{w'} \in W'$  is close to the boundary of this intersection, where small perturbations may break the constraints.

To address this, we consider moving slightly along a geodesic  $\gamma$ , starting at  $X_{w'}$  in a direction  $\xi \in T_{X_{w'}} M$ , to reach a point  $z = \gamma(s)$  for some small  $s > 0$ , thereby moving away from the boundary of the intersection region.

Strategically, the goal is to find such a direction  $\xi$  so that the resulting point  $z$  satisfies

$$B_{c_{4.5}\eta^2}(z) \cap W \subseteq W',$$

thus ensuring that a robust neighborhood of  $z$  lies entirely within  $W'$ .

For a fixed  $w' \in W'$ , define

$$\sigma_0 = \begin{cases} 1 & \text{if } N_{U_{\rho_0}}(w') < p(r), \\ -1 & \text{otherwise.} \end{cases}$$

We then choose  $\xi \in T_{X_{w'}} M$  such that

$$\text{sign}(\langle \xi, \hat{\rho}_0 \rangle) = \sigma_0,$$

where  $\hat{\rho}_0 \in T_{X_{w'}} M$  is the tangent vector satisfying  $\exp_{X_{w'}}(\hat{\rho}_0) = X_{\rho_0}$ . Suppose  $N_{U_{\rho_0}}(w') < p(r)$ , this suggests that  $|X_{\rho_0} X_{w'}| > r$ . In this case, setting  $\sigma_0 = 1$  means choosing  $\xi$  so that  $\langle \xi, \hat{\rho}_0 \rangle > 0$ , which moves  $z$  toward  $X_{\rho_0}$  along the geodesic and decreases  $|X_{\rho_0} z|$ , bringing it closer to  $r$ . The reverse case  $N_{U_{\rho_0}}(w') > p(r)$  corresponds to  $\sigma_0 = -1$ , so that moving in the opposite direction increases the distance toward  $r$ .

Following the same reasoning, for each  $i \in [k]$  we set

$$\sigma_i = \begin{cases} 1 & \text{if } N_{U_{\rho_i}}(w') < p(\sqrt{2}r), \\ -1 & \text{otherwise.} \end{cases}$$

Next, choose  $\xi' \in \text{span}(\{\tilde{\rho}_i\}_{i \in [0, k]})$  such that

$$\text{sign}(\langle \xi', \tilde{\rho}_i \rangle) = \sigma_i \quad \text{for all } i \in [0, k],$$

where

$$\tilde{\rho}_i = \frac{\hat{\rho}_i}{\|\hat{\rho}_i\|} \in T_{X_{w'}} M \quad \text{for } i \in [0, k].$$

Finally, normalize to obtain

$$\xi = \frac{\xi'}{\|\xi'\|}.$$

In other words, we have

$$\langle \xi, \tilde{\rho}_i \rangle = \frac{\sigma_i}{\|\xi'\|}.$$

We need to ensure that  $\|\xi'\|$  is not too large, so that moving along the direction  $\xi$  with a small distance still produces a non-trivial change in the distance to each  $X_{\rho_i}$ .

**Lemma 5.7.** *Under the assumptions of Proposition 5.2, for any fixed  $w' \in W'$ , the vector  $\xi$  defined above satisfies*

$$\|\xi'\|_2 \leq \sqrt{2(k+1)}, \quad \text{and} \quad \frac{\langle \xi, \tilde{\rho}_i \rangle}{\sigma_i} \geq \frac{1}{2(k+1)}$$

provided  $c_{5.2}$  is chosen sufficiently small.

*Proof.* Let  $\xi' = \sum_{i=0}^k a_i \tilde{\rho}_i$  with  $a_i \in \mathbb{R}$ . Then the coefficients satisfy

$$B\vec{a} = \vec{\sigma}, \quad B := (\langle \tilde{\rho}_i, \tilde{\rho}_j \rangle)_{i,j=0}^k, \quad \vec{a} := (a_i)_{i=0}^k, \quad \vec{\sigma} := (\sigma_i)_{i=0}^k.$$

In particular,  $\vec{a} = B^{-1}\vec{\sigma}$ . Hence

$$\|\xi'\|_2^2 = \left\| \sum_{i=0}^k a_i \tilde{\rho}_i \right\|_2^2 = \vec{a}^\top B \vec{a} = \vec{\sigma}^\top B^{-1} \vec{\sigma} \leq \|B^{-1}\|_{\text{op}} \|\vec{\sigma}\|_2^2 \leq \|B^{-1}\|_{\text{op}} (k+1).$$

Therefore, it suffices to show that  $s_{\min}(B) \geq \frac{1}{2(k+1)}$ , i.e.  $\|B^{-1}\|_{\text{op}} \leq 2(k+1)$ , which then yields  $\|\xi'\| \leq \sqrt{2(k+1)}$ .

From law of cosine to inner products. For any  $\{i, j\} \in \binom{[0, k]}{2}$ , consider the geodesic triangle with vertices  $(X_{w'}, X_{\rho_i}, X_{\rho_j})$ . Let

$$\theta_{ij} = \angle(X_{w'} X_{\rho_i} X_{\rho_j}) \in [0, \pi] \quad \text{and notice that} \quad B_{ij} = \langle \tilde{\rho}_i, \tilde{\rho}_j \rangle = \cos(\theta_{ij}).$$

Now we invoke Lemma A.7, a comparison version of law of cosines, to get

$$\left| |X_{\rho_i} X_{\rho_j}|^2 - (|X_{w'} X_{\rho_i}|^2 + |X_{w'} X_{\rho_j}|^2 - 2 |X_{w'} X_{\rho_i}| |X_{w'} X_{\rho_j}| \cos \theta_{ij}) \right| \leq C \kappa (|X_{w'} X_{\rho_i}|^4 + |X_{w'} X_{\rho_j}|^4).$$

Under the event defining  $W'$  and  $\mathcal{E}_\perp$ , we have, for  $i, j \in [k]$ ,

$$|X_{\rho_i} X_{\rho_j}| = \sqrt{2r} \pm \delta, \quad |X_{w'} X_{\rho_i}| = \sqrt{2r} \pm 0.92\delta,$$

and for pairs involving 0,

$$|X_{\rho_0} X_{\rho_i}| = r \pm \delta, \quad |X_{w'} X_{\rho_0}| = r \pm 0.92\delta, \quad |X_{w'} X_{\rho_i}| = \sqrt{2r} \pm 0.92\delta.$$

Plugging these into the display above and rearranging yields the following bounds.

Case 1:  $i, j \in [k]$ . Using  $(\sqrt{2r})^2 = 2r^2$  and collecting terms,

$$4r^2(1 - \cos \theta_{ij}) = 2r^2 + \text{Err}_{ij},$$

where  $|\text{Err}_{ij}| \leq C_1(\delta r + \kappa r^4)$ . Hence

$$|\cos \theta_{ij} - \frac{1}{2}| \leq C_2 \left( \frac{\delta}{r} + \kappa r^2 \right).$$

Case 2: one of  $i, j$  equals 0. Here

$$3r^2 - 2\sqrt{2}r^2 \cos \theta_{0i} = r^2 + \text{Err}_{0i}, \quad |\text{Err}_{0i}| \leq C_3(\delta r + \kappa r^4),$$

so

$$|\cos \theta_{0i} - 2^{-1/2}| \leq C_4\left(\frac{\delta}{r} + \kappa r^2\right).$$

On the diagonal we have  $B_{ii} = 1$  for all  $i$ .

By the parameter regime in Proposition 5.2 (namely  $\delta = cr$  and  $r \leq c \min\{r_{\text{inj}}(M), \varpi_\kappa\}$  so that  $\kappa r^2 \lesssim c^2$ ), there is a constant  $C_*$  such that

$$\varepsilon := C_*\left(\frac{\delta}{r} + \kappa r^2\right) = O(c_{5.2}),$$

and consequently

$$B_{ij} = \begin{cases} 1 & (i = j), \\ \frac{1}{2} \pm \varepsilon & (i, j \in [k], i \neq j), \\ \frac{\sqrt{2}}{2} \pm \varepsilon & (\text{one of } i, j \text{ is 0}). \end{cases}$$

Equivalently,  $B$  is a small perturbation (of size  $\varepsilon$ ) of the deterministic Gram matrix with entries  $\{1, \frac{1}{2}, \frac{\sqrt{2}}{2}\}$ , i.e. angles  $\{\pi/3, \pi/4\}$  up to  $O(\varepsilon)$ .

Least singular value of the baseline matrix. Let  $\bar{B}$  be the  $(k+1) \times (k+1)$  Gram matrix with

$$\bar{B}_{ii} = 1, \quad \bar{B}_{0i} = \bar{B}_{i0} = \alpha := \frac{\sqrt{2}}{2} \quad (i \geq 1), \quad \bar{B}_{ij} = \beta := \frac{1}{2} \quad (i \neq j, i, j \geq 1).$$

We decompose  $\mathbb{R}^{k+1}$  as the orthogonal sum

$$H := \text{span}\{e_0, (0, \frac{1}{\sqrt{k}}\mathbf{1})\} \quad \text{and} \quad H^\perp := \{(0, v) : \mathbf{1}^\top v = 0\}.$$

On  $H^\perp$  (dimension  $k-1$ ),  $\bar{B}$  acts by multiplication with  $1 - \beta = \frac{1}{2}$ , so there are  $k-1$  eigenvalues equal to  $\frac{1}{2}$ . On  $H$ , with basis  $\{e_0, (0, \frac{1}{\sqrt{k}}\mathbf{1})\}$ , the restriction of  $\bar{B}$  is

$$D = \begin{pmatrix} 1 & \frac{\sqrt{2}}{2}\sqrt{k} \\ \frac{\sqrt{2}}{2}\sqrt{k} & \frac{k+1}{2} \end{pmatrix}.$$

This  $2 \times 2$  matrix has trace  $\frac{k+3}{2}$  and determinant  $\frac{1}{2}$ , hence its eigenvalues are

$$\lambda_\pm(D) = \frac{\frac{k+3}{2} \pm \sqrt{\left(\frac{k+3}{2}\right)^2 - 2}}{2}. \quad \Rightarrow \quad \lambda_{\min}(\bar{B}) = \lambda_-(D) = \frac{\frac{k+3}{2} - \sqrt{\left(\frac{k+3}{2}\right)^2 - 2}}{2} \geq \frac{1}{k+3}.$$

Perturbation to the actual  $B$ . The actual Gram matrix satisfies  $B = \bar{B} + E$ , where

$$|E_{ij}| \leq \varepsilon, \quad \varepsilon = C\left(\frac{\delta}{r} + \kappa r^2\right) = O(c_{5.2}).$$

Thus  $\|E\|_{\text{op}} \lesssim k\varepsilon$ , and Weyl's inequality gives

$$\sigma_{\min}(B) \geq \sigma_{\min}(\bar{B}) - \|E\|_{\text{op}} \geq \frac{1}{k+3} - C'k\varepsilon \geq \frac{1}{2(k+1)},$$

provided that  $c_{5.2}$  is chosen small enough.  $\square$

**Lemma 5.8.** *Under the assumptions of Proposition 5.2 and fix  $w' \in W'$ . Let  $\xi = \frac{\xi'}{\|\xi'\|} \in T_{X_{w'}} M$  be the unit direction chosen as in the construction above, i.e.*

$$\xi' \in \text{span}(\{\tilde{\rho}_i\}_{i \in [0, k]}) \quad \text{satisfies} \quad \langle \xi, \tilde{\rho}_i \rangle = \sigma_i \quad \text{for all } i \in [0, k],$$

with  $\tilde{\rho}_i = \hat{\rho}_i / \|\hat{\rho}_i\|$  and  $\exp_{X_{w'}}(\hat{\rho}_i) = X_{\rho_i}$ . Set  $z = \exp_{X_{w'}}(t \xi)$  for  $0 < t \leq \delta$ . When  $c_{5.2}$  is chosen sufficiently small, we have

$$(38) \quad \frac{\ell_p}{2} \frac{t}{\|\xi'\|} \leq \sigma_i \cdot \frac{1}{|U_{\rho_i}|} \sum_{v \in U_{\rho_i}} \left( p(|z X_v|) - p(|X_{w'} X_v|) \right) \leq 2L_p \frac{t}{\|\xi'\|}.$$

*Proof.* Fix  $w' \in W'$  and let  $\xi \in T_{X_{w'}} M$  be the unit vector chosen above. For  $t > 0$  small, set

$$z := \exp_{X_{w'}}(t \xi).$$

For each  $i \in [0, k]$ , we apply Lemma 3.6 with

$$p := X_{w'}, \quad x := X_{\rho_i}, \quad q := z.$$

Verification of the conditions of Lemma 3.6. We have  $|p q| = |X_{w'} z| = t$  and

$$|\cos(\angle(p_x^q))| = |\langle \xi, \tilde{\rho}_i \rangle| \geq \frac{1}{2(k+1)}$$

by Lemma 5.7. Moreover, by Remark 5.3 we have

$$|X_{w'} X_{\rho_i}| = \begin{cases} r \pm 0.92\delta, & i = 0, \\ \sqrt{2}r \pm 0.92\delta, & i \geq 1, \end{cases}$$

and both quantities are  $O(r)$ . Thus, provided  $t \leq \delta$  and  $c_{5.2}$  is chosen sufficiently small, the two conditions from Lemma 3.6 are satisfied:

$$16 \frac{t}{|X_{w'} X_{\rho_i}|} \leq \frac{1}{2(k+1)} \leq |\cos(\angle(X_{w'} z_{X_{\rho_i}}))|,$$

and

$$|X_{w'} X_{\rho_i}|, \quad t < \frac{1}{8} \min\{r_{\text{inj}}(M), \varpi_\kappa\}.$$

Choose  $x' = X_v$  for  $v \in U_{\rho_i}$ . Now we take  $x'$  to be any  $X_v$  with  $v \in U_{\rho_i}$  in the setting of Lemma 3.6.

Within the event  $\mathcal{E}_{\text{clu}}(U_{\rho_i}, \rho_i)$ ,

$$|X_{\rho_i} X_v| \leq \eta \quad \text{for all } v \in U_{\rho_i}.$$

Since  $|p x| = |X_{w'} X_{\rho_i}| \geq r - 0.92\delta$  and  $\eta = c_{5.2} \delta$  with  $c_{5.2}$  small, we obtain

$$16 \frac{|x x'|}{|p x|} \leq 16 \frac{\eta}{r - 0.92\delta} \leq 16 \frac{c_{5.2} \delta}{r - 0.92\delta} \ll \frac{1}{2(k+1)} \leq |\cos(\angle(p_x^q))|,$$

where the last inequality uses Lemma 5.7. Thus the second hypothesis of Lemma 3.6 is also satisfied for this choice of  $x'$ .

Consequence. Recalling  $\cos(\angle(p_x^q)) = \langle \xi, \tilde{\rho}_i \rangle = \sigma_i / \|\xi'\|$ , Lemma 3.6 yields

$$(39) \quad \frac{\ell_p}{2} \frac{t}{\|\xi'\|} \leq \sigma_i \left( p(|q x'|) - p(|p x'|) \right) \leq 2L_p \frac{t}{\|\xi'\|}.$$

Averaging (39) over  $v \in U_{\rho_i}$  gives

$$\frac{\ell_p}{2} \frac{t}{\|\xi'\|} \leq \sigma_i \cdot \frac{1}{|U_{\rho_i}|} \sum_{v \in U_{\rho_i}} \left( p(|z X_v|) - p(|X_{w'} X_v|) \right) \leq 2L_p \frac{t}{\|\xi'\|}.$$

□

*Proof of Proposition 5.2.* Choose  $t$  so that

$$(40) \quad \frac{\ell_p}{2} \frac{t}{\|\xi'\|} = 2\varepsilon_m + L_p c'_{5.2} c_{4.5} \eta^2.$$

By (36) and  $\|\xi'\| \leq \sqrt{2(k+1)} \leq \sqrt{2(d+1)}$  from Lemma 5.7, we can assume this choice satisfies

$$t \leq \frac{1}{2} c_{5.2} \eta^2 \leq 0.01\delta,$$

provided that  $c_{5.2}$  and  $c'_{5.2}$  are sufficiently small. In particular, it is admissible in Lemma 5.8.

Let  $z$  be the point from Lemma 5.8 corresponding to this  $t$ . And by Lemma 5.8, It is rather straightforward to see that  $z$  satisfies the conditions in Proposition 5.2:

$$|X_{\rho_0} z - r| < 0.92\delta \quad \text{and} \quad |X_{\rho_i} z - \sqrt{2}r| < 0.92\delta \quad (\forall i \in [k]),$$

since  $X_{w'}$  satisfies the same conditions with  $0.92\delta$  replaced by  $0.91\delta$  and  $|X_{w'} z| = t \leq 0.01\delta$ .

Take any  $w \in W$  with

$$|z X_w| \leq c'_{5.2} c_{4.5} \eta^2,$$

whose existence is guaranteed by the definition of  $W$ .

Example case  $i = 0$  and  $\sigma_0 = 1$ . From  $\mathcal{E}_{\text{navi}}(U_{\rho_0}, w) \supseteq \mathcal{E}_{\text{navi}}(U_{\rho_0}, W)$ , we have

$$N_{U_{\rho_0}}(w) \geq p_{U_{\rho_0}}(w) - \varepsilon_m \geq \frac{1}{|U_{\rho_0}|} \sum_{v \in U_{\rho_0}} p(|z X_v|) - L_p |z X_w| - \varepsilon_m.$$

Then, invoke Lemma 5.8 with  $i = 0$  and  $\sigma_0 = 1$  to get

$$(*) \geq p_{U_{\rho_0}}(w') + \frac{\ell_p}{2} \frac{t}{\|\xi'\|} - L_p |z X_w| - \varepsilon_m \geq p_{U_{\rho_0}}(w') + \varepsilon_m \geq N_{U_{\rho_0}}(w') \geq p(r + 0.91\delta),$$

where the second last inequality uses the definition of  $W'$ . For the bound on the other side, notice  $\sigma_0 = 1$  also implies  $N_{U_{\rho_0}}(w') < p(r)$ , the analogous application of Lemma 5.8 to the upper bound yields

$$N_{U_{\rho_0}}(w) \leq N_{U_{\rho_0}}(w') + 2L_p \frac{t}{\|\xi'\|} + L_p |z X_w| + \varepsilon_m \leq p(r) + 0.1\delta$$

when  $c_{4.5}$  is sufficiently small.

Conclusion. We have shown that for  $i = 0$  and  $\sigma_0 = 1$ ,  $w$  satisfies the  $U_{\rho_0}$  constraints in the definition of  $W'$ . The same argument works verbatim for all  $i \in [0, k]$  and both signs of  $\sigma_i$ . Hence

$$B_{c'_{5.2} c_{4.5} \eta^2}(z) \cap W \subseteq W'.$$

Moreover,  $|z X_w| \leq t + c'_{5.2} c_{4.5} \eta^2 \leq c_{4.5} \eta^2$ , so

$$|B_{c_{4.5} \eta^2}(z) \cap W'| \geq |B_{c'_{5.2} c_{4.5} \eta^2}(z) \cap W| \geq m'.$$

□

## 6. SCENARIO 2: POINT NAVIGATION VIA APPROXIMATE ORTHOGONAL BASIS

This is a continuation of the previous scenario, where we will focus on point navigation once we established the approximate orthogonal basis. The main result of this section is the following proposition

**Proposition 6.1.** *There exists a constant  $c_{6.1} = c(d, L_p, \ell_p) > 0$  such that the following holds. Suppose  $0 < c < c_{6.1}$ ,  $0 < \eta < \delta < r$  and a positive integer  $m$  satisfies the parameter configuration (20):*

$$(41) \quad \varepsilon_m := \frac{\log n}{\sqrt{s_n m}} < c\eta^2, \quad \eta = c\delta, \quad \delta = cr, \quad r \leq c \min\{r_{\text{inj}}(M), \varpi_\kappa, r_p\}.$$

Suppose we are given  $d + 1$  pairs  $\mathbf{F} = (\rho_i, U_{\rho_i})_{i \in [0, d]}$  with  $\rho_i \subseteq U_{\rho_i} \subset \mathbf{V}$  and  $|U_{\rho_i}| \geq m$ . Further, let  $w_1, w_2 \in \mathbf{V}$  are two vertices disjoint from  $U_{\rho_i}$  for all  $i \in [0, d]$ . Let

$$\psi_{\mathbf{F}}(w_1, w_2) := \sqrt{\sum_{i \in [d]} (\mathcal{N}_{U_{\rho_i}}(w_2) - \mathcal{N}_{U_{\rho_i}}(w_1))^2}.$$

Then, on the event

$$\mathcal{E} := \mathcal{E}_{\perp}(r, \delta, \eta, (\rho_i)_{i \in [0, d]}) \cap \bigcap_{i \in [0, d]} \mathcal{E}_{\text{clu}}(\eta, U_{\rho_i}, \rho_i) \cap \bigcap_{i \in [0, d]} \mathcal{E}_{\text{navi}}(U_{\rho_i}, \{w_1, w_2\}),$$

the following implication holds:

$$\begin{aligned} & \max\{\mathcal{N}_{U_{\rho_0}}(w_1), \mathcal{N}_{U_{\rho_0}}(w_2)\} \geq p(9\delta) \\ \Rightarrow \quad & \frac{1}{3} L_p^{-1}(\psi_{\mathbf{F}}(w_1, w_2) - 2\sqrt{d}\varepsilon_m) \leq |X_{w_1} X_{w_2}| \leq 3 L_p^{-1}(\psi_{\mathbf{F}}(w_1, w_2) + 2\sqrt{d}\varepsilon_m). \end{aligned}$$

In other words, if we have two vertices  $w_1$  and  $w_2$  whose latent distances are sufficiently close to  $X_{\rho_0}$ , then we can estimate their latent distance using  $(\mathcal{N}_{U_{\rho_i}}(w_j))_{i \in [0, k], j \in [1, 2]}$  with a high precision.

**Lemma 6.2.** *Let  $0 < \delta < r$  satisfy the same condition as in Proposition 6.1. Suppose  $p, x_0, x_1, \dots, x_d \in M$  satisfy*

$$|x_0 x_i| - r \leq \delta, \quad ||x_i x_j| - \sqrt{2}r| \leq \delta \quad \forall i \neq j \in [d], \quad |p x_0| \leq 10\delta.$$

Let

$$\hat{x}_i = (\hat{x}_i)_p, \quad \vec{y}_i := \frac{\hat{x}_i}{\|\hat{x}_i\|} \in T_p M, \quad i \in [d],$$

and define the Gram matrix

$$A = (a_{ij})_{i,j \in [d]} = (\langle \vec{y}_i, \vec{y}_j \rangle_p)_{i,j \in [d]}.$$

Then

$$\|A - I_d\| \leq C d \left( \frac{\delta}{r} + \kappa r^2 \right),$$

for some universal constant  $C > 0$ .

*Proof.* To present the proof without focus on the computation, we will use  $C, C', C''$  to denote universal constants that may change from line to line.

Off-diagonal entries. For  $i \neq j$ , by definition

$$\langle \vec{y}_i, \vec{y}_j \rangle_p = \cos(\angle(p_{x_j}^{x_i})).$$

By triangle inequality, we know that

$$|p x_i| - r \leq |p x_0| + |x_0 x_i| - r \leq 10\delta + \delta \leq 11\delta.$$

Applying Lemma A.7 to the triangle  $\triangle p x_i x_j$  gives

$$\left| \cos(\angle(p_{x_j}^{x_i})) - \frac{|p x_i|^2 + |p x_j|^2 - |x_i x_j|^2}{2 |p x_i| |p x_j|} \right| \leq \frac{\kappa (r + 11\delta)^4}{(r - 11\delta)^2} \leq 2\kappa r^2.$$

From the distance assumptions,

$$\left| |p x_i|^2 + |p x_j|^2 - |x_i x_j|^2 \right| \leq C\delta r.$$

Then,

$$\left| \frac{|p x_i|^2 + |p x_j|^2 - |x_i x_j|^2}{2 |p x_i| |p x_j|} \right| \leq \frac{C \delta r}{2(r - 11\delta)^2} \leq C' \frac{\delta}{r}.$$

Combining with the curvature error term above yields

$$|\langle \vec{y}_i, \vec{y}_j \rangle_p| \leq C'' \left( \frac{\delta}{r} + 2\kappa r^2 \right), \quad i \neq j,$$

Diagonal entries. By construction  $\|\vec{y}_i\|_p = 1$ , so  $a_{ii} = 1$  for all  $i$ .

