

Euclidean Ideas Beyond Euclidean Geometry

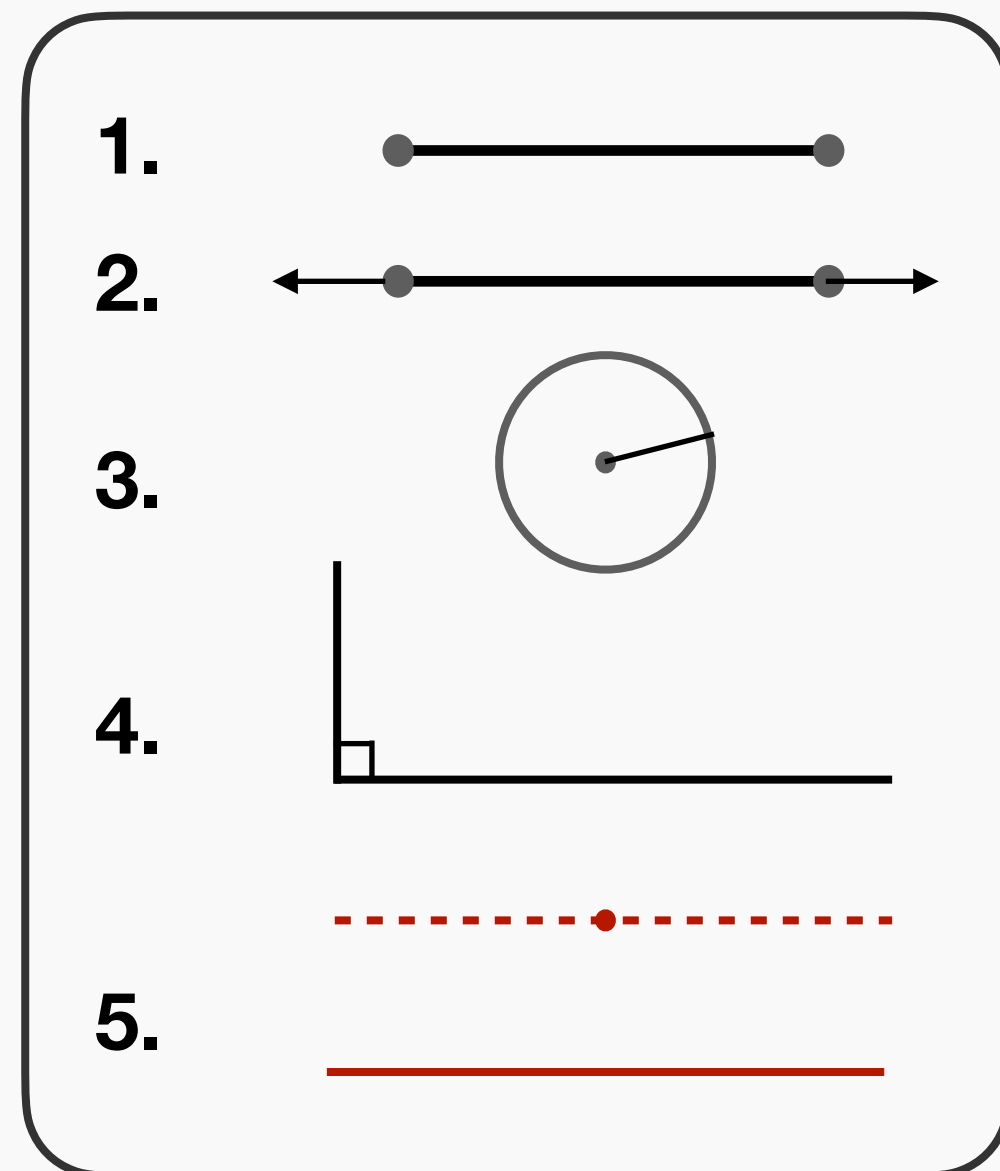
Chengyuan