Norm bound. The off-diagonal bound gives

$$\|A - I_d\| \leq \|A - I_d\|_{\text{HS}} \leq C''d \left( \frac{\delta}{r} + \kappa r^2 \right),$$

as claimed.  $\square$

**Corollary 6.3.** *Under the assumptions of Lemma 6.2, the vectors  $\{\vec{y}_i\}_{i=1}^d$  form a basis of  $T_p M$  and, for all  $\vec{q} \in T_p M$ ,*

$$(1 - \varepsilon) \sum_{i=1}^d \langle \vec{y}_i, \vec{q} \rangle_p^2 \leq \|\vec{q}\|^2 \leq (1 + \varepsilon) \sum_{i=1}^d \langle \vec{y}_i, \vec{q} \rangle_p^2,$$

where  $\varepsilon := 2Cd \left( \frac{\delta}{r} + \kappa r^2 \right) = \Theta(c_{6.1})$ .

*Proof.* Let  $A$  be the Gram matrix from Lemma 6.2. That lemma gives

$$\|A - I_d\| \leq Cd \left( \frac{\delta}{r} + \kappa r^2 \right) := \alpha \lesssim dc_{6.1}.$$

When  $c_{6.1}$  is chosen small enough, we have  $A$  is positive definite and  $\{\vec{y}_i\}$  is a basis. For  $\vec{q} = \sum_i q^i \vec{y}_i$  we have  $\|\vec{q}\|^2 = q^\top A q$  and  $\sum_i \langle \vec{y}_i, \vec{q} \rangle_p^2 = q^\top q$ . The bound on  $\|A - I_d\|$  implies  $\|A^{-1} - I_d\| \leq \frac{\alpha}{1-\alpha} \leq 2\alpha$ , hence

$$|(q^\top q) - (q^\top A^{-1} q)| \leq 2\alpha \|q\|^2,$$

which is exactly the stated inequality with  $\varepsilon = 2\alpha$ .  $\square$

*Proof of Proposition 6.1.* Define

$$\mathbf{b} = (b_i)_{i \in [d]} := \left( p_{U_{\rho_i}}(w_2) - p_{U_{\rho_i}}(w_1) \right)_{i \in [d]}.$$

We **claim** that

$$(42) \quad \frac{1}{\sqrt{5}} L_p^{-1} \|\mathbf{b}\|_2 \leq |X_{w_1} X_{w_2}| \leq \sqrt{5} \ell_p^{-1} \|\mathbf{b}\|_2$$

Once the claim is established, we can simply compare  $\psi_{\mathbf{F}}(w_1, w_2)$  and  $\|\mathbf{b}\|_2$ .

Let us fix an instance of  $X_{U_{\rho_i}} = x_{U_{\rho_i}}$  for  $i \in [0, d]$ ,  $X_{w_j} = x_{w_j}$  for  $j \in \{1, 2\}$ , and  $\mathcal{U}_{U_{\rho_i}, \{w_1, w_2\}} = \mathbf{u}_{U_{\rho_i}, \{w_1, w_2\}}$  such that the events described in Proposition 6.1 holds. We want to apply Corollary 6.3 with

$$p = x_{w_1}, q = x_{w_2}, x_i = x_{\rho_i} \quad \forall i \in [0, d].$$

From edge count to distance. By the Calibration Lemma 2.6,

$$p(|x_{w_j} x_0|) \geq N_{U_{\rho_0}}(w_j) - 2\eta \geq p(9\delta) - 2\eta \geq p(10\delta),$$

where the last inequality holds when  $c_{6.1}$  is chosen small enough. Consequently, we have  $|x_{w_1} x_0| \leq 10\delta$ , which satisfies the assumptions of Lemma 6.2 (and its Corollary 6.3). Now let us write  $p = x_{w_1}$ ,  $q = x_{w_2}$  and  $x_i = x_{\rho_i}$ .

Notation in  $T_p M$ . For each  $i \in [d]$  and  $v \in U_{\rho_i}$ , let

$$\hat{x}_i := (\hat{x}_i)_p, \quad \hat{x}_v := (\hat{x}_v)_p \text{ for } v \in U_{\rho_i}, \quad \hat{q} := (\hat{x}_{w_2})_p.$$

Define unit directions

$$\vec{y}_i := \frac{\hat{x}_i}{\|\hat{x}_i\|}, \quad \vec{y}_v := \frac{\hat{x}_v}{\|\hat{x}_v\|} \quad (v \in U_{\rho_i}),$$

and set

$$t_i := \langle \vec{y}_i, \hat{q} \rangle_p = \|\hat{q}\| \cos(\angle(p_{x_i}^q)), \quad \mathbf{t} := (t_i)_{i \in [d]}.$$

Equivalence between  $\|\mathbf{t}\|_2^2$  and  $\|\hat{\mathbf{q}}\|^2$ . By Corollary 6.3,

$$(43) \quad \left(1 - \varepsilon_{6.3}\right) \sum_{i=1}^d t_i^2 \leq \|\hat{\mathbf{q}}\|^2 \leq \left(1 + \varepsilon_{6.3}\right) \sum_{i=1}^d t_i^2,$$

where  $\varepsilon_{6.3} = \Theta(c_{6.1})$ .

Split indices by angle. Let

$$I := \left\{ i \in [d] : |\cos(\angle(p_{x_i}^q))| \geq \frac{1}{10\sqrt{d}} \right\}.$$

Then

$$\sum_{i \notin I} t_i^2 = \|\hat{\mathbf{q}}\|^2 \sum_{i \notin I} \cos^2(\angle(p_{x_i}^q)) \leq d \cdot \frac{1}{100d} \|\hat{\mathbf{q}}\|^2 = \frac{1}{100} \|\hat{\mathbf{q}}\|^2,$$

and together with (43) this yields

$$(44) \quad \frac{1}{2} \sum_{i \in I} t_i^2 \leq \|\hat{\mathbf{q}}\|^2 \leq 2 \sum_{i \in I} t_i^2,$$

provided that  $c_{6.1}$  is sufficiently small.

Relating  $b_i$  to  $t_i$  for  $i \in I$ . Fix  $i \in I$ . Because  $|\cos(\angle(p_{x_i}^q))| \geq \frac{1}{10\sqrt{d}}$ , and for all  $v \in U_{\rho_i}$  we have  $|x_i x_v| \leq \eta$  and  $|p x_i| \asymp r$ , the hypotheses of Lemma 3.6 (with  $p, q, x = x_i, x' = x_v$ ) are satisfied. Using its second conclusion,

$$\frac{\ell_p}{2} \leq \frac{p(|q x_v|) - p(|p x_v|)}{\cos(\angle(p_{x_i}^q)) |p q|} \leq 2L_p.$$

Averaging over  $v \in U_{\rho_i}$  and recalling  $t_i = \cos(\angle(p_{x_i}^q)) |p q|$ ,

$$(45) \quad \frac{\ell_p}{2} |t_i| \leq |b_i| \leq 2L_p |t_i|, \quad i \in I.$$

Summing (45) over  $i \in I$  and combining with (44),

$$(46) \quad \frac{1}{4} L_p^{-2} \sum_{i \in I} b_i^2 \leq \|\hat{\mathbf{q}}\|^2 \leq 4 \ell_p^{-2} \sum_{i \in I} b_i^2.$$

Bounding the complement  $i \notin I$ . Fix  $i \notin I$  and  $v \in U_{\rho_i}$ . Applying Corollary A.4 to the triangle  $\Delta p q x_v$ , obtaining

$$\begin{aligned} |p x_v| - |q x_v| &\leq |\cos(\angle(p_{x_v}^q))| \cdot |p q| + 4 \frac{|p q|^2}{|p x_v|} \\ &\leq (|\cos(\angle(p_{x_i}^q))| + \angle(p_{x_v}^{x_i})) \cdot |p q| + 4 \frac{|p q|^2}{|p x_v|} \end{aligned}$$

Using triangle comparison from Lemma 3.1, and then Lemma 3.5 yields

$$\angle(p_{x_v}^{x_i}) \leq \angle^\kappa(p_{x_v}^{x_i}) \leq 2 \frac{|x_i x_v|}{|p x_i|}.$$

Together with  $|\cos(\angle(p_{x_i}^q))| \leq \frac{1}{10\sqrt{d}}$ ,

$$|p x_v| - |q x_v| \leq \left( |\cos(\angle(p_{x_i}^q))| + 2 \frac{|x_i x_v|}{|p x_i|} + 4 \frac{|p q|}{|p x_v|} \right) |p q| \leq \left( \frac{1}{10\sqrt{d}} + \Theta(c_{6.1}) \right) |p q| \leq \frac{1}{5\sqrt{d}} |p q|,$$

under the small  $c_{6.1}$  assumption (using  $|x_i x_v| \leq \eta$ ,  $|p x_i| \gtrsim r$ , and  $|p q| \leq \delta$ ). By Lipschitzness of  $p$ ,

$$|p(|q x_v|) - p(|p x_v|)| \leq L_p \frac{1}{5\sqrt{d}} \|\hat{\mathbf{q}}\|.$$

Averaging over  $v \in U_{\rho_i}$  gives

$$(47) \quad |b_i| \leq L_p \frac{1}{5\sqrt{d}} \|\hat{q}\|, \quad i \notin I, \quad \Rightarrow \quad \sum_{i \notin I} b_i^2 \leq \frac{L_p^2}{25} \|\hat{q}\|^2.$$

Assembling the bounds for the claim. From (46) and (47),

$$\|\hat{q}\|^2 \geq \frac{1}{4} L_p^{-2} \sum_{i \in I} b_i^2 \geq \frac{1}{4} L_p^{-2} \left( \sum_{i=1}^d b_i^2 - \frac{L_p^2}{25} \|\hat{q}\|^2 \right).$$

Rearranging,

$$\left(1 + \frac{1}{100}\right) \|\hat{q}\|^2 \geq \frac{1}{4} L_p^{-2} \sum_{i=1}^d b_i^2 \Rightarrow \|\hat{q}\|^2 \geq \frac{1}{5} L_p^{-2} \sum_{i=1}^d b_i^2.$$

The upper bound follows from (46) and  $\sum_{i \in I} b_i^2 \leq \sum_{i=1}^d b_i^2$ :  $\|\hat{q}\|^2 \leq 4 L_p^{-2} \sum_{i=1}^d b_i^2$ . Therefore, the claim holds.

Comparing  $\sqrt{\sum_{i \in [d]} b_i^2}$ . Within the event  $\mathcal{E}_{\text{navi}}(U_{\rho_i}, w_j)$ , we have

$$|b_i - \underbrace{(\mathbf{N}_{U_{\rho_i}}(w_2) - \mathbf{N}_{U_{\rho_i}}(w_1))}_{:= a_i}| \leq 2\varepsilon_m,$$

Viewing  $\mathbf{b} = (b_i)_{i \in [d]}$  and  $\mathbf{a} = (a_i)_{i \in [d]}$  as vectors in  $\mathbb{R}^d$ , the standard triangle inequality gives

$$|\|\mathbf{b}\|_2 - \|\mathbf{a}\|_2| \leq 2\sqrt{d}\varepsilon_m \Leftrightarrow |\|\mathbf{b}\|_2 - \psi_{\mathbf{F}}(w_1, w_2)| \leq 2\sqrt{d}\varepsilon_m.$$

□

**Algorithm 2: cluster-net**


---

**Input :**  $V_{\text{cn}}, V_{\text{net}}, V_{\text{ortho}} \subseteq \mathbf{V}$  with  $V_{\text{net}} \cap V_{\text{ortho}} = \emptyset$ ,  $|V_{\text{net}}| = |V_{\text{ortho}}| = n$ .  
 Parameters  $r, \delta, \eta$  coupled by the standing choice with  $c_{7.1}$ .  
 Integers  $m, s_{\max}$  (minimum cluster size; max # orthogonal clusters to extract).

**Output:** Clusters  $\{(U_i, u_i)\}_{i \in [k]} \subseteq V_{\text{net}}$  (a  $2\delta$ -net).  
 Clusters  $\{(V_j, v_j)\}_{j \in [s]} \subseteq V_{\text{ortho}}$ .  
 Maps  $\text{ax}_i : [d] \rightarrow [\ell]$  for  $i \in [k]$  (orthogonal frames).

1 Initialization:  $t \leftarrow 1; k \leftarrow 1; \ell \leftarrow 1; (U_1, u_1) \leftarrow \text{Algorithm 1}(V_{\text{cn}}, V_{\text{net}}, \eta)$ ;

2 **while** true **do**

3    $t \leftarrow t + 1$ ;

4   **if**  $\text{ax}_k$  is not defined for all  $s \in [d]$  **then**

5      $s \leftarrow$  the smallest index such that  $\text{ax}_k(s)$  is not defined. Define the candidate set  

$$\mathcal{J}_{k,s} := \left\{ j \in [\ell] : \begin{array}{l} p(r + 0.93\delta) \leq N(u_k, v_j) \leq p(r - 0.93\delta), \\ p(\sqrt{2}r + 0.93\delta) \leq N(v_{\text{ax}_k(\alpha)}, v_j) \leq p(\sqrt{2}r - 0.93\delta) \quad \forall \alpha \in [s-1] \end{array} \right\}.$$

6     **if**  $\mathcal{J}_{k,s} \neq \emptyset$  **then**

7       pick any  $j^* \in \mathcal{J}_{k,s}$ ; set  $\text{ax}_k(s) \leftarrow j^*$ ;

8     **end**

9     **else**

10       Define  

$$W_{k,s} = W'(r, \delta; V_{\text{ortho}} \setminus L_{\leq t-1}(U_k, u_k), (V_{\text{ax}_k(1)}, v_{\text{ax}_k(1)}), \dots, (V_{\text{ax}_k(s-1)}, v_{\text{ax}_k(s-1)}))$$
.  
 from (37).

11        $(V_{\ell+1}, v_{\ell+1}) \leftarrow \text{Algorithm 1}(V_{\text{cn}}, W_{k,s}, \eta)$ ;

12        $\ell \leftarrow \ell + 1$ .

13     **end**

14   **end**

15   **else**

16     Define the candidate set  

$$\mathcal{V}_{k+1} = \left\{ (v, i_v) \in (V_{\text{net}} \setminus L_{\leq t-1}) \times [k] : N_{U_i}(v) \leq p(1.1\delta) \quad \forall i \in [k] \quad \text{and } N_{U_{i_v}}(v) \geq p(1.9\delta) \right\}$$

17     **if**  $\mathcal{V}_{k+1} \neq \emptyset$  **then**

18       pick any  $(u^*, i^*) \in \mathcal{V}_{k+1}$ ;  $u_{k+1} \leftarrow u^*$ ; Define  

$$U^* := \left\{ v \in V_{\text{net}} \setminus L_{\leq t-1} : \psi_{i^*}(u_{k+1}, v) \leq \frac{\ell_p}{6}\eta \text{ and } N_{U_{i^*}}(v) \geq p(3\delta) \right\}$$

19       where  

$$\psi_{i^*}(u_{k+1}, v) := \sqrt{\sum_{s \in [d]} (N_{V_{\text{ax}_{i^*}(s)}}(v) - N_{V_{\text{ax}_{i^*}(s)}}(u_{k+1}))^2}.$$

20        $(U_{k+1}, u_{k+1}) \leftarrow (U^*, u^*)$ ;

21        $k \leftarrow k + 1$ ;

22     **end**

23 **end**

---

## 7. STAGE 1: ALGORITHM **CLUSTER-NET** FOR EXTRACTING CLUSTERS NETS WITH ORTHOGONAL FRAMES

In this section we introduce the algorithm **cluster-net**, the goal of the algorithm is to extracts which form a  $2\delta$ -net of  $M$  where each element in a net are equipped with an orthogonal frame of clusters. The main result of this section is the following:

**Proposition 7.1.** *There exists a constant  $0 < c_{7.1} \leq \min\{c_{5.2}, c_{6.1}\}$  so that the following holds. For any  $0 < c_* \leq c_{7.1}$ ,*

*Set the parameters*

$$r := c_{7.1} r_G \quad \delta = c_{7.1} r, \quad \eta = c_{7.1} \delta,$$

*and*

$$m := \mu_{\min}(c'_{5.2} c_{4.5} \eta^2 / 3) \frac{n}{2}.$$

*Suppose  $V_{\text{cn}}$ ,  $V_{\text{net}}$ , and  $V_{\text{ortho}}$  are disjoint vertex subsets of  $\mathbf{V}$  with*

$$(48) \quad |V_{\text{cn}}| = n_{\text{cn}} := \left\lceil \frac{C_{4.5}^4 \log^2(n)}{\varsigma_n^2 \eta^4} \right\rceil = \Theta(\varsigma_n^{-2} \log^2 n) \quad \text{and} \quad |V_{\text{net}}| = |V_{\text{ortho}}| = n,$$

*Then with high probability, the algorithm **cluster-net** extracts a  $2\delta$ -net  $(U_i, u_i)_{i \in [k]}$  from  $V_{\text{net}}$  and a collection of orthogonal clusters  $(V_j, v_j)_{j \in [\ell]}$  from  $V_{\text{ortho}}$ , such that the following holds:*

- (1) *The set  $\{u_k\}_{k \in [k]}$  form a  $2\delta$ -net of  $M$  with  $\delta$ -separation:*

$$\forall p \in M, \exists i \in [k] \text{ such that } |p X_{u_i}| \leq 2\delta, \quad \text{and} \quad |X_{u_i} X_{u_j}| > \delta, \forall i \neq j \in [k].$$

- (2) *For each  $i \in [k]$  and  $j \in [\ell]$ , the cluster  $(U_i, u_i)$  satisfies the event  $\mathcal{E}_{\text{clu}}(\eta, U_i, u_i)$  with  $|U_i| \geq m$  and, similarly,  $\mathcal{E}_{\text{clu}}(\eta, V_j, v_j)$  hold with  $|V_j| \geq m$ .*

- (3) *For each  $i \in [k]$ , the clusters  $(U_i, u_i), (V_{\text{ax}_i(1)}, v_{\text{ax}_i(1)}), \dots, (V_{\text{ax}_i(d)}, v_{\text{ax}_i(d)})$  satisfies the events*

$$\mathcal{E}_{\perp}(r, \delta, \eta, (u_i, v_{\text{ax}_i(1)}, \dots, v_{\text{ax}_i(d)})).$$

**Remark 7.2.** The choice of  $m$  was arrives from Proposition 5.2. And the choice of  $|V_{\text{cn}}|$  is due to (32) from Proposition 4.5.

**Remark 7.3.** We further remark that the choice of  $c_{7.1}, r, \delta, \eta$  and  $m$  satisfy the assumptions of Propositions 5.2 and 6.1.

Algorithm **cluster-net** Outline. The algorithm attempts to extract clusters in a sequential process. We index the steps by

$$t = 1, 2, \dots, T,$$

where for concreteness we may assume  $T \leq n(d+1)$ . At each step we might extract a new cluster from  $V_{\text{net}}$  or from  $V_{\text{ortho}}$ .

At the end of the process we obtain two sequences of clusters:

$$\{(U_i, u_i)\}_{i \in [k]}, \quad k \leq d, \quad U_i \subseteq V_{\text{net}},$$

and

$$\{(V_j, v_j)\}_{j \in [\ell]}, \quad V_j \subseteq V_{\text{ortho}}.$$

Cluster property. By a cluster  $(U_i, u_i)$  we mean that the cluster event

$$\mathcal{E}_{\text{clu}}(\eta, U_i, u_i)$$

holds; i.e., every  $v \in U_i$  satisfies  $|X_{u_i} X_v| \leq \eta$ . Similarly, for each  $(V_j, v_j)$  we assume  $\mathcal{E}_{\text{clu}}(\eta, V_j, v_j)$  holds.

Orthogonal frames. For each  $(U_i, u_i)$  we construct an index map

$$\text{ax}_i : [d] \longrightarrow [\ell],$$

such that the collection

$$(U_i, u_i), (V_{\alpha x_i(1)}, v_{\alpha x_i(1)}), \dots, (V_{\alpha x_i(d)}, v_{\alpha x_i(d)})$$

forms an orthogonal frame of clusters. Precisely, they satisfy the event

$$\mathcal{E}_\perp(r, \delta, \eta, m; (U_i, u_i), (V_{\alpha x_i(1)}, v_{\alpha x_i(1)}), \dots, (V_{\alpha x_i(d)}, v_{\alpha x_i(d)})).$$

Moreover, the set of centers  $\{u_i\}_{i \in [k]}$  forms a  $2\delta$ -net of  $M$ .

Extraction times and notation. For each  $u_i$  (resp.  $v_j$ ) we denote by  $t(u_i)$  (resp.  $t(v_j)$ ) the step at which it is extracted. We set

$$L_t := \begin{cases} U_i, & \text{if } t(u_i) = t, \\ V_j, & \text{if } t(v_j) = t, \\ \emptyset, & \text{if no cluster is extracted at step } t, \end{cases} \quad \ell_t := \begin{cases} u_i, & \text{if } t(u_i) = t, \\ v_j, & \text{if } t(v_j) = t, \\ \emptyset, & \text{if no cluster is extracted at step } t. \end{cases}$$

Thus  $(L_t, \ell_t)$  is the cluster extracted at step  $t$  (when one exists). We also define the revealed sets

$$L_{\prec t} := \bigcup_{s < t} L_s, \quad L_{\leq t} := \bigcup_{s \leq t} L_s.$$

Edge revealing process. Initially the algorithm only observes the edges connecting  $V_{\text{cn}}$  to  $V_{\text{net}}$  and  $V_{\text{ortho}}$ . Whenever at step  $t$  a cluster  $L_t$  is extracted, we reveal all edges from  $L_t$  to

$$V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\leq t-1}.$$

(Note that edges from  $L_t$  to  $L_{\leq t-1}$  have already been revealed when previous clusters were extracted.) This convention makes independence transparent: *The data generated up to step  $t$  does not rely on edges that have not been revealed yet, so that these data are independent of the edges that will be revealed in the future.* In particular, whenever we reveal the edges from  $L_t$  to  $V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\leq t-1}$ , we expect the fluctuation event

$$\mathcal{E}_{\text{navi}}(L_t, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\leq t-1})$$

to hold. This means that for every  $v \in V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\leq t-1}$ , if  $N_{L_t}(v) \geq p(r_G)$ , then  $p^{-1}(N_{L_t}(v))$  estimate the distance of  $\ell_t$  to  $v$  up to an error  $2\eta$ .

Distance between centers. Given extracted cluster centers  $u_k, v_i, v_j$ , we define

$$(49) \quad n(u_k, v_j) := \begin{cases} N_{U_{u_k}}(v_j), & \text{if } t(u_k) < t(v_j), \\ N_{V_j}(u_k), & \text{if } t(v_j) < t(u_k). \end{cases} \quad \text{and} \quad n(v_i, v_j) := \begin{cases} N_{V_i}(v_j), & \text{if } t(v_i) < t(v_j), \\ N_{V_j}(v_i), & \text{if } t(v_j) < t(v_i). \end{cases}$$

That is,  $n(u_k, v_j)$  is always measured using the earlier cluster's vertex set, so that the corresponding fluctuation event (of the form  $\mathcal{E}_{\text{navi}}(L_{t(\cdot)}, \cdot)$ ) was assumed, but not the later, to avoid justification of independence.

Step  $t$  of the algorithm. At step  $t$ , suppose we have extracted  $(U_i, u_i)$  for  $i \in [k]$  and  $(V_j, v_j)$  for  $j \in [\ell]$ , where  $k = k(t)$  and  $\ell = \ell(t)$  are the number of clusters extracted so far from  $V_{\text{net}}$  and  $V_{\text{ortho}}$ , respectively. If  $\alpha x_k$  is not defined for all  $s \in [d]$ . Then the goal of  $t$ -th step is to define  $\alpha x_k(s)$  where  $s$  is the smallest index such that  $\alpha x_k(s)$  is not yet defined:

- It first checks whether there exists a cluster  $(V_j, v_j)$  already extracted from  $V_{\text{ortho}}$  that is a valid candidate for  $\alpha x_k(s)$ . If so, we set  $\alpha x_k(s) = j$ .
- If not, then we extract a new cluster  $(V_\ell, v_\ell)$  from  $V_{\text{ortho}}$  and set  $\alpha x_k(s) = \ell$ . (This corresponds to Scenario 1.)

On the other hand, if  $\alpha x_k$  is already defined for all  $s \in [d]$ , then the goal of  $t$ -th step is to extract a new cluster from  $V_{\text{net}}$ . This new cluster must be at least  $2\delta$  away from all previously extracted clusters in  $V_{\text{net}}$ , but within  $3\delta$  of some previously extracted cluster in  $V_{\text{net}}$ . (This corresponds to Scenario 2.)

Proof roadmap. We define a sequence of “good” events  $\{\Omega(t)\}_{t \geq 0}$  adapted to the sequential extraction, where  $\Omega(0)$  summarizes the initial conditions, and  $\Omega(t)$  collects the desired properties after step  $t$  (e.g., cluster validity, edge fluctuation bounds on all revealed edges, and the orthogonal frame properties). Write  $\Omega(\preceq t) := \bigcap_{t' \leq t} \Omega(t')$ . We first show that

$$\Pr(\Omega(0)) = 1 - n^{-\omega(1)}.$$

Inductively, assume  $\Omega(\preceq t)$  holds at the beginning of step  $t+1$ . If the algorithm does not terminate at step  $t+1$ , we prove the one-step estimate

$$(50) \quad \Pr(\Omega(t+1) \mid \Omega(\preceq t)) = 1 - n^{-\omega(1)}.$$

If the algorithm does terminate at step  $t+1$ , then  $\Omega(\preceq t)$  will imply that the output satisfies the conclusion of the proposition.

Since in each iteration we either (i) extend a frame coordinate and/or (ii) add a new net cluster, the procedure terminates after at most  $T \leq n(d+1)$  steps. A union bound over the at most  $T$  exposures in (50) yields

$$\Pr(\Omega(\preceq T)) \geq 1 - T \cdot n^{-\omega(1)} = 1 - n^{-\omega(1)},$$

Consequently, with probability  $1 - n^{-\omega(1)}$  either the algorithm terminates early with the stated output, or it maintains  $\Omega(t)$  at every step up to termination, proving the proposition.

Definition of the events  $\Omega(t)$ . Now let us define the events  $\Omega(t)$  more precisely. First, the initial event  $\Omega(0)$  is defined as follows:

$$\Omega(0) := \mathcal{E}_{\text{cn}}(V_{\text{cn}}, V_{\text{net}}) \cap \mathcal{E}_{\text{cn}}(V_{\text{cn}}, V_{\text{ortho}}) \cap \mathcal{E}_{\text{pt}}(V_{\text{net}}) \cap \mathcal{E}_{\text{pt}}(V_{\text{ortho}}).$$

As for  $\Omega(t)$ , it contains two components: The first event concerns about a new extracted cluster

$$(51) \quad \Omega_{\text{clu}}(t) := \begin{cases} \{|L_t| \geq m\} \cap \mathcal{E}_{\text{clu}}(\eta, L_t, \ell_t) \cap \mathcal{E}_{\text{navi}}(L_t, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq t-1}) & \text{if } L_t \neq \emptyset, \\ \text{True} & \text{otherwise.} \end{cases}$$

The second event concerns the orthogonal frame property:

$$(52) \quad \Omega_{\perp}(t) := \begin{cases} \mathcal{E}_{\perp}(r, \delta, \eta, (u_k, v_{\text{ax}_k(1)}, \dots, v_{\text{ax}_k(s)})_{j \in [s]}) & \text{if } \text{ax}_k(s) \text{ is defined at step } t, \\ \text{True} & \text{otherwise.} \end{cases}$$

The third event is a net property that is only checked when the algorithm terminates:

$$(53) \quad \Omega_{\text{net}}(t) := \begin{cases} \{u_i\}_{i \in [k]} \text{ is a } 2\delta\text{-net of } M & \text{if the algorithm terminates at step } t, \\ \text{True} & \text{otherwise.} \end{cases}$$

The  $\Omega(t)$  event is then defined as the intersection of these three events:

$$\Omega(t) := \Omega_{\text{clu}}(t) \cap \Omega_{\perp}(t) \cap \Omega_{\text{net}}(t),$$

Remark on randomness and data structure. It is worth to remark that all  $(U_i, u_i)$ ,  $(V_j, v_j)$ , and  $\text{ax}_i$  are all random. In particular, the revealing order matters. To analyze the algorithm, we will define  $\mathcal{D}_t$  to be a set of possible data that can be generated by the algorithm up to step  $t$ :

**Definition 7.4.** *A data set  $\mathcal{D}_t$  is a collection of tuples*

$$\mathcal{D}_t = \left\{ (U_i^\sharp, u_i^\sharp, t_{u_i^\sharp})_{i \in [k^\sharp]}, (V_j^\sharp, v_j^\sharp, t_{v_j^\sharp})_{j \in [\ell^\sharp]}, \text{ax}_i^\sharp : [d] \rightarrow [\ell^\sharp], i \in [k^\sharp - 1], \text{ax}_{k^\sharp}^\sharp : [s^\sharp] \rightarrow [\ell^\sharp] \right\}.$$

Further, we say that a data set  $\mathcal{D}_t$  is compatible with the Algorithm 2 at step  $t$  if

- (1)  $k(t) = k^\sharp$  and  $\ell(t) = \ell^\sharp$ ;
- (2)  $(U_i, u_i) = (U_i^\sharp, u_i^\sharp)$  and  $t(u_i) = t_{u_i^\sharp}$  for all  $i \in [k^\sharp]$ ;
- (3)  $(V_j, v_j) = (V_j^\sharp, v_j^\sharp)$  and  $t(v_j) = t_{v_j^\sharp}$  for all  $j \in [\ell^\sharp]$ .

(4)  $\text{ax}_i = \text{ax}_i^\sharp$  for all  $i \in [k^\sharp - 1]$  and  $\text{ax}_k = \text{ax}_{k^\sharp}^\sharp$  for all  $i \in [s^\sharp]$ . In particular,  $\text{ax}_k$  is only defined for  $[s^\sharp]$ .

Here,  $k^\sharp$  and  $\ell^\sharp$  indicates the number of clusters extracted from  $V_{\text{net}}$  and  $V_{\text{ortho}}$ , respectively, and  $s^\sharp$  indicates the number of coordinates of the orthogonal frame that have been defined for the last extracted cluster from  $V_{\text{net}}$ . For each  $i \in [k^\sharp]$  and  $j \in [\ell^\sharp]$ , the pair  $(U_i^\sharp, u_i^\sharp)$  (resp.  $(V_j^\sharp, v_j^\sharp)$ ) is a cluster extracted from  $V_{\text{net}}$  (resp.  $V_{\text{ortho}}$ ) at step  $t_{u_i^\sharp}$  (resp.  $t_{v_j^\sharp}$ ).

Basic properties. Before we proceed, let us make a few observations.

**Fact 7.5.** *From the Algorithm 2, whenever  $L_t$  is extracted at step  $t$ , we have*

$$L_{t+1} \subseteq V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq t},$$

whenever  $L_{t+1} \neq \emptyset$ .

Before proceeding, we also make a remark on the running time.

**Remark 7.6.** First, we remark that given  $\delta$  is a constant. The  $\delta$ -separation set of  $M$  has size at most  $O(1)$  (independent of  $n$ ). Therefore, the algorithm **cluster-net** will terminate in  $O(1)$  steps. In each iteration, the most time-consuming step is to run the cluster-finding algorithm (Algorithm 1) which takes time  $O(n^2(\log(n) + s_n^2|V_{\text{cn}}|)) = O(n^2 \log^2 n)$ . Therefore, the total running time of Algorithm 2 is  $O(n^2 \log^2 n)$ .

**7.1. Main Technical Lemma.** This subsection is devoted to prove the following lemma.

**Lemma 7.7.** *For each step  $t \geq 0$ ,*

$$\Pr(\Omega(t+1) \mid \Omega(\preceq t)) = 1 - n^{-\omega(1)}.$$

First of all, if the algorithm is already terminated before or at  $t$ , then  $\Omega(t+1)$  is trivially true. Therefore, we only need to consider the case that the algorithm does not terminate at step  $t$ .

A realization of partial information of the graph. From now on, we assume that the algorithm does not terminate at step  $t$  and we fix a data structure  $\mathcal{D}_t$ , we will first condition on the event  $\Omega(\preceq t)$  and the data set  $\mathcal{D}_t$ . If we run the algorithm with  $t$  iterations, then the edges that have been revealed are

$$E(V_{\text{cn}}, V_{\text{net}} \cup V_{\text{ortho}}) \cup \bigcup_{\alpha \leq t} E(L_\alpha, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq \alpha-1}),$$

which are functions of the independent random variables

$$X := (X_{V_{\text{cn}}}, X_{V_{\text{net}}}, X_{V_{\text{ortho}}}), \quad \mathcal{U}_0 := \{\mathcal{U}_{u,v} : u \in V_{\text{cn}}, v \in V_{\text{net}} \cup V_{\text{ortho}}\}$$

and  $\mathcal{U}_\alpha := \{\mathcal{U}_{u,v} : u \in L_\alpha, v \in V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq \alpha-1}\}$ , for  $\alpha \in [t]$ .

We remark that the sets are  $L_\alpha$  are all determined by  $\mathcal{D}_t$  for  $\alpha \leq t$ . At the same time,  $\Omega(\preceq t)$  is an event of  $X, \mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_t$ . Now, let us fix a realization

$$X = x \quad \text{and} \quad \mathcal{U}_{[0,t]} = \mathcal{U}_{[0,t]}^\sharp,$$

of the random variables so that it is within the event  $\Omega(\preceq t)$ , and with this realization, the data set generated by the algorithm up to step  $t$  is exactly  $\mathcal{D}_t$ .

From the data set  $\mathcal{D}_t$ , we can determine what's the goal of the algorithm at step  $t+1$ . There are two cases: either the algorithm tries to define  $\text{ax}_k(s)$  for some  $s \in [d]$ , which may involve extracting a new cluster from  $V_{\text{ortho}}$ , or it tries to extract a new cluster from  $V_{\text{net}}$ . Let us consider it case by case by following the structure of the algorithm.

Special case:  $t = 0$ . At the first iteration, we simply generate the first cluster  $(U_1, u_1)$  by

$$(U_1, u_1) = \text{Algorithm 1}(V_{\text{net}}, V_{\text{ortho}}, \eta).$$

In this case,

$$\Omega(1) = \Omega_{\text{clu}}(1).$$

From  $\mathcal{E}_{\text{pt}}(V_{\text{net}}) \supseteq \Omega(0)$ , we know that for every  $q \in M$ ,

$$|B_{c_{4.5}\eta^2}(q) \cap \{X_v\}_{v \in V_{\text{net}}} \geq m.$$

Together with  $\mathcal{E}_{\text{cn}}(V_{\text{cn}}, V_{\text{net}}) \supseteq \Omega(0)$ , we can apply Proposition 4.5 to conclude that with probability  $1 - n^{-\omega(1)}$ , the event  $\mathcal{E}_{\text{clu}}(\eta, U_1, u_1)$  holds with  $|U_1| \geq m$ .

So it remains to check the event  $\mathcal{E}_{\text{navi}}(L_1, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_1)$  where  $L_1 = U_1$ . Given that all the edges from  $L_1$  to  $V_{\text{net}} \cup V_{\text{ortho}} \setminus L_1$  have not been revealed yet, we simply apply Lemma 2.2, which gives

$$\Pr(\mathcal{E}_{\text{navi}}(L_1, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_1) \mid X = x, \mathcal{U}_{[0,t]} = \mathcal{U}_{[0,t]}^\sharp) = 1 - n^{-\omega(1)}.$$

Since this probability estimate is true for any realization  $X = x$  and  $\mathcal{U}_{[0,t]} = \mathcal{U}_{[0,t]}^\sharp$  within the event  $\Omega(\preceq t)$ , and also every data set  $\mathcal{D}_t$  compatible with the algorithm at step  $t$  and  $\Omega(0)$ , we have

$$\Pr(\Omega(1) \mid \Omega(0)) = 1 - n^{-\omega(1)}.$$

This concludes the proof of the case  $t = 0$ . From now on, we assume  $t \geq 1$ .

Case 1:  $\text{ax}_k$  is not defined for all  $s \in [d]$ . In this case, the algorithm tries to define  $\text{ax}_k(s)$  where  $s$  is the smallest index such that  $\text{ax}_k(s)$  is not yet defined. The candidate set

$$\mathcal{J}_{k,s} := \left\{ j \in [\ell] : \begin{array}{l} p(r + 0.93\delta) \leq n(u_k, v_j) \leq p(r - 0.93\delta), \\ p(\sqrt{2}r + 0.93\delta) \leq n(v_{\text{ax}_k(\alpha)}, v_j) \leq p(\sqrt{2}r - 0.93\delta) \quad \forall \alpha \in [s-1] \end{array} \right\}$$

is then constructed in the algorithm.

Case 1.1:  $\mathcal{J}_{k,s} \neq \emptyset$ . First, the event  $\Omega_{\text{clu}}(t+1)$  is trivially true since no new cluster will be extracted. Now one needs to verify the event

$$\Omega_{\perp}(t) = \mathcal{E}_{\perp}(r, \delta, \eta, (u_k, v_{f_1(1)}, \dots, v_{\text{ax}_k(s-1)}, v_{\text{ax}_k(s)})).$$

To see this, we first remark that

$$\mathcal{E}_{\perp}(r, \delta, \eta, (u_k, v_{f_1(1)}, \dots, v_{\text{ax}_k(s-1)})) \supseteq \Omega(\preceq t)$$

It remains to verify that

$$|X_{u_k} X_{v_{\text{ax}_k(s)}}| - r \leq \delta \quad \text{and} \quad |X_{v_{\text{ax}_k(\alpha)}} X_{v_{\text{ax}_k(s)}}| - \sqrt{2}r \leq \delta \quad \text{for } \alpha \in [s-1],$$

which follows directly from the definition of  $\mathcal{J}_{k,s}$  and (14) in Calibration Lemma 2.6 provided that  $c_{7.1}$  is sufficiently small. In other words,

$$\mathbb{P}(\Omega(t+1) \mid \mathcal{J}_{k,s} \neq \emptyset, X = x, \mathcal{U}_{[0,t]} = \mathcal{U}_{[0,t]}^\sharp) = 1.$$

Case 1.2:  $\text{ax}_k(s)$  is defined by extracting a new cluster  $(V_{\ell+1}, v_{\ell+1})$  from  $V_{\text{ortho}}$ . In this case, (14) in Calibration Lemma 2.6, we know that for every  $j \in [\ell] \setminus \{\text{ax}_k(1), \dots, \text{ax}_k(s-1)\}$ , and any vertex  $v' \in V_j$  DOES NOT satisfies one of the following condition

$$|X_{u_k} X_{v'}| - r \leq 0.928\delta \quad \text{or} \quad |X_{v_{\text{ax}_k(\alpha)}} X_{v'}| - \sqrt{2}r \leq 0.928\delta \quad \text{for some } \alpha \in [s-1],$$

provided that  $c_{7.1}$  is sufficiently small. Further, the above property also holds for  $v' \in V_{\text{ax}_k(\alpha)}$  for every  $\alpha \in [s-1]$ , since it clearly violates the second condition. Therefore, for any  $q \in M$  such that

$$|X_{u_k} q| \in [r - 0.92\delta, r + 0.92\delta] \quad \text{and} \quad |X_{v_{\text{ax}_k(\alpha)}} q| \in [\sqrt{2}r - 0.92\delta, \sqrt{2}r + 0.92\delta] \quad \text{for } \alpha \in [s-1],$$

every  $q' \in B_{c_{5.2}\eta^2}(q)$  satisfies the above inequality with 0.92 replaced by 0.921, by triangle inequalities. Hence, we conclude that

$$\left( V_{\text{ortho}} \setminus L(\preceq t) \right) \cap B_{c_{5.2}\eta^2}(q) = V_{\text{ortho}} \cap B_{c_{5.2}\eta^2}(q),$$

which in turn implies that

$$\left| \left( V_{\text{ortho}} \setminus L(\preceq t) \right) \cap B_{c_{5.2}\eta^2}(q) \right| \geq m,$$

from  $\mathcal{E}_{\text{pt}}(V_{\text{ortho}}) \supseteq \Omega(0)$  and our choice of  $m$ . In particular, the condition in Proposition 5.2 is satisfied with  $m = m'$ , and so does the event  $\mathcal{E}_\perp(r, \delta, \eta, (u_k, v_{\alpha x_k(1)}, \dots, v_{f_k(s-1)}))$ . Therefore, by Proposition 5.2, Proposition 4.5, and Remark 5.3, the output  $(L_{t+1}, \ell_{t+1}) = (V_{\ell+1}, v_{\ell+1})$  of the subalgorithm Algorithm 1( $V_{\text{cn}}, W', \eta$ ) satisfies the event  $\mathcal{E}_{\text{clu}}(\eta, V_{\ell+1}, v_{\ell+1})$  with  $|V_{\ell+1}| \geq m$  and

$$\mathcal{E}_\perp(r, \delta, \eta, (u_k, v_{\alpha x_k(1)}, \dots, v_{f_k(s-1)}, v_{\ell+1})).$$

In particular, all events in the definition of  $\Omega_{\text{clu}}(t+1)$  and  $\Omega_\perp(t+1)$  hold with the exception of the event  $\mathcal{E}_{\text{navi}}(L_{t+1}, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq t-1})$ . For the latter, we note that all edges from  $L_{t+1}$  to  $V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq t}$  have not been revealed yet. To be precise, the variables  $\mathcal{U}_{u,v}$  for  $u \in L_{t+1}$  and  $v \in V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq t}$  have not been revealed yet. Therefore, by Lemma 2.2, we have

$$\begin{aligned} & \mathbb{P}(\Omega(t+1) \mid \mathcal{J}_{k,s} = \emptyset, X = x, \mathcal{U}_{[0,t]} = \mathcal{U}_{[0,t]}^\sharp) \\ &= \mathbb{P}(\mathcal{E}_{\text{navi}}(L_{t+1}, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq t}) \mid X_{L_{t+1}} = x_{L_{t+1}}, X_{V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq t}} = x_{V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq t}}) \\ &= 1 - n^{-\omega(1)}. \end{aligned}$$

Case 2:  $f_k$  is already defined for all  $s \in [d]$  and the algorithm tries to extract a new cluster from  $V_{\text{net}}$ .  
In this case, the candidate set

$$\mathcal{V}_{k+1} = \left\{ (v, i_v) \in (V_{\text{net}} \setminus L_{\preceq t-1}) \times [k] : N_{U_i}(v) \leq p(1.1\delta) \quad \forall i \in [k] \quad \text{and} \quad N_{U_{i_v}}(v) \geq p(1.9\delta) \right\}$$

is defined. For each  $(v, i_v) \in \mathcal{V}_{k+1}$ , by  $\mathcal{E}_{\text{navi}}(L_\alpha, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\preceq \alpha-1}) \supseteq \Omega(\preceq t)$  for all  $\alpha \leq t$ . Then invoking the annulus calibration (14) in Calibration Lemma 2.6 we have

$$|X_{u_i} X_v| \geq \delta \quad \text{for all } i \in [k], \quad \text{and} \quad |X_{u_{i_v}} X_v| \leq 2\delta,$$

provided that  $c_{7.1}$  is sufficiently small.

Case 2.1:  $\mathcal{V}_{k+1} \neq \emptyset$ . For each  $v \in U^*$ , the Calibration Lemma 2.6 implies that

$$|X_v X_{u_{i^*}}| \leq 3.1\delta \quad \Rightarrow \quad U^* \subseteq (V_{\text{net}} \setminus L_{\preceq t}) \cap B_{3.1\delta}(u_{i^*}).$$

where we use  $B_{3.1\delta}(u_{i^*})$  to denote the set of vertices  $v \in \mathbf{V}$  such that  $|X_v X_{u_{i^*}}| \leq 3.1\delta$ . Now, we want to apply Proposition 6.1 for

$$\begin{aligned} ((U_{\rho_i}, \rho_i))_{i \in [0,d]} &= ((U_{i^*}, u_{i^*}), (V_{\alpha x_{i^*}(1)}, v_{\alpha x_{i^*}(1)}), \dots, (V_{\alpha x_{i^*}(d)}, v_{\alpha x_{i^*}(d)})) \\ w_1 &= u^* \quad \text{and} \quad w_2 \in (V_{\text{net}} \setminus L_{\preceq t}) \cap B_{3.1\delta}(u_{i^*}). \end{aligned}$$

Clearly, the assumed events in Proposition 6.1 are satisfied from  $\Omega(\preceq t)$ . Further, we also have, for each  $v \in (V_{\text{net}} \setminus L_{\preceq t}) \cap B_{4\delta}(X_{u_{i^*}})$ , we have

$$p(N_{U_{i^*}}(v)) \geq p(4\delta + \eta) - \varepsilon_m \geq p(9\delta),$$

satisfying the condition in Proposition 6.1. Therefore, we conclude that each  $v \in U^*$  satisfies

$$|X_{u^*} X_v| \leq \frac{3}{\ell_p} (\psi_{i^*}(u^*, v) + 2\sqrt{d}\varepsilon_m) \leq \frac{3}{\ell_p} \left( \frac{\ell_p}{6} \eta + 2\sqrt{d}\varepsilon_m \right) \leq \eta,$$

provided that  $c_{7.1}$  is sufficiently small. Therefore, the event

$$\mathcal{E}_{\text{clu}}(\eta, U_{k+1}, u_{k+1}) = \mathcal{E}_{\text{clu}}(\eta, U^*, u^*)$$

holds.

Next, consider

$$V_{\text{net}} \setminus L_{\leq t} = V_{\text{net}} \setminus \left( \bigcup_{i \in [k]} U_i \right).$$

For each  $u \in U_i$  for some  $i \in [k]$ , by the event  $\mathcal{E}_{\text{clu}}(\eta, U_i, u_i) \supseteq \Omega(\leq t)$ , we have

$$|X_u X_{u^*}| \geq |X_{u_i} X_{u^*}| - |X_u X_{u_i}| \geq \delta - \eta \geq 0.9\delta.$$

Thus,

$$B_{0.9\delta}(u^*) \cap V_{\text{net}} \subseteq V_{\text{net}} \setminus L_{\leq t}.$$

Fix any  $v \in V_{\text{net}}$  such that

$$|X_v X_{u^*}| \leq \frac{\ell_p}{18L_p} \eta.$$

First,  $v \in V_{\text{net}} \setminus L_{\leq t}$ . Second, by triangle inequalities, we have

$$|X_v X_{u_{i^*}}| \leq |X_{u^*} X_{u_{i^*}}| + \frac{\ell_p}{18L_p} \eta \leq 2.1\delta,$$

and therefore, by the event  $\mathcal{E}_{\text{navi}}(L_{i^*}, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\leq i^*-1}) \supseteq \Omega(\leq t)$  and (14) in Calibration Lemma 2.6, we have

$$\mathbf{p}(N_{U_{i^*}}(v)) \geq \mathbf{p}(3\delta).$$

Third, by Proposition 6.1, we have

$$\psi_{i^*}(u^*, v) \leq 3L_p |X_{u^*} X_v| \leq \frac{\ell_p}{6} \eta.$$

Together we conclude that

$$B_{\ell_p \eta / 18L_p}(u^*) \cap V_{\text{net}} \subseteq U^*,$$

and therefore, by  $\mathcal{E}_{\text{pt}}(V_{\text{net}}) \supseteq \Omega(0)$  and our choice of  $m$ , we have

$$|L_{t+1}| = |U^*| \geq |B_{\ell_p \eta / 18L_p}(u^*) \cap V_{\text{net}}| \geq m,$$

provided that  $c_{7.1}$  is sufficiently small. Therefore, the only missing event in the definition of  $\Omega_{\text{clu}}(t+1)$  and  $\Omega_{\perp}(t+1)$  is the event  $\mathcal{E}_{\text{navi}}(L_{t+1}, V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\leq t-1})$ . As in Case 1.2, all edges from  $L_{t+1}$  to  $V_{\text{net}} \cup V_{\text{ortho}} \setminus L_{\leq t}$  have not been revealed yet. Therefore, the same analysis as in Case 1.2 shows that

$$\mathbb{P}(\Omega(t+1) \mid \mathcal{V}_{k+1} \neq \emptyset, X = x, \mathcal{U}_{[0,t]} = \mathcal{U}_{[0,t]}^\sharp) = 1 - n^{-\omega(1)}.$$

Case 2.2:  $\mathcal{V}_{k+1} = \emptyset$ . Suppose there exists  $v \in V_{\text{net}} \setminus L_{\leq t}$  such that

$$|X_v X_{u_i}| \geq 2\delta \quad \text{for all } i \in [k].$$

Let  $\gamma$  be the geodesic from  $X_v$  to  $X_{u_1}$ . Let  $p$  be the first time on  $\gamma$  such that  $|p X_{u_i}| = 1.5\delta$  for some  $i' \in [0, k]$ . By the event  $\mathcal{E}_{\text{pt}}(V_{\text{net}}) \supseteq \Omega(0)$ , there exists  $v' \in V_{\text{net}}$  such that  $|X_{v'} p| \leq 0.1\delta$ . Therefore, we have

$$|X_v X_{u_i}| \geq 1.4\delta \quad \text{for all } i \in [k], \quad \text{and} \quad |X_v X_{u_{i'}}| \leq 1.6\delta.$$

In particular, from the event  $\mathcal{E}_{\text{clu}}(\eta, U_i, u_i) \supseteq \Omega(\leq t)$  for all  $i \in [k]$ , we have

$$v \in V_{\text{net}} \setminus L_{\leq t}$$

and consequently,  $(v, i') \in \mathcal{V}_{k+1}$ , contradicting the assumption that  $\mathcal{V}_{k+1} = \emptyset$ . Therefore, we conclude that when  $\mathcal{V}_{k+1} = \emptyset$ , there is no  $p \in M$  such that

$$|p X_{u_i}| \geq 2\delta \quad \text{for all } i \in [k].$$

Therefore, the event  $\Omega_{\text{net}}(t+1)$  is true.

In short, we conclude that

$$\mathbb{P}(\Omega(t+1) \mid X = x, \mathcal{U}_{[0,t]} = \mathcal{U}_{[0,t]}^\sharp)$$

for every realization  $x$  and  $\mathcal{U}_{[0,t]}^\sharp$  such that  $\Omega(\preceq t)$  holds and  $\mathcal{D}_t$  is the data set generated by the algorithm up to step  $t$ . Since the above holds for every such realization and  $\mathcal{D}_t$  that is compatible with  $\Omega(\preceq t)$ , we conclude that

$$\mathbb{P}(\Omega(t+1) \mid \Omega(\preceq t)) = 1 - n^{-\omega(1)}.$$

*Proof of Proposition 7.1.* By the Lemma, it is clear that

$$\mathbb{P}(\Omega(\preceq (d+1)n)) = 1 - n^{-\omega(1)}.$$

Within the event  $\Omega(\preceq (d+1)n)$ , the algorithm terminates at step  $t \leq (d+1)n$ . This is because at each step  $t$ , either a new cluster  $(U_i, u_i)$  from  $V_{\text{net}}$  is extracted, or a cluster  $(V_j, v_j)$  is assigned to form an orthogonal pair. There can be at most  $n$  clusters extracted from  $V_{\text{net}}$ , given the size of  $V_{\text{net}}$ , and thus the algorithm must terminate by step  $(d+1)n$ .  $\square$

## 8. STAGE 2: DISTILLING FINE CLUSTERS

**Proposition 8.1.** *There exists  $c_{8.1} > 0$  such that the following holds. Recall the parameter choice*

$$r := c_{7.1} r_G, \quad \delta = c_{7.1} r, \quad \eta = c_{7.1} \delta,$$

and

$$m := \mu_{\min}(c_{4.5} \eta^2 / 3) \frac{n}{2}.$$

Let  $V_{\text{ocn}} \subseteq \mathbf{V}$  with  $|V_{\text{ocn}}| = 2n + n_{\text{cn}}$ , where  $n_{\text{cn}}$  is introduced in Algorithm 2. Let  $\Omega_{\text{ocn}}$  be the there event running partitioning  $V_{\text{ocn}}$  into 3 parts  $V_{\text{cn}}$ ,  $V_{\text{net}}$  and  $V_{\text{ortho}}$  and then running Algorithm 2 on  $V_{\text{cn}} \sqcup V_{\text{net}} \sqcup V_{\text{ortho}}$  with parameters  $r, \delta, \eta, m$  outputs

- A family  $\{(U_i, u_i)\}_{i=1}^k \subseteq V_{\text{ocn}}$  with  $|U_i| \geq m$  and  $\mathcal{E}_{\text{clu}}(\eta, U_i, u_i)$ . Their centers form a  $2\delta$ -net of  $M$ :  $\{X_{u_i}\}_{i=1}^k$  is  $2\delta$  net of  $M$  with  $\delta$ -separation.
- Let  $\{(V_j, v_j)\}_{j=1}^\ell \subseteq V_{\text{ocn}}$  with  $|V_j| \geq m$  and  $\mathcal{E}_{\text{clu}}(\eta, V_j, v_j)$ . For each  $i \in [k]$  there is a map  $\mathbf{ax}_i : [d] \rightarrow [\ell]$  such that the orthogonal-frame event  $\mathcal{E}_\perp(r, \delta, \eta, (u_i, v_{\mathbf{ax}_i(1)}, \dots, v_{\mathbf{ax}_i(d)}))$  holds, namely:

$$\begin{cases} |X_{u_i} X_{v_{\mathbf{ax}_i(\alpha)}}| - r \leq \delta & \forall \alpha \in [d] \\ |X_{v_{\mathbf{ax}_i(\alpha)}} X_{v_{\mathbf{ax}_i(\beta)}}| - \sqrt{2}r \leq \delta & \forall \alpha, \beta \in [d], \alpha \neq \beta \end{cases}$$

Let  $V_{\text{fine}} \subseteq \mathbf{V}$  with  $|V_{\text{fine}}| = n$  be disjoint from  $V_{\text{ocn}}$ . Assume  $\Omega_{\text{ocn}}$ ,  $\mathcal{E}_{\text{navi}}(U_i, V_{\text{fine}})$  and  $\mathcal{E}_{\text{navi}}(V_j, V_{\text{fine}})$  for all  $i \in [k]$ ,  $j \in [\ell]$ , and  $\mathcal{E}_{\text{pt}}(V_{\text{fine}})$ . We set

$$(54) \quad \xi := c_{8.1} \left( \frac{\log n}{s_n n} \right)^{1/(d+2)}.$$

The Algorithm 3, which takes the edges

$$E(U_i, V_{\text{fine}}), \quad E(V_j, V_{\text{fine}}) \quad (i \in [k], j \in [\ell]),$$

as input, returns clusters  $\{(W_w, w)\}_{w \in V_{\text{fine}}}$  with  $W_w \subseteq V_{\text{fine}}$  such that each  $W_w$  satisfies  $\mathcal{E}_{\text{clu}}(\xi, W_w, w)$  and its size satisfies the size calibration margin (12):

$$\varepsilon_{W_w} \leq \ell_p \xi.$$

**Remark 8.2.** If we denote by  $\Omega_{\text{fine}}$  the event that the composition of Algorithm 2 and Algorithm 3 satisfies the properties in Proposition 8.1, then we have

$$\mathbb{P}(\Omega_{\text{fine}}) = 1 - n^{-\omega(1)}.$$

This is simply due to conditional independence. First,  $\Omega_{\text{ocn}}$  holds with probability  $1 - n^{-\omega(1)}$  by Proposition 7.1, and is an event of  $X_{V_{\text{ocn}}}$  and  $\mathcal{U}_{V_{\text{ocn}}, V_{\text{ocn}}}$ . Second,  $\mathcal{E}_{\text{pt}}(V_{\text{fine}})$ , as an event of  $X_{V_{\text{fine}}}$ , is independent of  $\Omega_{\text{ocn}}$  and holds with probability  $1 - n^{-\omega(1)}$  by Lemma 2.9, which implies

$$\Pr(\Omega_{\text{ocn}} \cap \mathcal{E}_{\text{pt}}(V_{\text{fine}})) = 1 - n^{-\omega(1)}.$$

Third, the events  $\mathcal{E}_{\text{navi}}(U_i, V_{\text{fine}})$  and  $\mathcal{E}_{\text{navi}}(V_j, V_{\text{fine}})$  for all  $i \in [k]$  and  $j \in [\ell]$  are all events of  $\mathcal{U}_{U_i, V_{\text{fine}}}$  and  $\mathcal{U}_{V_j, V_{\text{fine}}}$ , which are independent, conditioning of  $X_{V_{\text{ocn}} \cup V_{\text{fine}}}$  and  $\mathcal{U}_{V_{\text{ocn}}, V_{\text{ocn}}}$ . Therefore, by Lemma 2.2 and union bound, we have

$$\Pr\left(\bigcap_{i=1}^k \mathcal{E}_{\text{navi}}(U_i, V_{\text{fine}}) \cap \bigcap_{j=1}^\ell \mathcal{E}_{\text{navi}}(V_j, V_{\text{fine}}) \mid \Omega_{\text{ocn}} \cap \mathcal{E}_{\text{pt}}(V_{\text{fine}})\right) = 1 - n^{-\omega(1)},$$

which implies  $\Omega_{\text{fine}}$  holds with probability  $1 - n^{-\omega(1)}$ .

The algorithm is straightforward. For each  $w \in V_{\text{fine}}$  and choose  $i(w) \in [k]$  such that

$$N_{U_{i(w)}}(w) \geq p(3\delta),$$

and set

$$u^w := u_{i(w)}, \quad U^w := U_{i(w)}, \quad v_\alpha^w := v_{\alpha x_{i(w)}(\alpha)}, \quad V_\alpha^w := V_{\alpha x_{i(w)}(\alpha)} \quad (\alpha \in [d]).$$

Define, for  $v \in V_{\text{fine}}$ ,

$$\psi_w(v) := \sqrt{\sum_{\alpha=1}^d (N_{V_\alpha^w}(v) - N_{V_\alpha^w}(w))^2},$$

and for a threshold  $\lambda$  (to be specified below),

$$W_w := \left\{ v \in V_{\text{fine}} : N_{U^w}(v) \geq p(9\delta) \text{ and } \psi_w(v) \leq \lambda \right\}.$$

We assume throughout that

$$4\sqrt{d}\varepsilon_m \leq \lambda \leq \eta^2,$$

which is feasible for all large  $n$  under our parameter configuration (since  $\eta$  is a fixed small constant and  $\varepsilon_m = o(1)$ ). The precise choice of  $\lambda$  is given in the algorithm and derived in the proof below.

**Algorithm 3: refine-fine:** axis-guided local clustering on  $V_{\text{fine}}$ 


---

**Input :**  $\{(U_i, u_i)\}_{i \in [k]}$ ,  $\{(V_j, v_j)\}_{j \in [\ell]}$ , and axis maps  $\{\alpha_i : [d] \rightarrow [\ell]\}_{i \in [k]}$   
 Fresh set  $V_{\text{fine}} \subseteq \mathbf{V}$  with  $|V_{\text{fine}}| = n$ , disjoint from all clusters.  
 Edges  $E(U_i, V_{\text{fine}})$  and  $E(V_j, V_{\text{fine}})$  for all  $i \in [k]$ ,  $j \in [\ell]$ .  
 Parameters  $r, \delta, \eta$  (standing choice tied to  $c_{7.1}$ ), integer  $m$ .

**Output:** Refined local clusters  $\{(W_w, w)\}_{w \in V_{\text{fine}}}$  with  $W_w \subseteq V_{\text{fine}}$ .

```

1 Initialization:
2  $\lambda \leftarrow \left( \frac{\sqrt{2} (18L_p)^{d/2}}{6\sqrt{c_\mu}} \cdot \frac{\log n}{\sqrt{s_n n}} \right)^{\frac{2}{d+2}}$  ;
3 foreach  $w \in V_{\text{fine}}$  do
4   // Choose reference net cluster via a coarse proximity test
5    $i(w) \leftarrow \arg \max_{i \in [k]} \mathbf{1}\{N_{U_i}(w) \geq p(3\delta)\};$ 
6   if no such  $i$  exists then
7     | return failure (no cluster found around  $w$ );
8   end
9    $U^w \leftarrow U_{i(w)}$ ;  $u^w \leftarrow u_{i(w)}$ ;
10  for  $\alpha \in [d]$  do
11    |  $V_\alpha^w \leftarrow V_{\alpha_{i(w)}}(\alpha)$ ;  $v_\alpha^w \leftarrow v_{\alpha_{i(w)}}(\alpha)$ 
12  end
13  // Define the axis-based difference functional  $\psi_w$ 
14  foreach  $v \in V_{\text{fine}}$  do
15    |  $\psi_w(v) \leftarrow \sqrt{\sum_{\alpha=1}^d (N_{V_\alpha^w}(v) - N_{V_\alpha^w}(w))^2}$ ;
16  end
17  // Select the refined cluster around  $w$ 
18   $W_w \leftarrow \{v \in V_{\text{fine}} : N_{U^w}(v) \geq p(9\delta) \wedge \psi_w(v) \leq \lambda\};$ 
19 end
20 return  $\{(W_w, w)\}_{w \in V_{\text{fine}}}$ .

```

---

**Remark 8.3.** The running time of Algorithm 3 is  $O(n^2)$  where the dominant step is to compute the edges  $N_{U^w}(v)$  and  $N_{V_\alpha^w}(v)$  for all  $w, v \in V_{\text{fine}}$  and  $\alpha \in [d]$ .

*Proof.* Step 1: From  $N_{U^w}(w) \geq p(3\delta)$  to a bound on  $|X_w X_{u^w}|$ . First, we know there exists  $i \in [k]$  such that  $|X_w X_{u_i}| \leq 2\delta$  by the event that  $\{X_{u_i}\}_{i=1}^k$  is a  $2\delta$ -net of  $M$ . Fix such  $i$ , with  $\mathcal{E}_{\text{clu}}(\eta, U_i, u_i)$  and  $\mathcal{E}_{\text{navi}}(U_i, V_{\text{fine}})$ , the Calibration Lemma 2.6 gives

$$N_{U_i}(w) \geq p(2\delta + \eta) - \varepsilon_{U_i} \geq p(3\delta),$$

which guarantees  $i(w)$  is well-defined. Now, with the choice of  $i(w)$ , the Calibration Lemma 2.6 gives

$$|X_w X_{u^w}| \leq 3\delta + 2\eta \leq 4\delta,$$

for  $c_{7.1}$  small enough.

Step 2: A convenient geometric inclusion. For any  $v \in V_{\text{fine}}$  with  $|X_v X_w| \leq \delta$ , the triangle inequality and Step 1 give

$$|X_v X_{u^w}| \leq |X_v X_w| + |X_w X_{u^w}| \leq \delta + 4\delta = 5\delta.$$

By monotonicity of  $p$ , the bracket  $p(|X_v X_{u^w}| + \eta) \leq p_{U^w}(v) \leq p(|X_v X_{u^w}| - \eta)$  (see (15)) and  $\varepsilon_{U^w} \leq \ell_p \eta$ , we obtain

$$N_{U^w}(v) \geq p_{U^w}(v) - \varepsilon_{U^w} \geq p(5\delta + \eta) - \varepsilon_{U^w} \geq p(9\delta).$$

Therefore,

$$(55) \quad B_{5\delta}(w) \cap V_{\text{fine}} \subseteq \{v \in V_{\text{fine}} : N_{U^w}(v) \geq p(9\delta)\}.$$

Step 3: Navigation via the orthogonal frame (Scenario 2).. On  $\Omega_{\text{ocn}}$ , the family  $\{(U^w, u^w), (V_\alpha^w, v_\alpha^w)\}_{\alpha \in [d]}$  is an approximate orthogonal frame. Thus Proposition 6.1 applies (using  $\mathcal{E}_{\text{navi}}(U^w, \{w, v\})$  and  $\mathcal{E}_{\text{navi}}(V_\alpha^w, \{w, v\})$ ), yielding for all  $v \in V_{\text{fine}}$  with  $N_{U^w}(v) \geq p(9\delta)$ ,

$$(56) \quad \frac{\ell_p}{3} |X_w X_v| - 2\sqrt{d} \varepsilon_m \leq \psi_w(v) \leq 3L_p |X_w X_v| + 2\sqrt{d} \varepsilon_m.$$

Step 4: Inner and outer radii for  $W_w$ . Assume  $\lambda \geq 4\sqrt{d} \varepsilon_m$ . Then from the upper bound in (56),

$$|X_w X_v| \leq \frac{\lambda}{6L_p} \implies \psi_w(v) \leq \frac{\lambda}{2} + 2\sqrt{d} \varepsilon_m \leq \lambda,$$

so every  $v \in V_{\text{fine}}$  with  $|X_v X_w| \leq \lambda/(6L_p)$  and  $N_{U^w}(v) \geq p(9\delta)$  lies in  $W_w$ . Since  $\lambda \leq \eta \ll \delta$ , this ball is contained in  $B_{5\delta}(w)$ , hence by (55),

$$(57) \quad B_{\lambda/(6L_p)}(w) \cap V_{\text{fine}} \subseteq W_w.$$

Conversely, if  $|X_v X_w| \geq \frac{6\lambda}{\ell_p}$ , then the lower bound in (56) and  $\lambda \geq 4\sqrt{d} \varepsilon_m$  imply

$$\psi_w(v) \geq 2\lambda - 2\sqrt{d} \varepsilon_m \geq \frac{3}{2}\lambda > \lambda,$$

so such  $v$  cannot belong to  $W_w$ . Therefore,

$$(58) \quad W_w \subseteq B_{6\lambda/\ell_p}(w) \cap V_{\text{fine}}.$$

In particular,  $\mathcal{E}_{\text{clu}}(6\lambda/\ell_p, W_w, w)$  holds.

Step 5: Size lower bound for  $W_w$ . By (57) and the point-count event  $\mathcal{E}_{\text{pt}}(V_{\text{fine}})$ ,

$$(59) \quad |W_w| \geq |B_{\lambda/(6L_p)}(w) \cap V_{\text{fine}}| \geq \mu_{\min}\left(\frac{\lambda}{18L_p}\right) \frac{n}{2}.$$

With  $\lambda/(18L_p) \leq \frac{\eta}{18L_p} \ll r_\mu$ , the standing assumption  $\mu_{\min}(r) \geq c_\mu r^d$  yields

$$|W_w| \geq \frac{c_\mu}{2} \left(\frac{\lambda}{18L_p}\right)^d n.$$

Step 6: Fluctuation level inside  $W_w$  and the balanced choice of  $\lambda$ . Using the definition  $\varepsilon_U = \frac{\log n}{\sqrt{s_n |U|}}$  and (59), we get

$$\varepsilon_{W_w} \leq \frac{\sqrt{2} \log n}{\sqrt{c_\mu s_n n}} \left(\frac{18L_p}{\lambda}\right)^{d/2}.$$

Recall that if we treat  $(W^w, w)$  as a cluster and within the event  $\mathcal{E}_{\text{navi}}(W^w, v)$  for some vertex  $v$ , then

$$|N_{W^w}(v) - p_{W^w}(v)| \leq \frac{\varepsilon_{W^w}}{\ell_p} + \frac{6\lambda}{\ell_p}.$$

So a convenient ‘‘balanced’’ choice of  $\lambda$  is obtained by imposing  $\varepsilon_{W_w} = 6\lambda$ . Define

$$(60) \quad \lambda_{\text{bal}} := \left(\frac{\sqrt{2} (18L_p)^{d/2}}{6\sqrt{c_\mu}} \cdot \frac{\log n}{\sqrt{s_n n}}\right)^{\frac{2}{d+2}}.$$

Finally set

$$(61) \quad \lambda := \min \left\{ \eta, \max \{4\sqrt{d} \varepsilon_m, \lambda_{\text{bal}}\} \right\} = \max \{4\sqrt{d} \varepsilon_m, \lambda_{\text{bal}}\} = \lambda_{\text{bal}},$$

where the last inequality holds for all sufficiently large  $n$  since

$$\lambda_{\text{bal}} = \Theta\left(\left(\frac{\log n}{\sqrt{s_n n}}\right)^{2/(d+2)}\right) \text{ while } \varepsilon_m = \Theta\left(\frac{\log n}{\sqrt{s_n n}}\right).$$

For all sufficiently large  $n$ ,  $\lambda$  satisfies  $4\sqrt{d}\varepsilon_m \leq \lambda \leq \eta$ , hence Steps 4–6 apply and give

$$\mathcal{E}_{\text{clu}}\left(\frac{6\lambda}{\ell_p}, W_w, w\right), \quad |W_w| \geq \mu_{\min}\left(\frac{\lambda}{18L_p}\right) \frac{n}{2}, \quad \varepsilon_{W_w} \leq 6\lambda.$$

Replacing  $\lambda$  by  $\frac{6\lambda}{\ell_p}$  concludes the proof of Proposition 8.1.  $\square$

We conclude this section by stating a byproduct which will be used later estimate the pairwise distances between any two points.

---

**Algorithm 4: refine-fine-net:** greedy extraction of a  $\zeta$ -net on  $V_{\text{fine}}$ 


---

**Input :**  $\{(U_i, u_i)\}_{i \in [k]}, \{(V_j, v_j)\}_{j \in [\ell]}$ , and axis maps  $\{\alpha_i : [d] \rightarrow [\ell]\}_{i \in [k]}$ ;  
 Fresh set  $V_{\text{fine}} \subseteq \mathbf{V}$  with  $|V_{\text{fine}}| = n$ , disjoint from  $V_{\text{ocn}}$ ;  
 Edges  $E(U_i, V_{\text{fine}})$  and  $E(V_j, V_{\text{fine}})$  for all  $i \in [k], j \in [\ell]$ ;  
 Parameters  $r, \delta, \eta, m$  (standing choice tied to [c<sub>7.1</sub>](#));  
 A scale  $\zeta$  with  $C_{8.4}\varepsilon_m \leq \zeta \leq \eta$ .

**Output:** A subset  $S \subseteq V_{\text{fine}}$  such that  $\{X_w : w \in S\}$  is a  $\zeta$ -net of  $M$  with  $\frac{1}{18L_p}\zeta$ -separation (on  $\Omega_{\text{ocn}}$ ).

1 Precomputation (axis setup and local gauges):

2 **foreach**  $w \in V_{\text{fine}}$  **do**

- // Choose a reference net cluster for  $w$
- 3  $i(w) \leftarrow \arg \max_{i \in [k]} \mathbf{1}\{N_{U_i}(w) \geq p(3\delta)\};$
- 4  $U^w \leftarrow U_{i(w)}$ ;  $u^w \leftarrow u_{i(w)}$ ;
- 5 **for**  $\alpha \in [d]$  **do**
- 6   |  $V_\alpha^w \leftarrow V_{\alpha_i(w)}(w)$ ;  $v_\alpha^w \leftarrow v_{\alpha_i(w)}(w)$ ;
- 7 **end**
- 8   // Define the axis-based difference functional  $\psi_w$
- 9   **foreach**  $v \in V_{\text{fine}}$  **do**
- 10     |  $\psi_w(v) \leftarrow \sqrt{\sum_{\alpha=1}^d (N_{V_\alpha^w}(v) - N_{V_\alpha^w}(w))^2}$ ;
- 11 **end**

12 Greedy selection:

13  $\tau \leftarrow 10L_p\zeta$ ; // separation threshold from Lemma 8.4

14  $S \leftarrow \emptyset$ ;  $A \leftarrow V_{\text{fine}}$ ; //  $A$  = uncovered points

15  $t \leftarrow 0$  **while**  $A \neq \emptyset$  **do**

16   |  $t \leftarrow t + 1$ ; pick any  $w_t \in A$ ; // arbitrary choice

17   |  $S \leftarrow S \cup \{w_t\}$ ;

18   | // Mark points covered by  $w_t$  (close in the coarse test & small axis discrepancy)

19   |  $\text{Cover}(w_t) \leftarrow \{v \in A : N_{U^{w_t}}(v) \geq p(9\delta) \wedge \psi_{w_t}(v) < \tau\}$ ;

20   |  $A \leftarrow A \setminus \text{Cover}(w_t)$ ;

21 **end**

22 **return**  $(S, (\text{Cover}(w))_{w \in S})$ .

---

**Lemma 8.4.** Under the same notation and assumptions as in Proposition 8.1, there exists a constant  $C_{8.4} > 0$ , depending on  $L_p, \ell_p$ , and  $d$ , such that the following holds. For any

$$C_{8.4}\varepsilon_m \leq \zeta \leq \eta,$$

on the event  $\Omega_{\text{ocn}}$  and  $\mathcal{E}_{\text{pt}}(V_{\text{fine}})$ , the Algorithm 4 returns a collection  $(\text{Cover}(w), w)_{w \in S}$  for some subset  $S \subseteq V_{\text{fine}}$  such that  $S$  is a  $2\zeta$ -separation. Further, for each  $w \in S$  and  $v \in \text{Cover}(w)$ ,

$$|X_w X_v| \leq C_{8.4} \zeta.$$

*Proof of Lemma 8.4.* As established in the proof of Proposition 8.1, for any  $w, v \in V_{\text{fine}}$  with  $|X_w X_v| \leq \delta$ , we have

$$N_{U^w}(w') \geq p(9\delta), \quad \text{and} \quad \frac{\ell_p}{3} |X_w X_v| - 2\sqrt{d} \varepsilon_m \leq \psi_w(v) \leq 3L_p |X_w X_v| + 2\sqrt{d} \varepsilon_m.$$

We now construct  $S$  via a greedy selection procedure. Start with an arbitrary singleton  $\{w_1\} \subseteq V_{\text{fine}}$ , and iteratively add  $w_{i+1} \in V_{\text{fine}}$  such that for all  $s \in [i]$ ,

$$\psi_{w_s}(w_{i+1}) \geq 10L_p \zeta \quad \text{or} \quad N_{U^{w_s}}(w_{i+1}) < p(9\delta).$$

This construction ensures that for any two distinct points  $w, w' \in S$ ,

$$|X_w X_{w'}| \geq 2\zeta$$

where the last inequality follows from the assumption  $\zeta \geq C\varepsilon_m$ . (In the case when  $N_{U^{w_s}}(w_{i+1}) < p(9\delta)$ , we simply have  $|X_{w_s} X_{w_{i+1}}| \geq 8\delta > \zeta$ .)

On the other hand, for each  $w' \in \text{Cover}(w)$ , we have

$$|X_{w'} X_w| \leq C\zeta,$$

when  $C$  is a sufficiently large constant depending on  $L_p$  and  $\ell_p$ .

Next, suppose that there exists a vertex  $w \in V_{\text{fine}} \setminus S$  such that  $|X_{w'} X_w| \geq \zeta/2$  for all  $w' \in S$ . Then for each  $w' \in S$ , we must have either  $\psi_{w'}(w) \leq \frac{\ell_p \zeta}{3} - 2\sqrt{d} \varepsilon_m$  or  $N_{U^{w'}}(w) \leq p(9\delta)$ , contradicting the maximality of  $S$ . Therefore,  $S$  is a  $\zeta/2$ -net of  $\{X_v\}_{v \in V_{\text{fine}}}$ .

Finally, by the event  $\mathcal{E}_{\text{pt}}(V_{\text{fine}})$ , every point  $q \in M$  lies within distance

$$\left(\frac{\log^2 n}{n}\right)^{1/d} \ll \frac{\zeta}{2}$$

of some  $X_v$ , implying that  $S$  is indeed a  $\zeta$ -net of  $M$ .  $\square$

## 9. FROM CLUSTERS TO DISTANCE ESTIMATES

Setup and objective. From now on assume  $|\mathbf{V}| = 4n$ . Suppose we fix a partition  $\mathbf{V} = V_{\text{ocn}} \dot{\cup} V_{\text{fine}}$  with  $|V_{\text{ocn}}| = 3n$  and  $|V_{\text{fine}}| = n$ . Applying the previous two stages (Algorithm 2 on  $V_{\text{ocn}}$  and the refinement procedure on  $V_{\text{fine}}$ ) yields a family of clusters

$$\{(W_w, w)\}_{w \in V_{\text{fine}}}, \quad W_w \subseteq V_{\text{fine}},$$

such that each  $W_w$  satisfies

$$\mathcal{E}_{\text{clu}}(\xi, W_w, w) \quad \text{and} \quad \varepsilon_{W_w} \leq \ell_p \xi,$$

These clusters allow us to estimate pairwise distances within  $V_{\text{fine}}$  with high accuracy whenever the latent distance lies in the bi-Lipschitz window of  $p$  (via the Calibration Lemma). The goal of this section is to push beyond these two limitations:

- (1) to extend the estimator beyond the bi-Lipschitz regime of  $p$ , by stitching together local distance estimates along paths, and
- (2) to produce a fine distance estimator for every pair  $v, w \in \mathbf{V}$  (rather than just those in  $V_{\text{fine}}$ ).

The resolution of the both issues are rather straightforward. Suppose we have a distance estimator  $d(w, v)$  for all  $w, v \in V_{\text{fine}}$  whose latent distances are within  $r_G$ . Then, we build a weighted graph on  $V_{\text{fine}}$  with edge weights  $d(w, v)$  for all  $w, v \in V_{\text{fine}}$  with  $|X_w X_v| \lesssim r_G$ , and  $\infty$  otherwise. The shortest-path distance on this graph will be a good distance estimator for all pairs  $w, v \in V_{\text{fine}}$ , with no restriction on the latent distance.

For the second issues, we will simply partition  $\mathbf{V}$  into 8 equal parts with each part of size  $n$ . Then we run the above two-stage algorithms on every combination of  $V_{\text{fine}}$  being the union of two parts and  $V_{\text{ochn}}$  being the union of the other six parts. This way, every pair  $v, w \in \mathbf{V}$  will be contained in some version of  $V_{\text{fine}}$ . Overall, we will run the two-stage algorithms  $\binom{8}{2}$  times. The overall procedure is summarized in Algorithm 5.

---

**Algorithm 5: fine distance**


---

**Input :**  $\mathbf{V}$  with  $|\mathbf{V}| = 4n$ .  
   Edges  $E(\mathbf{V}, \mathbf{V})$ .  
   Parameters  $r, \delta, \eta$ , integer  $m$ .

**Output:**  $s$

```

1   $\xi \leftarrow c_{8.1} \left( \frac{\log n}{s_n n} \right)^{1/(d+2)}$ ; 
2  Initialize  $w^{\alpha, \beta} : \mathbf{V} \rightarrow \mathbb{R}_{\geq 0}$  by  $w^{\alpha, \beta}(w, v) \leftarrow \infty$  for all  $w, v \in \mathbf{V}$  and  $\{\alpha, \beta\} \in \binom{[8]}{2}$ ;
3  Partition  $\mathbf{V}$  into 8 equal parts with each part of size  $n$ :  $\mathbf{V}_1, \dots, \mathbf{V}_8$ ;
4  foreach  $\{\alpha, \beta\} \subseteq \binom{[8]}{2}$  do
5     $V_{\text{fine}} \leftarrow \mathbf{V}_\alpha \cup \mathbf{V}_\beta$ ;
6     $V_{\text{ochn}} \leftarrow \mathbf{V} \setminus V_{\text{fine}}$ ;
7    Randomly select disjoint subsets  $V_{\text{cn}}, V_{\text{net}}, V_{\text{ortho}}$  from  $V_{\text{ochn}}$  with  $|V_{\text{cn}}| = n_{\text{cn}}$  (see (48)),  

       $|V_{\text{net}}| = |V_{\text{ortho}}| = n$ ;
8     $(\{(U_i, u_i)\}_{i \in [k]}, \{(V_j, v_j)\}_{j \in [\ell]}, \{\text{ax}_i : [d] \rightarrow [\ell]\}_{i \in [k]}) \leftarrow$   

      cluster-net( $V_{\text{cn}}, V_{\text{net}}, V_{\text{ortho}}, r, \delta, \eta, m, E(V_{\text{ochn}})$ );
9     $\{(W_w^{\alpha, \beta}, w)\}_{w \in V_{\text{fine}}} \leftarrow$   

      refine-fine( $\{(U_i, u_i)\}_{i \in [k]}, \{(V_j, v_j)\}_{j \in [\ell]}, \{\text{ax}_i\}_{i \in [k]}, V_{\text{fine}}, E(U_i, V_{\text{fine}}), E(V_j, V_{\text{fine}}), r, \delta, \eta, m$ );
10    $(S, (\text{Cover}(w))_{w \in S}) \leftarrow$   

      refine-fine-net( $\{(U_i, u_i)\}_{i \in [k]}, \{(V_j, v_j)\}_{j \in [\ell]}, \{\text{ax}_i\}_{i \in [k]}, V_{\text{fine}}, E(U_i, V_{\text{fine}}), E(V_j, V_{\text{fine}}), r, \delta, \eta, m, \zeta =$   

       $\xi$ );
11   foreach  $w \in V_{\text{fine}}$  do
12      $w' \leftarrow$  the unique element in  $S$  such that  $w \in \text{Cover}(w')$ ;
13     foreach  $v \in V_{\text{fine}}$  do
14        $v' \leftarrow$  the unique element in  $S$  such that  $v \in \text{Cover}(v')$ ;
15       if  $w' = v'$  then
16          $w^{\alpha, \beta}(w, v) \leftarrow C_{8.4}\xi$ ;
17       end
18       else if  $N_{W_w^{\alpha, \beta}}(v) \geq p(r_G)$  then
19          $w^{\alpha, \beta}(w, v) \leftarrow p^{-1}(N_{W_{w'}^{\alpha, \beta}}(v)) + 2\xi + 2C_{8.4}\xi$ ;
20       end
21     end
22   end
23 end
24  $w(w, v) \leftarrow \min\{w^{\alpha, \beta}(w, v), w^{\alpha, \beta}(v, w)\}_{\{\alpha, \beta\} \in \binom{[8]}{2}} \quad \forall w, v \in \mathbf{V}$ ;
25  $(d(w, v))_{w, v \in \mathbf{V}} \leftarrow \text{Dijkstra}((w(w, v))_{w, v \in \mathbf{V}})$ ;
26 return  $(d(w, v))_{w, v \in \mathbf{V}}$ ;

```

---

**Remark 9.1.** We remark that the running time of Algorithm 5 is  $O(n^2 \log^2 n)$  where the dominant step is to run Algorithm 3 for  $\binom{8}{2}$  times.

Here we restate Theorem 1.7 and provide its proof.

**Theorem.** *Assuming that  $s_n n = \omega(\log^2 n)$ . The following holds when  $n$  is sufficiently large. Running Algorithm 5 on  $\mathbf{V}$  with  $|\mathbf{V}| = 4n$  and edges  $E(\mathbf{V}, \mathbf{V})$  returns a distance estimator  $(d(w, v))_{w, v \in \mathbf{V}}$*

such that, with probability  $1 - n^{-\omega(1)}$ , for all  $w, v \in \mathbf{V}$ ,

$$|X_w X_v| - d(w, v) \leq C \left( \frac{\log^2 n}{s_n n} \right)^{\frac{1}{d+2}}$$

where  $C$  is a constant depending only on  $d, L_p, \ell_p, c_\mu, r_G$ , and  $\text{diam}(M)$ .

*Proof of Theorem 1.7.* Good event for each run of the two-stage algorithms. Fix  $\{\alpha, \beta\} \in \binom{[8]}{2}$ . Let

$$\Omega_{\text{fine}}^{\alpha, \beta}$$

be the event of  $X_{\mathbf{V}}$  and  $\mathcal{U}_{\mathbf{V} \setminus \mathbf{V}_\alpha \cup \mathbf{V}_\beta, \mathbf{V}}$  (excluding  $\mathcal{U}_{\mathbf{V}_\alpha \cup \mathbf{V}_\beta, \mathbf{V}_\alpha \cup \mathbf{V}_\beta}$ ) that

$$\mathcal{E}_{\text{pt}}(\mathbf{V}_\alpha \cup \mathbf{V}_\beta), \quad \mathcal{E}_{\text{clu}}(\xi, W_w^{\alpha, \beta}, w), \quad \varepsilon_{W_w^{\alpha, \beta}} \leq \ell_p \xi \quad \text{for all } w \in \mathbf{V}_\alpha \cup \mathbf{V}_\beta,$$

where

$$\xi = c_{8.1} \left( \frac{\log n}{s_n n} \right)^{1/(d+2)}$$

was defined in Proposition 8.1. This event holds with probability  $1 - n^{-\omega(1)}$  by Proposition 8.1 and Remark 8.2. Now, fix any realization of  $X_{\mathbf{V}} = x_{\mathbf{V}}$  and  $\mathcal{U}_{\mathbf{V} \setminus \mathbf{V}_\alpha \cup \mathbf{V}_\beta, \mathbf{V}} = u_{\mathbf{V} \setminus \mathbf{V}_\alpha \cup \mathbf{V}_\beta, \mathbf{V}}$  such that  $\Omega_{\text{fine}}^{\alpha, \beta}$ . Given that  $\mathcal{U}_{\mathbf{V}_\alpha \cup \mathbf{V}_\beta, \mathbf{V}_\alpha \cup \mathbf{V}_\beta}$  is independent of  $X_{\mathbf{V}}$  and  $\mathcal{U}_{\mathbf{V} \setminus \mathbf{V}_\alpha \cup \mathbf{V}_\beta, \mathbf{V}}$ , by Lemma 2.2 and union bound, we have

$$\Pr \left( \bigcap_{w \in \mathbf{V}_\alpha \cup \mathbf{V}_\beta} \mathcal{E}_{\text{navi}}(W_w^{\alpha, \beta}, \mathbf{V}_\alpha \cup \mathbf{V}_\beta \setminus W_w^{\alpha, \beta}) \mid X_{\mathbf{V}} = x_{\mathbf{V}}, \mathcal{U}_{\mathbf{V} \setminus \mathbf{V}_\alpha \cup \mathbf{V}_\beta, \mathbf{V}} = u_{\mathbf{V} \setminus \mathbf{V}_\alpha \cup \mathbf{V}_\beta, \mathbf{V}} \right) = 1 - n^{-\omega(1)}.$$

Given this conditional probability holds for all realizations of  $X_{\mathbf{V}}$  and  $\mathcal{U}_{\mathbf{V} \setminus \mathbf{V}_\alpha \cup \mathbf{V}_\beta, \mathbf{V}}$  such that  $\Omega_{\text{fine}}^{\alpha, \beta}$ , we have

$$\Omega_{\text{dist}}^{\alpha, \beta} := \Omega_{\text{fine}}^{\alpha, \beta} \cap \bigcap_{w \in \mathbf{V}_\alpha \cup \mathbf{V}_\beta} \mathcal{E}_{\text{navi}}(W_w^{\alpha, \beta}, \mathbf{V}_\alpha \cup \mathbf{V}_\beta \setminus W_w^{\alpha, \beta})$$

holds with probability  $1 - n^{-\omega(1)}$ .

Distance estimates with  $w^{\alpha, \beta}$ . Fix any realization of the graph such that  $\Omega_{\text{dist}}^{\alpha, \beta}$  holds. From  $\mathcal{E}_{\text{clu}}(\xi, W_w^{\alpha, \beta}, w)$ ,  $\varepsilon_{W_w^{\alpha, \beta}} \leq \ell_p \xi$ , and  $\mathcal{E}_{\text{navi}}(W_w^{\alpha, \beta}, \mathbf{V}_\alpha \cup \mathbf{V}_\beta \setminus W_w^{\alpha, \beta})$ , the Calibration Lemma 2.6 gives, for all  $w, v \in \mathbf{V}_\alpha \cup \mathbf{V}_\beta$  with  $v \notin W_w^{\alpha, \beta}$  and  $N_{W_w^{\alpha, \beta}}(v) \geq p(r_G)$ ,

$$|p^{-1}(N_{W_w^{\alpha, \beta}}(v)) - |X_w X_v|| \leq 2\xi.$$

In the case  $v \in W_w^{\alpha, \beta}$ , we have  $|X_w X_v| \leq \xi$  by  $\mathcal{E}_{\text{clu}}(\xi, W_w^{\alpha, \beta}, w)$ . Therefore, from our definition

$$w^{\alpha, \beta}(w, v) = \begin{cases} \xi, & v \in W_w^{\alpha, \beta}, \\ p^{-1}(N_{W_w^{\alpha, \beta}}(v)) + 2\xi, & v \notin W_w^{\alpha, \beta}, N_{W_w^{\alpha, \beta}}(v) \geq p(r_G), \\ +\infty, & \text{otherwise,} \end{cases}$$

we have the implication

$$(62) \quad w^{\alpha, \beta}(w, v) < +\infty \implies w^{\alpha, \beta}(w, v) \geq |X_w X_v| \geq w^{\alpha, \beta}(w, v) - 4\xi.$$

Small distance guarantees for  $w^{\alpha, \beta}$ . Further, whenever  $|X_w X_v| \leq r_G/2$ , if  $v \notin W_w^{\alpha, \beta}$ , then our Calibration Lemma 2.6 gives

$$N_{W_w^{\alpha, \beta}}(v) \geq p(|X_w X_v| + 2\eta) \geq p(r_G) \quad (\text{since } \eta \ll r_G).$$

Therefore,

$$(63) \quad |X_w X_v| \leq r_G/2 \implies w^{\alpha, \beta}(w, v) < +\infty.$$

Conclusion for  $w(w, v)$ . Now, we assume the good event

$$\Omega_{\text{dist}} := \bigcap_{\{\alpha, \beta\} \in \binom{[8]}{2}} \Omega_{\text{dist}}^{\alpha, \beta}$$

holds, which happens with probability  $1 - n^{-\omega(1)}$  by union bound. Then, from (62), the definition

$$w(w, v) := \min\{w^{\alpha, \beta}(w, v), w^{\alpha, \beta}(v, w)\}_{\{\alpha, \beta\} \in \binom{[8]}{2}} \quad \forall w, v \in \mathbf{V},$$

implies that

$$(64) \quad w(w, v) < +\infty \implies w(w, v) \geq |X_w X_v| \geq w(w, v) - 4\xi.$$

Furthermore, whenever  $w, v \in \mathbf{V}$  satisfy  $|X_w X_v| \leq r_G/2$ , there exists  $\{\alpha, \beta\} \in \binom{[8]}{2}$  such that  $w, v \in \mathbf{V}_\alpha \cup \mathbf{V}_\beta$ . Then, from (63), we have

$$(65) \quad |X_w X_v| \leq r_G/2 \implies w(w, v) < +\infty.$$

Final distance estimator  $d(w, v)$ . Finally, we define the distance estimator  $d(w, v)$  for all  $w, v \in \mathbf{V}$  as the shortest-path distance with respect to the edge weights  $w(w, v)$ . Fix any  $w, v \in \mathbf{V}$ . Let  $\gamma$  be a shortest geodesic from  $X_w$  to  $X_v$  in  $M$  and let  $t = \lceil \frac{|X_w X_v|}{r_G} \rceil$ . Let  $p_0 = X_w, p_1, \dots, p_t = X_v$  be equidistant points on  $\gamma$  such that  $|p_{i-1} p_i| = |X_w X_v|/t \leq r_G/2$  for all  $i = 1, \dots, t$ . By the event  $\mathcal{E}_{\text{pt}}(\mathbf{V})$ , there exist  $w = u_0, u_1, \dots, u_t = v$  in  $\mathbf{V}$  such that

$$|X_{u_i} p_i| \leq \left( \frac{\log^2(n)}{n} \right)^{1/d} = o(\xi)$$

for all  $i = 1, \dots, t-1$ . Then, by triangle inequality,

$$\begin{aligned} |X_w X_v| &= \sum_{i=1}^t |p_{i-1} p_i| \geq \sum_{i=1}^t \left( |X_{u_{i-1}} X_{u_i}| - 2 \left( \frac{\log^2(n)}{n} \right)^{1/d} \right) \\ &\geq \sum_{i=1}^t \left( w(u_{i-1}, u_i) - 5\xi \right) \geq d(w, v) - \frac{\text{diam}(M)}{r_G/2} \cdot 5\xi. \end{aligned}$$

On the other hand, let  $w = z_0, z_1, \dots, z_s = v$  in  $\mathbf{V}$  be the path that attains the shortest-path distance  $d(w, v)$ . Then by (65), we have

$$|X_w X_v| \leq \sum_{i=1}^s |X_{z_{i-1}} X_{z_i}| \leq \sum_{i=1}^s w(z_{i-1}, z_i) = d(w, v).$$

Together we conclude that under the event  $\Omega_{\text{dist}}$ , for all  $w, v \in \mathbf{V}$ ,

$$|X_w X_v| \leq d(w, v) \leq |X_w X_v| + 10 \frac{\text{diam}(M)}{r_G} \xi.$$

□

## REFERENCES

- [ADC19] Ernesto Araya and Yohann De Castro. Latent distance estimation for random geometric graphs. *Advances in Neural Information Processing Systems*, 32, 2019.
- [AKP24] Stephanie Alexander, Vitali Kapovitch, and Anton Petrunin. *Alexandrov geometry—foundations*, volume 236 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, [2024] ©2024.
- [BBN20] Matthew Brennan, Guy Bresler, and Dheeraj Nagaraj. Phase transitions for detecting latent geometry in random graphs. *Probability Theory and Related Fields*, 178:1215–1289, 2020.
- [BDER16] Sébastien Bubeck, Jian Ding, Ronen Eldan, and Miklós Z Rácz. Testing for high-dimensional geometry in random graphs. *Random Structures & Algorithms*, 49(3):503–532, 2016.
- [Bol11] Béla Bollobás. Random graphs. In *Modern graph theory*, pages 215–252. Springer, 2011.
- [CE08] Jeff Cheeger and David G. Ebin. *Comparison theorems in Riemannian geometry*. AMS Chelsea Publishing, Providence, RI, 2008. Revised reprint of the 1975 original.
- [Che70] J. Cheeger. Finiteness theorems for Riemannian manifolds. *Am. J. Math.*, 92:61–74, 1970.
- [DDC23] Quentin Duchemin and Yohann De Castro. Random geometric graph: some recent developments and perspectives. In *High dimensional probability IX—the ethereal volume*, volume 80 of *Progr. Probab.*, pages 347–392. Birkhäuser/Springer, Cham, 2023.
- [DG16] Carl P Dettmann and Orestis Georgiou. Random geometric graphs with general connection functions. *Physical Review E*, 93(3):032313, 2016.
- [EMP22] Ronen Eldan, Dan Mikulincer, and Hester Pieters. Community detection and percolation of information in a geometric setting. *Combin. Probab. Comput.*, 31(6):1048–1069, 2022.
- [FK15] Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2015.
- [HJM24] Han Huang, Pakawut Jiradilok, and Elchanan Mossel. Reconstructing the geometry of random geometric graphs. In *The Thirty Seventh Annual Conference on Learning Theory*, pages 2519–2521. PMLR, 2024.
- [JLR11] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*. John Wiley & Sons, 2011.
- [Lee18] John M. Lee. *Introduction to Riemannian manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer, Cham, 2018.
- [LMSY22] Siqi Liu, Sidhanth Mohanty, Tselil Schramm, and Elizabeth Yang. Testing thresholds for high-dimensional sparse random geometric graphs. In *STOC ’22—Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 672–677. ACM, New York, 2022.
- [Pen03] Mathew Penrose. *Random geometric graphs*, volume 5 of *Oxford Studies in Probability*. Oxford University Press, Oxford, 2003.
- [Pet16] Peter Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*. Springer, Cham, third edition, 2016.

## APPENDIX A. AUXILIARY GEOMETRIC RESULTS

This section provides some auxiliary results from Riemannian geometry. We start with a standard proposition about the exponential map.

**Proposition A.1** ([CE08, Corollary 1.9]). *Let  $B_r(\vec{0}) \subseteq T_p M$  be the open ball of radius  $r$  on which the exponential map  $\exp_p$  is an embedding. Then*

- (1) *For every  $v \in B_r(\vec{0})$ , the curve*

$$\gamma_v : [0, 1] \longrightarrow M, \quad \gamma_v(t) = \exp_p(tv),$$

*is the unique curve satisfying*

$$L[\gamma_v] = \rho(p, \exp_p(v)) = \|v\|$$

*such that  $\gamma_v(0) = p$  and  $\gamma_v(1) = \exp_p(v)$ . In particular, if a curve  $c$  satisfies  $L[c] = |c(0)c(1)|$ , then—up to reparameterization— $c$  is a smooth geodesic.*

- (2) *If  $q \notin \exp_p(B_r(\vec{0})) = B_r(p)$ , then there exists  $q' \in \partial B_r(p)$  (the boundary of  $B_r(p)$ ) such that*

$$|pq| = r + |q'q|.$$

*In particular,  $|pq| \geq r$ .*

A.0.1. *Proof of Lemma 3.1.* The lemma follows quickly from two standard corollaries of the Rauch comparison theorem (Theorem A.2 and Theorem A.3 below). While the lemma is standard, we did not locate a reference that states it in the form we need, so we provide a proof for completeness.

**Theorem A.2** ([CE08, Corollary 1.35 (“Corollary of Rauch I”)]). *Let  $M_1, M_2$  be Riemannian manifolds with  $\dim(M_2) \geq \dim(M_1)$ . Let  $p_1 \in M_1$  and  $p_2 \in M_2$ . Assume  $K_{M_2} \geq K_{M_1}$ ; i.e., for every plane section  $\sigma_1$  in  $M_1$  and for every plane section  $\sigma_2$  in  $M_2$ , we have  $K(\sigma_2) \geq K(\sigma_1)$ . Let  $r > 0$  be such that  $\exp_{p_1}|_{B_r(\vec{0})}$  is an embedding and  $\exp_{p_2}|_{B_r(\vec{0})}$  is nonsingular. Let  $I : T_{p_1}M_1 \rightarrow T_{p_2}M_2$  be a linear injection preserving inner products. Then for any curve  $\gamma : [0, 1] \rightarrow \exp_{p_1}(B_r(\vec{0}))$ , we have*

$$L[\gamma] \geq L[\exp_{p_2} \circ I \circ \exp_{p_1}^{-1}(\gamma)] = L[\gamma_2(t)],$$

*where  $\gamma_2(t) = \exp_{p_2} \circ I \circ \exp_{p_1}^{-1}(\gamma(t))$  and  $L[\cdot]$  denotes the length of a curve in the corresponding manifold.*

**Theorem A.3** ([Lee18, Theorem 11.12]). *Suppose  $(M, g)$  is a Riemannian  $n$ -manifold whose sectional curvatures are all bounded above by a constant  $\kappa > 0$ . Then there is no conjugate point along any geodesic segment shorter than  $\varpi_\kappa$ .*

*Proof of Lemma 3.1.* Given  $r \leq r_{\text{inj}}(M)$ , the two tangent vectors

$$\hat{x} = \exp_p^{-1}(x), \quad \hat{y} = \exp_p^{-1}(y) \in T_p M$$

are unique and satisfy

$$\|\hat{x}\| = |px|, \quad \|\hat{y}\| = |py|.$$

Positive-curvature comparison. Recall that  $M_\kappa^d$  is the simply-connected  $d$ -dimensional space of constant curvature  $\kappa > 0$ . Pick any  $p_2 \in M_\kappa^d$  and choose a linear isometry

$$I : T_p M \longrightarrow T_{p_2} M_\kappa^d$$

and set

$$x_2 = \exp_{p_2}(I\hat{x}), \quad y_2 = \exp_{p_2}(I\hat{y}).$$

This is well-defined because  $I$  is a linear isometry, and the injectivity radius of  $M_\kappa^d$  is  $r_{\text{inj}}(M_\kappa^d) \geq \varpi_\kappa/2 = \pi/(2\sqrt{\kappa})$ . Following our notations, we have

$$\hat{x}_2 = (\hat{x}_2)_p = I\hat{x} \quad \text{and} \quad \hat{y}_2 = (\hat{y}_2)_p = I\hat{y}.$$

We have

$$\angle^\kappa(p_2 \hat{x}_2) = \angle(\hat{x}_2, \hat{y}_2) = \angle(\hat{x}, \hat{y}) := \theta,$$

and

$$|p_2 x_2| = \|\hat{x}_2\| = \|\hat{x}\| = |p x| := a, \quad |p_2 y_2| = \|\hat{y}_2\| = \|\hat{y}\| = |p y| := b.$$

Let  $\gamma$  be the distance-minimizing geodesic from  $x$  to  $y$  in  $M$ .

Claim:  $\gamma \subseteq B_r(p)$ . Assume for the sake of contradiction that the minimizing geodesic  $\gamma$  from  $x$  to  $y$  leaves  $B_r(p)$ . Let  $z \in \gamma$  be the first point with  $|p z| = r$ . Since  $z$  lies on the shortest geodesic  $\gamma$  connecting  $x$  and  $y$ , we have

$$|x y| = |x z| + |z y|.$$

By the triangle inequality,

$$|x z| \geq |p z| - |p x| = r - a, \quad |z y| \geq |p z| - |p y| = r - b,$$

whence

$$|x y| \geq (r - a) + (r - b) > r - \frac{r}{2} + r - \frac{r}{2} = r.$$

But another application of the triangle inequality gives

$$|x y| \leq d(x, p) + |p y| = a + b < \frac{r}{2} + \frac{r}{2} = r,$$

a contradiction. Therefore  $\gamma$  must be contained in  $B_r(p)$ .

Applying the Rauch comparison theorem. Because  $r \leq r_{\text{inj}}(M)$ , the restriction  $\exp_p|_{B_r(\vec{0})}$  is a diffeomorphism. Further, with  $r < \varpi_\kappa/2$ , we also have  $\exp_{p_2}|_{B_r(\vec{0})}$  is a diffeomorphism.

Let  $\gamma_2 = \exp_{p_2} \circ I \circ \exp_p^{-1}(\gamma)$  be the image of  $\gamma$  under the map  $I$ . Applying Theorem A.2 yields

$$|x y| = L[\gamma] \geq L[\gamma_2] \geq |x_2 y_2|.$$

Because  $a + b < \varpi_\kappa$ , the geodesic triangle  $\triangle p_2 x_2 y_2$  exists in  $M_\kappa^d$ , so

$$|x_2 y_2| = \gamma^\kappa(\theta; a, b).$$

Consequently,

$$|x y| \geq \gamma^\kappa(\theta; a, b),$$

which is exactly the first inequality claimed in (21).

Negative curvature comparison. To obtain the upper bound on  $|x y|$  we reverse the roles of  $M$  and the model space when we apply Theorem A.2.

Let  $M_1 := M_{-\kappa}^d$  and set  $M_2 := M$ . Choose any point  $p_1 \in M_1$  and a linear isometry  $I : T_{p_1} M_1 \rightarrow T_p M$ . Because  $I$  is invertible, put

$$x_1 = \exp_{p_1}(I^{-1}\hat{x}), \quad y_1 = \exp_{p_1}(I^{-1}\hat{y}).$$

and set

$$\theta := \angle^{\kappa_0}(p_1 \hat{x}_1).$$

Similarly to the previous case, we have

$$|p_1 x_1| = \|\hat{x}_1\| = \|\hat{x}\| = |p x| := a, \quad |p_1 y_1| = \|\hat{y}_1\| = \|\hat{y}\| = |p y| := b.$$

Let  $\gamma_1$  be the minimizing geodesic from  $x_1$  to  $y_1$  in  $M_1$ . By the same argument used earlier (triangle inequalities and the choice of  $r$ ),  $\gamma_1 \subseteq B_r(p_1)$ .

For the model space with negative curvature,  $\exp_{p_1}$  is a diffeomorphism without restrictions; and  $\exp_p|_{B_r(\vec{0})}$  for  $M$  is a diffeomorphism as well, hence we may apply Theorem A.2 to get

$$|x_1 y_1| = L[\gamma_1] \geq L[\gamma] \geq |x y|.$$

Since the comparison triangle  $\triangle p_1 x_1 y_1$  exists (indeed, hyperbolic space admits geodesic triangles of any perimeter), we have

$$|x_1 y_1| = \gamma^{\kappa_0}(\theta; a, b),$$

and therefore

$$|x y| \leq \gamma^{\kappa_0}(\theta; a, b),$$

which is the second inequality in (21).

From opposite-side to angle comparison. Because  $a + b < \varpi_\kappa$ , the opposite-side function

$$\alpha \mapsto \gamma^\kappa(\alpha; a, b)$$

is smooth and strictly increasing on  $(0, \pi)$ , with boundary values

$$\gamma^\kappa(0; a, b) = |b - a|, \quad \gamma^\kappa(\pi; a, b) = a + b.$$

From triangle inequality also have  $|b - a| < |x y| < a + b$ , so by strict monotonicity there is a unique  $\alpha_1 \in (0, \pi)$  satisfying

$$\gamma^\kappa(\alpha_1; a, b) = |x y|.$$

By construction this  $\alpha_1$  is the model-space angle:

$$\alpha_1 = \angle^\kappa(p_y^x).$$

Because  $\gamma^\kappa(\cdot; a, b)$  is increasing and  $\gamma^\kappa(\alpha_1; a, b) = |x y| \geq \gamma^\kappa(\angle(p_y^x); a, b)$ , we must have

$$\angle^\kappa(p_y^x) = \alpha_1 \geq \angle(p_y^x),$$

which yields the second inequality in (22).

The first inequality is obtained in exactly the same way, replacing the positive curvature model space  $M_\kappa$  by the lower-curvature model space  $M_{-\kappa}$  and using the inequality in the opposite direction.  $\square$

#### A.0.2. Quantitative estimates of opposite side length with 1 small adjacent side.

*Proof of Lemma 3.3.* First, we apply the triangle comparison lemma (Lemma 3.1) to the triangle  $\triangle p x q$  in  $M$ . Comparing simultaneously with the positively curved model space  $M_\kappa$  and the negatively curved model space  $M_{-\kappa}$  yields

$$\gamma^\kappa(\angle(p_q^x); |p x|, |p q|) \leq |q x| \leq \gamma^{-\kappa}(\angle(p_q^x); |p x|, |p q|).$$

The rest of the proof follows by estimating the opposite side lengths in the two model spaces, which was stated in Lemma A.5 and Lemma A.6.  $\square$

**Corollary A.4.** *With the same assumptions as in Lemma 3.3, we have*

$$|a - c - b \cos(\theta)| \leq 4 \frac{b^2}{a}.$$

*Proof.* Given the assumptions on  $a$  and  $b$ ,

$$\frac{\kappa b^2}{3} = \frac{\kappa a^2}{3} \frac{b^2}{a^2} \leq \frac{1}{16 \cdot 3} \frac{1}{16} \frac{b}{a}.$$

Together with  $|\cos(\theta)| \leq 1$ , the estimate follows immediately.  $\square$

**Lemma A.5.** *Fix  $\kappa > 0$  and consider a geodesic triangle in  $M_\kappa$  whose two adjacent sides have lengths  $a$  and  $b$ . Assume*

$$0 < a, b < \frac{1}{2} \varpi_\kappa \quad \text{and} \quad b \leq \frac{1}{4} a.$$

Let  $\theta \in [0, \pi]$  be the angle between the two sides of lengths  $a$  and  $b$  at the vertex opposite to the side of length  $c = \gamma^\kappa(\theta; a, b)$ . Then

$$a - c - b \cos \theta \leq \pi \frac{b}{a} \cdot b |\cos \theta|.$$

*Proof.* Throughout, write

$$c := \gamma^\kappa(\theta; a, b),$$

where  $\gamma^\kappa(\theta; a, b)$  denotes—by definition—the length of the side opposite the angle  $\theta$  in a geodesic triangle of  $M_\kappa$  whose adjacent sides have lengths  $a$  and  $b$ .

Normalized variables and cosine law. Set

$$\tilde{a} := \sqrt{\kappa} a, \quad \tilde{b} := \sqrt{\kappa} b, \quad \tilde{c} := \sqrt{\kappa} c.$$

From the spherical law of cosines in  $M_\kappa$  (see (23)), we have  $\tilde{a}, \tilde{b}, \tilde{c}$  and  $\theta$  satisfies spherical law of cosines on the unit sphere  $M_1$ :

$$(66) \quad \cos \tilde{c} = \cos \tilde{a} \cos \tilde{b} + \sin \tilde{a} \sin \tilde{b} \cos \theta.$$

Define

$$\Delta := \cos \tilde{a} (\cos \tilde{b} - 1) + \sin \tilde{a} \sin \tilde{b} \cos \theta,$$

so that (66) is  $\cos \tilde{c} = \cos \tilde{a} + \Delta$ . By the fundamental theorem of calculus,

$$(67) \quad \tilde{c} = \arccos(\cos \tilde{a} + \Delta) = \tilde{a} - \int_{\cos \tilde{a}}^{\cos \tilde{a} + \Delta} \frac{1}{\sqrt{1-s^2}} ds.$$

Case  $\theta \in [0, \pi/2]$ . Here  $\cos \theta \geq 0$ .

If  $\Delta \leq 0$ , then  $\cos \tilde{c} \leq \cos \tilde{a}$  and hence  $\tilde{c} \geq \tilde{a}$ , whence  $\tilde{a} - \tilde{c} - \tilde{b} \cos \theta \leq 0$ , which already implies the desired bound. Now we assume  $\Delta > 0$ . For  $s \in [\cos \tilde{a}, \cos \tilde{c}] \subset (0, 1)$  we have

$$\frac{1}{\sqrt{1-s^2}} \leq \frac{1}{\sin \tilde{c}} \leq \frac{1}{\sin(\tilde{a} - \tilde{b})},$$

because  $\tilde{c} \geq \tilde{a} - \tilde{b}$  and  $\sin$  is increasing on  $[0, \pi/2]$ . Since  $\cos t$  decreases on  $[0, \pi/2]$ ,

$$\sin \tilde{a} - \sin(\tilde{a} - \tilde{b}) = \int_{\tilde{a}-\tilde{b}}^{\tilde{a}} \cos t dt \leq \int_0^{\tilde{b}} \cos t dt = \sin \tilde{b},$$

so  $\sin(\tilde{a} - \tilde{b}) \geq \sin \tilde{a} - \sin \tilde{b}$ . Combining these bounds yields

$$\tilde{c} \geq \tilde{a} - \frac{\Delta}{\sin \tilde{a} - \sin \tilde{b}} \geq \tilde{a} - \frac{\sin \tilde{a}}{\sin \tilde{a} - \sin \tilde{b}} \sin \tilde{b} \cos \theta.$$

Because  $0 < \tilde{b} \leq \tilde{a}/4 < \pi/8$ ,  $\sin \tilde{b}/\sin \tilde{a} \leq \frac{1}{2}$ , and for  $t \in [0, 1/2)$  we have  $(1-t)^{-1} \leq 1+2t$ . Hence

$$\frac{\sin \tilde{a}}{\sin \tilde{a} - \sin \tilde{b}} \leq 1 + 2 \frac{\sin \tilde{b}}{\sin \tilde{a}},$$

so that

$$\tilde{c} \geq \tilde{a} - \sin \tilde{b} \cos \theta - 2 \frac{\sin^2 \tilde{b}}{\sin \tilde{a}} \cos \theta \geq \tilde{a} - \tilde{b} \cos \theta - \pi \frac{\tilde{b}^2}{\tilde{a}} \cos \theta,$$

where we used  $\sin \tilde{b} \leq \tilde{b}$  and  $\sin \tilde{a} \geq \frac{\pi}{2} \tilde{a}$ . Rewriting, we obtain

$$\tilde{a} - \tilde{c} - \tilde{b} \cos \theta \leq \pi \frac{\tilde{b}}{\tilde{a}} \tilde{b} |\cos \theta| \iff a - c - b \cos \theta \leq \pi \frac{b}{a} b |\cos \theta|.$$

Case  $\theta \in (\pi/2, \pi]$ . Now  $\cos \theta \leq 0$ . Note that

$$\Delta = \underbrace{\cos \tilde{a}(\cos \tilde{b} - 1)}_{\leq 0} + \underbrace{\sin \tilde{a} \sin \tilde{b} \cos \theta}_{\leq 0} \leq 0, \quad |\Delta| \geq -\sin \tilde{a} \sin \tilde{b} \cos \theta.$$

Reversing the limits in (67),

$$\tilde{c} = \tilde{a} + \int_{\cos \tilde{a} + \Delta}^{\cos \tilde{a}} \frac{1}{\sqrt{1-s^2}} ds \geq \tilde{a} + \frac{|\Delta|}{\sin \tilde{a}} \geq \tilde{a} - \sin \tilde{b} \cos \theta.$$

Because  $0 < \tilde{b} \leq \pi/8 < 1/2$ , the alternating series for  $\sin$  yields  $\sin \tilde{b} \geq \tilde{b} - \tilde{b}^3/6$ . Hence

$$\tilde{c} \geq \tilde{a} - \tilde{b} \cos \theta + \frac{\tilde{b}^3}{6} \cos \theta, \quad \Rightarrow \quad a - c - b \cos \theta \leq \frac{\kappa b^2}{6} b |\cos \theta|.$$

Finally, since  $b \leq a/4$  and  $a < \varpi_\kappa/2 = \pi/(2\sqrt{\kappa})$ ,

$$\frac{\kappa b^2}{6} \leq \pi \frac{b}{a},$$

so the claimed inequality follows.  $\square$

**Lemma A.6.** Fix  $\kappa > 0$  and consider a geodesic triangle in  $M_{-\kappa}$  whose two adjacent sides have lengths  $a$  and  $b$  satisfying

$$0 < a, b < \frac{1}{\sqrt{\kappa}}.$$

Let  $\theta \in [0, \pi]$  be the angle between the two sides of lengths  $a$  and  $b$  at the vertex opposite to the side of length  $c = \gamma^{-\kappa}(\theta; a, b)$ . Then

$$a - c - b \cos(\theta) \geq -\frac{7}{6} \frac{b^2}{a} - \frac{\kappa b^2}{3} \cdot b |\cos(\theta)|.$$

*Proof.* The proof in the hyperbolic case is indeed similar, thanks to the hyperbolic law of cosines (24):

$$\cosh(\sqrt{\kappa}c) = \cosh(\sqrt{\kappa}a) \cosh(\sqrt{\kappa}b) - \sinh(\sqrt{\kappa}a) \sinh(\sqrt{\kappa}b) \cos(\theta).$$

Again, let us write

$$\tilde{a} = \sqrt{\kappa}a, \quad \tilde{b} = \sqrt{\kappa}b, \quad \tilde{c} = \sqrt{\kappa}c.$$

Recalling that

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \quad \text{and} \quad \sinh(x) = \frac{e^x - e^{-x}}{2}.$$

Let us write

$$\Delta = \tilde{c} - \tilde{a}.$$

Given  $\cosh(t)$  is concave up with derivative  $\sinh(t)$ , we have

$$\cosh(\tilde{c}) \geq \cosh(\tilde{a}) + \sinh(\tilde{a})\Delta,$$

which implies

$$(68) \quad \Delta \leq \frac{\cosh(\tilde{a})}{\sinh(\tilde{a})} (\cosh(\tilde{b}) - 1) - \sinh(\tilde{b}) \cos(\theta).$$

Estimate  $\sinh$  and  $\cosh$  by comparing their Taylor series with Geometric series. Recall the Taylor series of  $\sinh$  and  $\cosh$  at 0:

$$\sinh(x) = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} = x + \frac{x^3}{6} + \frac{x^5}{120} + \dots,$$

$$\cosh(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} = 1 + \frac{x^2}{2} + \frac{x^4}{24} + \dots.$$

When  $x \in [0, 1]$ , both series converges faster than a geometric series with a rate  $1/2$ , so we can use the following estimates:

$$(69) \quad \forall x \in [0, 1] : \quad x \leq \sinh(x) \leq x + \frac{x^3}{3} \leq \frac{4}{3}x \quad \text{and} \quad \cosh(x) \leq 1 + \frac{x^2}{2} + \frac{x^4}{12} \leq 1 + \frac{7}{12}x^2.$$

Estimate  $\Delta$ . The first term in the right-hand side of (68) can be estimated as follows.

$$\frac{\cosh(\tilde{a})}{\sinh(\tilde{a})}(\cosh(\tilde{b}) - 1) \leq \frac{2}{\tilde{a}} \frac{7}{12} \tilde{b}^2 = \frac{7}{6} \frac{\tilde{b}^2}{\tilde{a}}.$$

The second term can be estimated depending on the sign of  $\cos(\theta)$ . When  $\theta \in [0, \pi/2]$ , we have  $\cos(\theta) \geq 0$  and thus we can rely on the simple bound  $\sinh(\tilde{b}) \geq \tilde{b}$  to get

$$-\sinh(\tilde{b}) \cos(\theta) \leq -\tilde{b} \cos(\theta).$$

This gives us the estimate

$$\tilde{c} - \tilde{a} \leq \frac{7}{6} \frac{\tilde{b}^2}{\tilde{a}} - \tilde{b} \cos(\theta) \Leftrightarrow a - c - b \cos(\theta) \geq -\frac{7}{6} \frac{b^2}{a}.$$

As for  $\theta \in (\pi/2, \pi]$  where  $\cos(\theta) < 0$ , we invoke (69) to get

$$-\sinh(\tilde{b}) \cos(\theta) \leq -\tilde{b} \cos(\theta) - \frac{\tilde{b}^3}{3} \cos(\theta).$$

This gives us the estimate

$$\tilde{c} - \tilde{a} \leq \frac{7}{6} \frac{\tilde{b}^2}{\tilde{a}} - \tilde{b} \cos(\theta) - \frac{\tilde{b}^3}{3} \cos(\theta) \Leftrightarrow a - c - b \cos(\theta) \geq -\frac{7}{6} \frac{b^2}{a} - \frac{\kappa b^2}{3} \cdot b |\cos(\theta)|.$$

□

#### A.0.3. Spherical angle estimate via side ratio.

*Proof of Lemma 3.5.* We can assume  $\kappa = 1$  without loss of generality.

Spherical law of cosines and Taylor expansion. Recall the spherical law of cosines in the model space  $M_1$ :

$$\cos c = \cos a \cos b + \sin a \sin b \cos \theta$$

For the right hand side, we apply  $\cos(a + b) = \cos a \cos b - \sin a \sin b$  to get

$$\cos a \cos b + \sin a \sin b \cos \theta = \cos(a + b) - \sin a \sin b(1 - \cos \theta) \leq 1 - \sin a \sin b(1 - \cos \theta).$$

Since  $c \leq \frac{1}{2}a$ , the triangle inequality yields  $b \geq a - c \geq \frac{1}{2}a$ . Together with the assumption that  $a \leq \frac{\pi}{4}$  and using

$$\sin(t) \geq \frac{1}{\sqrt{2}} \frac{4}{\pi} t \quad t \in [0, \pi/4],$$

we obtain

$$1 - \sin a \sin b(1 - \cos \theta) \leq 1 - \frac{8}{\pi^2} \cdot \frac{1}{2} a^2 (1 - \cos \theta).$$

On the other hand, the standard estimate for the cosine function gives the bound on the left hand side of the spherical law of cosines:

$$\cos c \geq 1 - \frac{1}{2} c^2.$$

Substituting the two inequalities into the spherical law of cosines yields

$$(70) \quad \begin{aligned} 1 - \frac{1}{2}c^2 &\leq 1 - \frac{1}{2}\frac{8}{\pi^2}a^2(1 - \cos\theta) \\ &\Rightarrow c \geq \frac{\sqrt{8}}{\pi}a\sqrt{1 - \cos(\theta)}. \end{aligned}$$

Claim:  $\theta \leq \pi/3$ . Substituting the assumption  $c \leq \frac{1}{2}a$  into the above inequality (70) gives

$$\frac{1}{2} \geq \frac{\sqrt{8}}{\pi}\sqrt{1 - \cos(\theta)}.$$

Assume, for contradiction, that  $\theta \geq \pi/3$ . Then,

$$\frac{\sqrt{8}}{\pi}\sqrt{1 - \cos(\theta)} \geq \frac{2}{\pi} > \frac{1}{2}.$$

This contradicts the above inequality, hence we conclude that  $\theta < \pi/3$ .

Bounding  $\cos(\theta)$  from below on  $[0, \pi/3]$ . We claim that

$$(71) \quad \cos x \leq 1 - \frac{1}{2}\left(\frac{3x}{\pi}\right)^2, \quad \forall x \in [0, \pi/3].$$

Equivalently, set

$$F(x) := 1 - \frac{1}{2}\left(\frac{3x}{\pi}\right)^2 - \cos x, \quad x \in [0, \pi/3],$$

and show  $F(x) \geq 0$  on this interval. First, a direct computation gives  $F(0) = F(\pi/3) = 0$ . Now consider its derivative

$$F'(x) = -\frac{9}{\pi^2}x + \sin x.$$

We remark that  $F'$  has the same sign as

$$G(x) := \frac{F'(x)}{x} = -\frac{9}{\pi^2} + \frac{\sin x}{x}, \quad x \in (0, \pi/3].$$

on  $(0, \pi/3]$ . Because  $x \mapsto \sin x/x$  is strictly decreasing on  $(0, \pi/2)$ ,  $G$  is strictly decreasing on  $(0, \pi/3]$ . Moreover,

$$\lim_{x \rightarrow 0^+} G(x) = 1 > 0, \quad G(\pi/3) = -\frac{9}{\pi^2} + \frac{3\sin(\pi/3)}{\pi} < 0.$$

Hence there exists a unique  $x_0 \in (0, \pi/3)$  such that  $G(x_0) = 0$ , i.e.  $F'(x_0) = 0$ . Because  $G$  changes sign from positive to negative at  $x_0$ , we have

$$F'(x) \begin{cases} > 0, & x \in [0, x_0), \\ < 0, & x \in (x_0, \pi/3], \end{cases}$$

so  $F$  attains its unique local (and global) maximum at  $x_0$ . Since  $F$  vanishes at both endpoints and is non-negative at its single interior extremum, it follows that  $F(x) \geq 0$  for all  $x \in [0, \pi/3]$ , proving (71).

Conclusion. With our claim that  $\theta \leq \pi/3$ , we can substitute the above bound on  $\cos(x)$  into (70) to obtain

$$c \geq \frac{\sqrt{8}}{\pi}a\sqrt{\frac{1}{2}\left(\frac{3}{\pi}\theta\right)^2} \quad \Rightarrow \quad \frac{c}{a} \geq \frac{6}{\pi^2}\theta \geq \frac{1}{2}\theta.$$

□

A.0.4. *Opposite side estimates with error in terms of maximal side length.*

**Lemma A.7.** *Let  $(M, g)$  be a Riemannian manifold whose sectional curvatures are uniformly bounded in absolute value by a constant  $\kappa > 0$ . Suppose  $p, x, y$  be points in  $M$  such that*

$$0 < |px|, |py| < \frac{1}{8} \min\{\varpi_\kappa, r_{\text{inj}}(M)\}.$$

Let  $\theta = \angle(px, py) \in [0, \pi]$ . Then, the following holds:

$$(72) \quad \left| |xy|^2 - (|px|^2 + |py|^2 - 2|px||py|\cos(\theta)) \right| \leq 2\kappa \max\{|px|^4, |py|^4\}.$$

*Proof.* Without loss of generality, we can assume that

$$|px| \geq |py|.$$

We first invoke Lemma 3.1 to get

$$\gamma^\kappa(\theta; |px|, |py|) \leq |xy| \leq \gamma^{-\kappa}(\theta; |px|, |py|).$$

Now, we will bound the opposite sides in the two cases using the spherical and hyperbolic law of cosines, respectively.

Spherical Cosine Law. Let  $a = \sqrt{\kappa}|px|$ ,  $b = \sqrt{\kappa}|py|$ , and  $c = \sqrt{\kappa}\gamma^\kappa(\theta; |px|, |py|)$ . In particular, we have  $b \leq a \leq 1$ . The spherical law of cosines in the model space  $M_\kappa$  states

$$\cos c = \cos a \cos b + \sin a \sin b \cos \theta.$$

For  $0 \leq x \leq \pi/4 < 1$ , the Taylor expansion of  $\cos x$  and  $\sin x$  together with the alternating-series remainder give

$$1 - x^2/2 \leq \cos x \leq 1 - x^2/2 + x^4/24, \quad \text{and} \quad x - \frac{x^3}{6} \leq \sin x \leq x.$$

Hence, for  $a, b \in [0, \pi/8]$ , we have

$$1 - c^2/2 \leq \cos c.$$

Using the upper bound for  $\cos$  and the fact that  $b \leq a \leq 1$ ,

$$\begin{aligned} \cos a \cos b &\leq \left(1 - \frac{a^2}{2} + \frac{a^4}{24}\right) \left(1 - \frac{b^2}{2} + \frac{b^4}{24}\right) \\ &= 1 - \frac{a^2}{2} - \frac{b^2}{2} + \frac{a^2 b^2}{4} + \left(1 - \frac{a^2}{2}\right) \frac{b^4}{24} + \left(1 - \frac{b^2}{2}\right) \frac{a^4}{24} + \frac{a^4 b^4}{24} \\ &\leq 1 - \frac{a^2}{2} - \frac{b^2}{2} + a^4 \left(\frac{1}{4} + 2 \cdot \frac{1}{24} + \frac{1}{(24)^2}\right) \\ &= 1 - \frac{a^2}{2} - \frac{b^2}{2} + \frac{a^4}{2}. \end{aligned}$$

Because  $\cos(\theta)$  may be negative, we combine the upper and lower Taylor bounds for  $\sin$ :

$$\sin(a) \sin(b) \cos(\theta) \leq ab \cos \theta + \frac{ab^3}{12} + \frac{a^3b}{12} + \frac{a^3b^3}{36} \leq ab \cos \theta + \frac{7}{36} a^4,$$

again using  $b \leq a$ . Collecting the above bounds and substituting into the spherical law of cosines gives

$$1 - \frac{c^2}{2} \leq 1 - \frac{a^2}{2} - \frac{b^2}{2} + \frac{a^4}{2} + ab \cos(\theta) + \frac{7}{36} a^4.$$

which simplifies to

$$c^2 \geq a^2 + b^2 - 2ab \cos(\theta) - 2a^4.$$

Rewriting in the original variables yields

$$\gamma^\kappa(\theta; |px|, |py|)^2 \geq |px|^2 + |py|^2 - 2|px||py|\cos(\theta) - 2\kappa|px|^4.$$

Hyperbolic Cosine Law. Similar, let  $a = \sqrt{\kappa}|px|$ ,  $b = \sqrt{\kappa}|py|$ , and  $c = \sqrt{\kappa}\gamma^{-\kappa}(\theta; |px|, |py|)$ . The hyperbolic law of cosines in the model space  $M_{-\kappa}$  states

$$\cosh c = \cosh a \cosh b - \sinh a \sinh b \cos \theta.$$

For every  $x \geq 0$ ,

$$\cosh x = 1 + \frac{1}{2}x^2 + \sum_{k=2}^{\infty} \frac{x^{2k}}{(2k)!} \geq 1 + \frac{1}{2}x^2, \quad \Rightarrow \quad \cosh c \geq 1 + \frac{1}{2}c^2.$$

For  $0 \leq x \leq \pi/4 < 1$ ,

$$\cosh x = 1 + \frac{x^2}{2} + \sum_{k=2}^{\infty} \frac{x^{2k}}{(2k)!} \leq 1 + \frac{x^2}{2} + \frac{x^4}{24} \left( \sum_{s=0}^{\infty} \left( \frac{1}{5 \cdot 6} \right)^s \right) = 1 + \frac{x^2}{2} + \frac{x^4}{24} \cdot \frac{30}{29} \leq 1 + \frac{x^2}{2} + \frac{x^4}{23}.$$

so that

$$\begin{aligned} \cosh a \cosh b &\leq \left( 1 + \frac{a^2}{2} + \frac{a^4}{23} \right) \left( 1 + \frac{b^2}{2} + \frac{b^4}{23} \right) \\ &\leq 1 + \frac{a^2}{2} + \frac{b^2}{2} + \frac{a^4}{2}, \end{aligned}$$

where the inequality follows from expanding the product and using  $b \leq a \leq 1$ .

For the second term on the right hand side, for the same reason that  $\cos(\theta)$  can be negative. For  $0 \leq x \leq \pi/4 < 1$ ,

$$\begin{aligned} |\sinh(x) - x| &= \sum_{k=1}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \leq \frac{x^3}{3!} \sum_{s=0}^{\infty} \left( \frac{1}{4 \cdot 5} \right)^s = \frac{x^3}{6} \frac{20}{19} \leq \frac{x^3}{5} \\ \Rightarrow |\sinh(a) \sinh(b) \cos(\theta) - ab \cos(\theta)| &\leq \frac{ab^3 + ba^3}{5} + \frac{a^3 b^3}{25} \leq \frac{6}{25} a^4. \end{aligned}$$

Combining the above estimates, we have

$$c^2 \leq a^2 + b^2 - ab \cos(\theta) + 2a^4,$$

or equivalently,

$$\gamma^{-\kappa}(\theta; |px|, |py|)^2 \leq |px|^2 + |py|^2 - 2|px||py| \cos(\theta) + 2\kappa|px|^4.$$

□

## APPENDIX B. LOWER BOUNDS

This section discusses the lower bounds for errors of approximating distances from random geometric graphs. We are going to construct two manifolds which are “very close” in the sense that our random geometric graph model, with high probability, cannot distinguish them; on the other hand, the two manifolds are sufficiently different, and so this gives a lower bound for errors of distance estimation.

More precisely, we are going to construct  $(M_1, g_1, \mu_1)$  and  $(M_2, g_2, \mu_2)$  and a coupling  $\pi$  of

$$G_1 \sim G(n, M_1, g_1, \mu_1, p, s_n = 1) \quad \text{and} \quad G_2 \sim G(n, M_2, g_2, \mu_2, p, s_n = 1),$$

so that asymptotically the random geometric graph model cannot distinguish  $G_1$  and  $G_2$ . We are going to determine two special points  $X, Y$  so that  $|XY|_{g_1}$  and  $|XY|_{g_2}$  are sufficiently different.

What does our construction show? The upshot of our construction is we are going to disprove the following statement: there exists an algorithm which takes the random geometric graph as input and with high probability produces the intrinsic distances of all pairs of latent points simultaneously with the uniform error bound  $\lesssim n^{-6/(d+2)}$ . Since we can couple the two manifolds so that the resulting random geometric graphs from the two models are the same graph with high probability, we know that the algorithm would output the same distance estimates with high probability. However, since the two manifolds are different, our calculation below shows that one of the two applications of the algorithm must have too much error.

**B.1. Riemannian manifolds.** Let  $d \geq 2$  be a positive integer. Fix a sectional curvature upper bound  $\kappa > 0$  so that our two manifolds  $M_1$  and  $M_2$  to be constructed below respect this global curvature upper bound. We take  $(M_1, g_1)$  to be the flat  $d$ -torus realized as a point set as

$$M_1 := \mathbb{R}^d / \mathbb{Z}^d.$$

As a point set, we take  $M_2$  to be the same set

$$M_2 := \mathbb{R}^d / \mathbb{Z}^d,$$

while  $g_2$  is to be modified. We take a small radius  $r_{\text{bump}} > 0$  (where  $r_{\text{bump}}$  is a small number to be determined such that  $r_{\text{bump}} \rightarrow 0$  as  $n \rightarrow \infty$ , where  $n$  is the number of latent points to sample; for the moment, we can imagine  $r_{\text{bump}} < 0.01 \cdot \min\{1, \pi/\sqrt{\kappa}\}$ ), and consider the ball  $B_{g_1}(\vec{0}, r_{\text{bump}}) \subseteq \mathbb{R}^d / \mathbb{Z}^d$ . Outside this ball, let  $g_2$  be the same as  $g_1$  (flat metric), and inside the ball (“small bump”),  $g_2$  is to be modified in a radially-symmetric manner about  $\vec{0}$  (the center of the small bump).

Let us now define the two special points. Throughout this appendix, we let  $X := (-0.02, 0, \dots, 0)$  and  $Y := (0.02, 0, \dots, 0)$ , which exist on both manifolds. Note that both  $X$  and  $Y$  are outside of the small bump.

**B.2. Conformal change to the flat metric.** We use the usual coordinates  $x^1, x^2, \dots, x^d$  in a coordinate chart near  $\vec{0}$  that contains  $B(\vec{0}, 0.03)$ . With these coordinates, the original flat metric  $g_1 = (dx^1)^2 + (dx^2)^2 + \dots + (dx^d)^2$  has the matrix representation

$$g_1 = I_d \in \mathbb{R}^{d \times d}.$$

While we use the rectangular coordinates, it is helpful to consider  $r := \sqrt{(x^1)^2 + \dots + (x^d)^2}$ . For a smooth function  $h : [0, 0.03] \rightarrow \mathbb{R}_{>0}$  to be determined, we define the modified metric  $g_2$  via the radially-symmetric conformal change

$$g_2(p) = f(p) \cdot g_1(p) = h(\|p\|) \cdot g_1(p),$$

for every point  $p$  on this coordinate chart so that for every  $r \geq r_{\text{bump}}$ , we have  $h(r) = 1$ . This last condition guarantees that  $g_1$  and  $g_2$  agree everywhere outside of the small bump. In matrix form,  $g_2 = f(p) \cdot I_d = h(r) \cdot I_d$ . Following the usual convention, let us write  $f(p) := e^{2\varphi(p)}$ . Let us also

write  $h(r) = e^{2\psi(r)}$ , where  $\psi : [0, 0.03] \rightarrow \mathbb{R}$  so that  $\psi(r) = 0$  for every  $r \geq r_{\text{bump}}$ . This is saying that both  $f$  and  $\varphi$  factor through the norm function.

**B.2.1. Upper bound on sectional curvature in  $M_2$ .** Consider an arbitrary point  $p \in B(\vec{0}, r_{\text{bump}}) \subseteq M_2$ . Consider two arbitrary nonparallel tangent vectors  $u = u^i \partial_i, v = v^j \partial_j \in T_p M_2$ . The sectional curvature for the plane spanned by  $u$  and  $v$  is given by

$$K(u, v) := \frac{g_2(R(u, v)v, u)}{g_2(u, u)g_2(v, v) - g_2(u, v)^2} = \frac{1}{f(p)} \cdot \frac{g_1(R(u, v)v, u)}{g_1(u, u)g_1(v, v) - g_1(u, v)^2},$$

where  $R$  is the Riemann curvature tensor on  $M_2$  (with respect to  $g_2$ ). If we take  $u$  and  $v$  to be an orthonormal basis of the plane (with respect to  $g_1$ ), then the sectional curvature simplifies to

$$K(u, v) = \frac{1}{f(p)} \cdot g_1(R(u, v)v, u) = \frac{g_2(R(u, v)v, u)}{f(p)^2}.$$

Let us recall the following formula for Christoffel symbols (with respect to the metric  $g_2$ ):

$$\begin{aligned} \Gamma_{k\ell}^i &= \frac{1}{2} g^{im} \left( \frac{\partial g_{mk}}{\partial x^\ell} + \frac{\partial g_{m\ell}}{\partial x^k} - \frac{\partial g_{k\ell}}{\partial x^m} \right) \\ &= \frac{1}{2 \cdot f(p)} \cdot \left( \frac{\partial f(p)}{\partial x^\ell} \cdot \mathbf{1}_{i=k} + \frac{\partial f(p)}{\partial x^k} \cdot \mathbf{1}_{i=\ell} - \frac{\partial f(p)}{\partial x^i} \cdot \mathbf{1}_{k=\ell} \right). \end{aligned}$$

Since  $f(p) = e^{2\varphi(p)}$ , the chain rule gives

$$\Gamma_{k\ell}^i = \varphi_\ell \delta_{ik} + \varphi_k \delta_{i\ell} - \varphi_i \delta_{k\ell},$$

where  $\varphi_i = \partial_i \varphi = \partial \varphi / \partial x^i$ , and  $\delta_{ij}$  denotes the Kronecker delta  $\mathbf{1}_{i=j}$ .

Following Petersen [Pet16, §§ 3.1.6], we have the following formula for the  $(1, 3)$ -Riemann curvature tensor (with respect to  $g_2$ ):

$$g_2(R(\partial_i, \partial_j)\partial_k, \partial_\ell) = R_{ijk}^\ell = \partial_i \Gamma_{jk}^\ell - \partial_j \Gamma_{ik}^\ell + \Gamma_{jk}^s \Gamma_{is}^\ell - \Gamma_{ik}^s \Gamma_{js}^\ell.$$

So we have that

$$\begin{aligned} R_{ijk}^\ell &= (\varphi_{ik} - \varphi_i \varphi_k) \delta_{j\ell} + (\varphi_{j\ell} - \varphi_j \varphi_\ell) \delta_{ik} - (\varphi_{i\ell} - \varphi_i \varphi_\ell) \delta_{jk} - (\varphi_{jk} - \varphi_j \varphi_k) \delta_{i\ell} \\ &\quad + (\delta_{ik} \delta_{j\ell} - \delta_{jk} \delta_{i\ell}) \|\nabla \varphi\|^2. \end{aligned}$$

Here, the function  $\varphi_{ij}$  denotes  $\partial_i \varphi_j = \partial_i \partial_j \varphi$ , and

$$\|\nabla \varphi\|^2 = \sum_s \varphi_s^2.$$

Since

$$K(u, v) = \frac{1}{f(p)^2} \sum_{i,j,k,\ell} u^i v^j v^k u^\ell R_{ijk}^\ell,$$

we obtain the following formula for orthonormal  $u = \sum_i u^i \partial_i$  and  $v = \sum_j v^j \partial_j$ :

$$-f(p)^2 \cdot K(u, v) = \sum_{i,j} (u^i u^j + v^i v^j) (\varphi_{ij} - \varphi_i \varphi_j) + \|\nabla \varphi\|^2.$$

Now we use radial symmetry. Recall that

$$f(p) = h(\|p\|) = h(r) \quad \text{and} \quad \varphi(p) = \psi(\|p\|) = \psi(r).$$

By the chain rule, we find

$$\|(\nabla \varphi)(p)\|^2 = (\psi'(r))^2.$$

Note that for any  $i \in [d]$ , we have

$$\varphi_i(p) = \psi'(r) \cdot \frac{x^i}{r},$$

which implies, for any  $i, j \in [d]$ ,

$$\varphi_i \varphi_j = \frac{\psi'(r)^2}{r^2} \cdot x^i x^j,$$

and

$$\varphi_{ij} = \frac{\psi''(r)}{r^2} \cdot x^i x^j - \frac{\psi'(r)}{r^3} \cdot x^i x^j + \frac{\psi'(r)}{r} \cdot \delta_{ij}.$$

Hence, for any  $i, j \in [d]$ , we have

$$\varphi_{ij} - \varphi_i \varphi_j = \zeta(r) \cdot x^i x^j + \frac{\psi'(r)}{r} \delta_{ij},$$

where

$$\zeta(r) := \frac{\psi''(r)}{r^2} - \frac{\psi'(r)^2}{r^2} - \frac{\psi'(r)}{r^3}.$$

This implies the following simple formula for sectional curvatures:

$$-f(p)^2 \cdot K(u, v) = \zeta(r) \left( \sum_i u^i x^i \right)^2 + \zeta(r) \left( \sum_i v^i x^i \right)^2 + 2 \frac{\psi'(r)}{r} + \psi'(r)^2.$$

**Lemma B.1.** *Let  $u^1, \dots, u^d, v^1, \dots, v^d$  be real numbers such that*

$$\sum_i (u^i)^2 = \sum_i (v^i)^2 = 1 \quad \text{and} \quad \sum_i u^i v^i = 0.$$

*For any real numbers  $x^1, \dots, x^d$ , we have*

$$\left( \sum_i u^i x^i \right)^2 + \left( \sum_i v^i x^i \right)^2 \leq \sum_i (x^i)^2.$$

*Proof.* Consider the usual  $\mathbb{R}^d$  with the usual Euclidean norm  $\|\cdot\|$  and Euclidean metric  $\langle \cdot, \cdot \rangle$ . Write  $\mathbf{x} := (x^1, \dots, x^d)$ ,  $\mathbf{u} := (u^1, \dots, u^d)$ , and  $\mathbf{v} := (v^1, \dots, v^d)$ . Note that  $\mathbf{u}$  and  $\mathbf{v}$  are orthonormal vectors. Therefore, we can write

$$\mathbf{x} = \langle \mathbf{x}, \mathbf{u} \rangle \cdot \mathbf{u} + \langle \mathbf{x}, \mathbf{v} \rangle \cdot \mathbf{v} + \mathbf{y},$$

where the three summands are pairwise orthogonal. Thus,

$$\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{u} \rangle^2 + \langle \mathbf{x}, \mathbf{v} \rangle^2 + \|\mathbf{y}\|^2 \geq \langle \mathbf{x}, \mathbf{u} \rangle^2 + \langle \mathbf{x}, \mathbf{v} \rangle^2,$$

which is the desired inequality.  $\square$

Combining the lemma and the formula above, we obtain the following bound:

$$f(p)^2 \cdot |K(u, v)| \leq \max \left\{ \left| 2 \cdot \frac{\psi'(r)}{r} + \psi'(r)^2 \right|, \left| \psi''(r) + \frac{\psi'(r)}{r} \right| \right\},$$

which gives

$$|K(u, v)| \leq e^{-4\psi(r)} \max \left\{ \left| 2 \cdot \frac{\psi'(r)}{r} + \psi'(r)^2 \right|, \left| \psi''(r) + \frac{\psi'(r)}{r} \right| \right\},$$

By the triangle inequality, the following proposition is immediate.

**Proposition B.2.** *In the above notations, the magnitude of the sectional curvature  $K$  is bounded above as follows.*

$$|K| \leq 3e^{-4\psi(r)} \cdot \max \left\{ \frac{|\psi'(r)|}{r}, \psi'(r)^2, |\psi''(r)| \right\}.$$

**B.3. Using a smooth transition function.** In this subsection, we construct an explicit  $\psi$  from a smooth transition function. To begin with, let us consider the standard smooth transition function

$\tau : \mathbb{R} \rightarrow \mathbb{R}$  given by

$$\tau(x) := \begin{cases} 0, & \text{if } x < 0, \\ \left(1 + \exp\left(\frac{1}{x} - \frac{1}{1-x}\right)\right)^{-1}, & \text{if } 0 \leq x < 1, \\ 1, & \text{if } x \geq 1. \end{cases}$$

The following facts follow from single-variable calculus.

**Lemma B.3.** *The standard smooth transition function  $\tau$  satisfies the following properties:*

- (i)  $\tau : \mathbb{R} \rightarrow \mathbb{R}$  is smooth everywhere on  $\mathbb{R}$ .
- (ii)  $\tau$  is strictly increasing on  $[0, 1]$ .
- (iii)  $\tau(1/2) = 1/2$ .
- (iv)  $|\tau'(x)| \leq 2$ , for every  $x \in \mathbb{R}$ .
- (v)  $|\tau''(x)| \leq 10$ , for every  $x \in \mathbb{R}$ .

By linear scaling (dilation), we can define scaled versions of transition function. For any real numbers  $a, b, h$  with  $a < b$  and  $h > 0$ , we define  $\tau_{a,b;h} : \mathbb{R} \rightarrow \mathbb{R}$  given by

$$\tau_{a,b;h}(x) := \begin{cases} 0, & \text{if } x < a, \\ h \cdot \tau\left(\frac{x-a}{b-a}\right), & \text{if } a \leq x < b, \\ h, & \text{if } x \geq b. \end{cases}$$

We immediately have a scaled version of Lemma B.3 above.

**Lemma B.4.** *The standard smooth transition function  $\tau$  satisfies the following properties:*

- (i)  $\tau_{a,b;h}$  is smooth everywhere on  $\mathbb{R}$ .
- (ii)  $\tau_{a,b;h}$  is strictly increasing on  $[a, b]$ .
- (iii)  $\tau((a+b)/2) = h/2$ .
- (iv)  $|\tau'_{a,b;h}(x)| \leq 2h(b-a)^{-1}$ , for every  $x \in \mathbb{R}$ .
- (v)  $|\tau''_{a,b;h}(x)| \leq 10h(b-a)^{-2}$ , for every  $x \in \mathbb{R}$ .

Now we construct the function  $\psi : [0, 0.03] \rightarrow \mathbb{R}$  explicitly:

$$\psi(x) := \begin{cases} 0, & \text{if } 0 \leq x < \frac{r_{\text{bump}}}{2} \text{ or } x \geq r_{\text{bump}}, \\ \tau_{\frac{1}{2}r_{\text{bump}}, \frac{2}{3}r_{\text{bump}}; \alpha r_{\text{bump}}^2}(x), & \text{if } \frac{r_{\text{bump}}}{2} \leq x < \frac{2}{3}r_{\text{bump}}, \\ \alpha r_{\text{bump}}^2, & \text{if } \frac{2}{3}r_{\text{bump}} \leq x < \frac{5}{6}r_{\text{bump}}, \\ 1 - \tau_{\frac{5}{6}r_{\text{bump}}, \frac{5}{6}r_{\text{bump}}; \alpha r_{\text{bump}}^2}(x), & \text{if } \frac{5}{6}r_{\text{bump}} \leq x < r_{\text{bump}}, \end{cases}$$

where we may take  $\alpha \in \mathbb{R}$  to be any real number which satisfies

$$0 < \alpha < 2^{-12}\kappa.$$

It is immediate from our use of smooth transition functions that  $\psi : [0, 0.03] \rightarrow \mathbb{R}$  is smooth. Furthermore, since  $\psi$  is smooth and is identically zero on an open neighborhood of 0, it follows from the chain rule that  $\varphi(p) := \psi(\|p\|)$  is smooth everywhere in  $B(\vec{0}, 0.03)$ .

**Proposition B.5.** *The manifold  $(M_2, g_2)$ , where the metric  $g_2$  is obtained by the conformal change as described above, has the sectional curvature bound*

$$-\kappa < K < \kappa,$$

*everywhere on the manifold.*

*Proof.* This is a consequence of Proposition B.2 and Lemma B.4. Since  $\psi$  is nonnegative, Proposition B.2 implies that

$$|K| \leq 3 \cdot \max\left\{\frac{|\psi'(r)|}{r}, \psi'(r)^2, |\psi''(r)|\right\},$$

whence it suffices to show

$$\max\left\{\frac{|\psi'(r)|}{r}, \psi'(r)^2, |\psi''(r)|\right\} < \frac{\kappa}{3}.$$

Since  $\psi'$  vanishes on

$$\left[0, \frac{1}{2}r_{\text{bump}}\right) \cup \left(\frac{2}{3}r_{\text{bump}}, \frac{5}{6}r_{\text{bump}}\right) \cup \left(r_{\text{bump}}, 0.03\right),$$

we only need to consider the case

$$r \in \left[\frac{1}{2}r_{\text{bump}}, \frac{2}{3}r_{\text{bump}}\right] \cup \left[\frac{5}{6}r_{\text{bump}}, r_{\text{bump}}\right].$$

Now we use Lemma B.4. First,

$$\frac{|\psi'(r)|}{r} \leq 2\alpha r_{\text{bump}}^2 \left(\frac{1}{6}r_{\text{bump}}\right)^{-1} \left(\frac{1}{2}r_{\text{bump}}\right)^{-1} = 24\alpha < \frac{\kappa}{3}.$$

Second,

$$\psi'(r)^2 \leq (2\alpha r_{\text{bump}}^2)^2 \left(\frac{1}{6}r_{\text{bump}}\right)^{-2} = 144\alpha^2 r_{\text{bump}}^2 < \frac{\kappa}{3},$$

where in the last step we have used our assumptions that

$$\alpha < 2^{-12}\kappa \quad \text{and} \quad r_{\text{bump}} < \frac{\pi}{100\sqrt{\kappa}}.$$

Third,

$$|\psi''(r)| \leq 10(\alpha r_{\text{bump}}^2) \left(\frac{1}{6}r_{\text{bump}}\right)^{-2} = 360\alpha < \frac{\kappa}{3}.$$

We have finished the proof.  $\square$

We have successfully shown that our construction provides a conformal change to the metric in a way that respects the desired sectional curvature upper bound.

**B.4. Lower bound on the injectivity radius  $r_{\text{inj}}(M_2)$ .** In the original flat torus  $(M_1, g_1)$ , it is clear that the injectivity radius is  $r_{\text{inj}}(M_1) = 1/2$ , which is exactly half the length of the shortest geodesic loop on  $M_1$ . Above we have constructed  $(M_2, g_2)$  with a proven sectional curvature two-sided bounds  $-\kappa < K_{M_2} < \kappa$ . Potentially the injectivity radius  $r_{\text{inj}}(M_2)$  might be different from that of  $M_1$  after we added the bump. In this subsection, we provide a lower bound on  $r_{\text{inj}}(M_2)$ .

A crucial feature of our lower bound is that it is independent of  $n$ . Hence, the injectivity radius, even as the size of the small bump vanishes asymptotically as  $n \rightarrow \infty$ , remains bounded away from zero.

Our main tool for this part is the following theorem of Cheeger.

**Theorem B.6** ([Che70]; also see [CE08, Thm. 5.8]). *For any  $d \in \mathbb{Z}_{\geq 1}$ ,  $D > 0$ ,  $V > 0$ , and  $H \in \mathbb{R}$ , there exists a constant  $c_{B.6} = c(d, D, V, H) > 0$ , depending only on  $d, D, V, H$ , such that for every  $d$ -dimensional Riemannian manifold  $M$  with  $\text{diam}(M) \leq D$ ,  $\text{Vol}(M) \geq V$ , and  $K_M \geq H$ , every smooth closed geodesic on  $M$  has length greater than  $c_{B.6}$ .*

We claim that  $\text{diam}(M_2) \leq \frac{\sqrt{d}}{2} + 0.03$ . To see this, note that  $\text{diam}(M_1) = \frac{\sqrt{d}}{2}$ . For any two points  $p, q \in M_2$ , consider the path  $\gamma$  from  $p$  to  $q$  such that  $L(\gamma; M_1) = \|pq\|_{M_1}$ . Note that the difference between  $L(\gamma; M_1)$  and  $L(\gamma; M_2)$  can be only when  $\gamma$  passes through the small bump, and this adds at most

$$2r_{\text{bump}} \cdot e^{\alpha r_{\text{bump}}^2} \leq 0.02 \cdot e^{2^{-12} \cdot \pi^2 \cdot 10^{-4}} < 0.03.$$

Thus,  $L(\gamma; M_2) \leq L(\gamma; M_1) + 0.03 \leq \frac{\sqrt{d}}{2} + 0.03$ .

We claim that  $\text{Vol}(M_2) \geq 1$ . To see this, note that outside the small bump, the volume form is just the Euclidean volume form. Inside the small bump, the volume form satisfies

$$\sqrt{|\deg(g_2)|} \Big|_p = e^{d \cdot \varphi(p)} \geq 1.$$

Hence,  $\text{Vol}(M_2) \geq \text{Vol}(M_1) = 1$ .

From the previous subsection, we have shown  $K_{M_2} \geq -\kappa$ . Now we can show the following result.

**Proposition B.7.** *Let  $c_{B.6} = c(d, \sqrt{d}/2 + 0.03, 1, -\kappa) > 0$  be the constant from Theorem B.6. Define*

$$c_{B.7} = c_{B.7}(d, \kappa) := \min \left\{ \frac{\pi}{\sqrt{\kappa}}, \frac{1}{2} c_{B.6} \right\} > 0.$$

*Then  $r_{\text{inj}}(M_2) \geq c_{B.7}$ . Note that  $c_{B.7}$  depends only on  $d$  and  $\kappa$ .*

*Proof.* This follows immediately from combining our discussions above with Klingenberg's lemma (see e.g. [Pet16, Thm. 6.4.7]).  $\square$

**B.5. Difference in distances.** Recall that we consider two special points  $X = (-0.02, 0, \dots, 0)$  and  $Y = (+0.02, 0, \dots, 0)$ . In  $M_1$ , the distance  $|XY|_{M_1}$  is 0.04.

To ensure that  $|XY|_{M_2}$  is sufficiently larger than  $|XY|_{M_1}$ , let us from now on take

$$\alpha := 3 \cdot 2^{-14} \kappa.$$

This subsection proves the following lower bound:

**Proposition B.8.** *We have*

$$|XY|_{M_2} \geq 0.04 + 2^{-14} \cdot \min\{\kappa, 1\} \cdot r_{\text{bump}}^3.$$

*Proof.* Take an arbitrary curve  $\gamma$  from  $X$  to  $Y$ . We consider two cases.

Case 1. Suppose that  $\gamma$  never visits  $B(\vec{0}, r_{\text{bump}}/2)$ , the “flat” part inside of the bump. In this case, note that

$$L(\gamma; M_2) \geq L(\gamma; M_1),$$

and elementary geometry gives

$$L(\gamma; M_1) \geq 2 \cdot \sqrt{(0.02)^2 - (r_{\text{bump}}/2)^2} + 25r_{\text{bump}}^2 \geq 0.04 + 12r_{\text{bump}}^2,$$

which yields the desired lower bound.

Case 2. Suppose that  $\gamma$  visits  $B(\vec{0}, r_{\text{bump}}/2)$ . Then from our construction,  $\gamma$  must spend at least  $2 \cdot (r_{\text{bump}}/6) = r_{\text{bump}}/3$   $g_1$ -distance over where  $\psi$  is  $\alpha r_{\text{bump}}^2$ . Hence,

$$L(\gamma; M_2) \geq L(\gamma; M_1) + \frac{r_{\text{bump}}}{3} \cdot \left( e^{\alpha r_{\text{bump}}^2} - 1 \right) \geq 0.04 + \frac{1}{3} \alpha r_{\text{bump}}^3 = 2^{-14} \kappa r_{\text{bump}}^3.$$

This finishes the proof of the proposition.  $\square$

Using a similar idea, we can obtain an upper bound of the difference between  $|pq|_{M_1}$  and  $|pq|_{M_2}$  for arbitrary points  $p, q$  on the torus.

**Proposition B.9.** *For any two points  $p, q \in \mathbb{R}^d/\mathbb{Z}^d$ , we have*

$$0 \leq |pq|_{M_2} - |pq|_{M_1} < 2^{-11} \kappa r_{\text{bump}}^3.$$

*Proof.* Take a minimizing  $g_1$ -geodesic  $\gamma$  from  $p$  to  $q$ , so that  $|pq|_{M_1} = L(\gamma; M_1)$ . Note that the length of  $\gamma \cap B(\vec{0}, r_{\text{bump}})$  is at most  $2r_{\text{bump}}$ . Hence,

$$L(\gamma; M_2) - L(\gamma; M_1) \leq 2r_{\text{bump}} \left( e^{\alpha r_{\text{bump}}^2} - 1 \right) \leq \frac{5}{2} \alpha r_{\text{bump}}^3.$$

Now because  $\alpha = 3 \cdot 2^{-14} \kappa$ , we have

$$L(\gamma; M_2) - L(\gamma; M_1) < 2^{-11} \kappa r_{\text{bump}}^3.$$

Since  $L(\gamma; M_2) \geq |pq|_{M_2}$ , we have finished the proof.  $\square$

**B.6. Coupling.** Let us now construct the measures  $\mu_1$  and  $\mu_2$  on  $(M_1, g_1)$  and  $(M_2, g_2)$ , respectively. Take a small mass  $m > 0$  (where  $m = \frac{\text{polylog}(n)}{n}$  is much smaller than 1 to be determined; for the moment, imagine  $m < 0.01$ ). Define  $\mu$  to be the sum of  $(1 - 2m)$  times the uniform measure on  $\mathbb{R}^d/\mathbb{Z}^d$  (induced from the standard Euclidean measure) and two point masses at  $X = (-0.02, 0, \dots, 0)$  and  $Y = (0.02, 0, \dots, 0)$ , each of mass  $m$ . We let both  $\mu_1$  and  $\mu_2$  to be the same probability measure  $\mu$ . We can do this because  $(M_1, g_1)$  and  $(M_2, g_2)$ , as point sets, are the same set, even though the Riemannian metrics are different.

For the distance-probability function, let us simply take  $p(x) := e^{-x}$ , for all  $x \geq 0$ .

Now we construct the coupling between

$$G_1 \sim G(n, M_1, g_1, \mu_1, p, s_n = 1) \quad \text{and} \quad G_2 \sim G(n, M_2, g_2, \mu_2, p, s_n = 1).$$

First sample  $n$  i.i.d. points  $X_1, X_2, \dots, X_n$  from  $\mu$  on  $\mathbb{R}^d/\mathbb{Z}^d$ . Let  $\mathcal{U}_{i,j}$  be i.i.d. uniform(0, 1) random variables. Using the same  $\{X_i\}$  and  $\{\mathcal{U}_{i,j}\}$ , we define the edge sets:

$$E(G_1) := \left\{ \{i, j\} \in \binom{[n]}{2} : \mathcal{U}_{i,j} < p(|X_i X_j|_{M_1}) \right\}$$

and

$$E(G_2) := \left\{ \{i, j\} \in \binom{[n]}{2} : \mathcal{U}_{i,j} < p(|X_i X_j|_{M_2}) \right\}.$$

Note that  $E(G_2) \subseteq E(G_1)$ . It is easy to see that the marginal distributions of  $G_1$  and  $G_2$  indeed follow the correct random geometric graph models.

Our analysis so far in the previous subsections has not explicitly imposed how  $r_{\text{bump}}$  depends on  $n$ . This is because so far we only consider the geometry. Now that we would like the coupling  $\pi$  to be such that  $G_1 = G_2$ , let us now impose

$$r_{\text{bump}} < C(d, \kappa) \cdot n^{-\frac{2}{d+2}},$$

where  $C(d, \kappa) > 0$  is some constant depending on  $d$  and  $\kappa$ , to be described below.

To help with notations in our analysis below, let us describe sampling from  $\mu$  as follows. First, let us define i.i.d.  $W_1, W_2, \dots, W_n \sim \text{uniform}(0, 1)$ . For each  $i$ , if  $W_i \leq 1 - 2m$ , then take  $X_i \sim \text{uniform}(\mathbb{R}^d/\mathbb{Z}^d)$  according to the induced Lebesgue measure. If  $1 - 2m < W_i \leq 1 - m$ , take  $X_i := X$ . If  $1 - m < W_i$ , take  $X_i := Y$ .

We would like to analyze the probability that there is a minimizing geodesic segment from  $X_i$  to  $X_j$  which passes through the bump  $B(\vec{0}, r_{\text{bump}})$ . To do so, we need some geometry. For each point  $p \in M_1$ , let us define the following set (which we call “geodesic cone”) of  $p$

$$\mathcal{GC}(p) := \bigcup \left\{ q \in \gamma : \gamma \text{ is a minimizing } g_1\text{-geodesic from } p \text{ to } q \text{ such that } \gamma \cap B(\vec{0}, r_{\text{bump}}) \neq \emptyset \right\}.$$

The main property of this set that we use is that for any point  $q \notin \mathcal{GC}(p)$ , any minimizing  $g_1$ -geodesic from  $p$  to  $q$  does not pass through the bump  $B(\vec{0}, r_{\text{bump}})$ . Note that since our conformal transformation satisfies  $g_2 \geq g_1$  everywhere on the set  $\mathbb{R}^d/\mathbb{Z}^d$ , we find that any minimizing  $g_2$ -geodesic from  $p$  and  $q$  does not pass through the bump either. In particular, if  $q \notin \mathcal{GC}(p)$ , we obtain

$$|pq|_{M_1} = |pq|_{M_2}.$$

The following lemma can be proved by standard geometry.

**Lemma B.10.** *For every positive real number  $R \in (r_{\text{bump}}, \frac{\sqrt{d}}{2}]$ , and for every point  $p \in \mathbb{R}^d/\mathbb{Z}^d$ , if the  $g_1$ -distance from  $p$  to  $\vec{0}$  is  $R$ , then*

$$\text{Vol}(\mathcal{GC}(p)) = C_{\text{B.10}} \cdot \left( \frac{r_{\text{bump}}}{R} \right)^{d-1} \sqrt{1 - \left( \frac{r_{\text{bump}}}{R} \right)^2},$$

where  $C_{\text{B.10}} = C_{\text{B.10}}(d) > 0$  is a constant depending only on  $d$ .

In particular, the lemma shows that

$$\text{Vol}(\mathcal{GC}(p)) \leq C_{\text{B.10}} \cdot \left( \frac{r_{\text{bump}}}{R} \right)^{d-1}.$$

Now consider a point  $Y \sim \text{uniform}(\mathbb{R}^d/\mathbb{Z}^d)$  according to the standard Lebesgue measure induced on the flat torus. Let  $\mathfrak{R}$  denote the  $g_1$ -geodesic distance from  $Y$  to  $\vec{0}$ . It is not hard to see that  $\mathfrak{R}$  is absolutely continuous, so let us write  $f_{\mathfrak{R}}$  for the unique continuous density of  $\mathfrak{R}$ . Then the following estimate on  $f_{\mathfrak{R}}$  of  $\mathfrak{R}$  follows from the volume formula for spheres:

**Lemma B.11.** *For  $0 \leq R \leq \frac{\sqrt{d}}{2}$ , we have*

$$f_{\mathfrak{R}}(R) \leq C_{\text{B.11}} \cdot R^{d-1},$$

where  $C_{\text{B.11}} = C_{\text{B.11}}(d) > 0$  is a constant depending only on  $d$ .

Now we are ready to bound the probability that  $G_1 \neq G_2$ . Using the union bound, we find

$$(73) \quad \mathbb{P}\{G_1 \neq G_2\} \leq \sum_{\{i,j\} \in \binom{[n]}{2}} \mathbb{P}\{\text{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \text{p}(|X_i X_j|_{M_1})\}.$$

Let us write  $\mathcal{T}(X_i, X_j)$  to denote the event where there is a minimizing  $g_1$ -geodesic from  $X_i$  to  $X_j$ . From our above discussion, if  $\mathcal{T}(X_i, X_j)$  does not happen, then  $|X_i X_j|_{M_1} = |X_i X_j|_{M_2}$ , and hence

$$\mathbb{P}\{\text{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \text{p}(|X_i X_j|_{M_1}) \mid \mathcal{T}(X_i, X_j)^c\} = 0.$$

Let us break into cases according to  $W_i, W_j$ .

First, if both  $W_i$  and  $W_j$  are in  $[0, 1 - 2m]$ , then  $X_i$  and  $X_j$  are simply i.i.d. uniform on the torus. Thus,

$$\begin{aligned} & \mathbb{P}\{\text{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \text{p}(|X_i X_j|_{M_1}) \mid W_i, W_j \in [0, 1 - 2m]\} \\ &= \mathbb{P}\{\text{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \text{p}(|X_i X_j|_{M_1}) \mid W_i, W_j \in [0, 1 - 2m], \mathcal{T}(X_i, X_j)\} \\ & \quad \times \mathbb{P}\{\mathcal{T}(X_i, X_j) \mid W_i, W_j \in [0, 1 - 2m]\}. \end{aligned}$$

Now we bound the two factors on the right-hand side. For the first factor, Proposition B.9 shows that

$$\mathbb{P}\{\text{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \text{p}(|X_i X_j|_{M_1}) \mid W_i, W_j \in [0, 1 - 2m], \mathcal{T}(X_i, X_j)\} \leq 2^{-11} \kappa r_{\text{bump}}^3,$$

where we have used the upper Lipschitz constant of  $\text{p}$ . For the second factor, let us define  $p$  and  $q$  to be i.i.d. uniform on the torus to avoid confusion. We have

$$\mathbb{P}\{\mathcal{T}(X_i, X_j) \mid W_i, W_j \in [0, 1 - 2m]\} = \mathbb{P}\{\mathcal{T}(p, q)\}.$$

Lemmas B.10 and B.11 yield

$$\begin{aligned} \mathbb{P}\{\mathcal{T}(p, q)\} &= \int_0^{\sqrt{d}/2} \mathbb{P}\left\{\mathcal{T}(p, q) \mid d_{g_1}(p, \vec{0}) = R\right\} \cdot f_{\mathfrak{R}}(R) \, dR \\ &\leq \int_{r_{\text{bump}}}^{\sqrt{d}/2} C_{\text{B.10}} \cdot \left( \frac{r_{\text{bump}}}{R} \right)^{d-1} \cdot C_{\text{B.11}} \cdot R^{d-1} \, dR + \int_0^{r_{\text{bump}}} C_{\text{B.11}} \cdot R^{d-1} \, dR \\ &\leq C(d) \cdot r_{\text{bump}}^{d-1}, \end{aligned}$$

where  $C(d) > 0$  is a constant depending only on  $d$ .

This shows that

$$\mathbb{P}\{\mathbf{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \mathbf{p}(|X_i X_j|_{M_1}) \mid W_i, W_j \in [0, 1 - 2m]\} \leq C(d, \kappa) \cdot r_{\text{bump}}^{d+2},$$

where  $C(d, \kappa) > 0$  is a constant depending only on  $d$  and  $\kappa$ .

Second, if one of  $W_i$  and  $W_j$  is in  $[0, 1 - 2m]$  and one is in  $(1 - 2m, 1]$ , then this is saying one of  $X_i$  and  $X_j$  is  $X$  or  $Y$ . For instance, we consider

$$\begin{aligned} & \mathbb{P}\{\mathbf{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \mathbf{p}(|X_i X_j|_{M_1}) \mid W_i \in [0, 1 - 2m], W_j \in (1 - 2m, 1]\} \\ &= \mathbb{P}\{\mathbf{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \mathbf{p}(|X_i X_j|_{M_1}) \mid W_i \in [0, 1 - 2m], W_j \in (1 - 2m, 1], \mathcal{T}(X_i, X_j)\} \\ &\quad \times \mathbb{P}\{\mathcal{T}(X_i, X_j) \mid W_i \in [0, 1 - 2m], W_j \in (1 - 2m, 1]\}. \end{aligned}$$

Once again, the first factor can be bounded by Proposition B.9. For the second factor, note that  $X_j$  is either  $X$  or  $Y$ , which is exactly 0.02 away from  $\vec{0}$ . Lemma B.10 yields

$$\mathbb{P}\{\mathcal{T}(X_i, X_j) \mid W_i \in [0, 1 - 2m], W_j \in (1 - 2m, 1]\} \leq C(d, \kappa) \cdot r_{\text{bump}}^{d+2},$$

where  $C(d, \kappa) > 0$  is a constant depending only on  $d$  and  $\kappa$ .

Third, if both  $W_i$  and  $W_j$  are in  $(1 - 2m, 1]$ , then we simply use Proposition B.9 to obtain

$$\mathbb{P}\{\mathbf{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \mathbf{p}(|X_i X_j|_{M_1}) \mid W_i, W_j \in (1 - 2m, 1]\} \leq 2^{-11} \kappa r_{\text{bump}}^3.$$

Combining all cases, we have shown:

$$\begin{aligned} & \mathbb{P}\{\mathbf{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \mathbf{p}(|X_i X_j|_{M_1})\} \\ & \leq C(d, \kappa) \left( (1 - 2m)^2 r_{\text{bump}}^{d+2} + 4m \cdot r_{\text{bump}}^{d+2} + 4m^2 \cdot r_{\text{bump}}^3 \right), \end{aligned}$$

where  $C(d, \kappa) > 0$  depends only on  $d$  and  $\kappa$ . We record this as a proposition below.

**Proposition B.12.** *For any  $i, j \in [n]$ , we have*

$$\mathbb{P}\{\mathbf{p}(|X_i X_j|_{M_2}) < \mathcal{U}_{i,j} \leq \mathbf{p}(|X_i X_j|_{M_1})\} \leq C_{B.12}(d, \kappa) \cdot \left( r_{\text{bump}}^{d+2} + m^2 \cdot r_{\text{bump}}^3 \right),$$

where  $C_{B.12}(d, \kappa) > 0$  depends only on  $d$  and  $\kappa$ .

Combining this proposition with the union bound (73), we conclude that if  $r_{\text{bump}} = o(n^{-2/(d+2)})$  and  $m = \frac{\text{polylog}(n)}{n}$ , then  $\mathbb{P}\{G_1 \neq G_2\} = o(1)$  as  $n \rightarrow \infty$ .

**B.7. Final step.** We have shown above that  $G_1 = G_2$  with high probability. It is not hard to see that our construction satisfies the regularity assumptions on  $\mathbf{p}$ ,  $\mu$ , and  $M_i$ . It remains to show that both  $X$  and  $Y$  appear as latent points with high probability.

This last step is simple. We have

$$\mathbb{P}\{\forall i \in [n], W_i \leq 1 - 2m\} = (1 - 2m)^n \leq e^{-2mn} \rightarrow 0,$$

since  $m = \frac{\text{polylog}(n)}{n}$ .

Recall from Proposition B.8 that the two points  $X$  and  $Y$  satisfy

$$|XY|_{M_2} - |XY|_{M_1} \geq C(d, \kappa) \cdot r_{\text{bump}}^3,$$

where  $C(d, \kappa) > 0$  depends on only  $d$  and  $\kappa$ .

Since we have shown above that we can pick any  $r_{\text{bump}}$  for which

$$r_{\text{bump}} = o\left(n^{-\frac{2}{d+2}}\right),$$

we have finished proving Theorem 1.8.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MISSOURI, COLUMBIA, MO 65203  
*Email address*, H. Huang: [hhuang@missouri.edu](mailto:hhuang@missouri.edu)

SCIENCE DIVISION, MAHIDOL UNIVERSITY INTERNATIONAL COLLEGE, NAKHON PATHOM 73170, THAILAND  
*Email address*, P. Jiradilok: [pakawut.jir@mahidol.edu](mailto:pakawut.jir@mahidol.edu)

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139  
*Email address*, E. Mossel: [elmos@mit.edu](mailto:elmos@mit.edu)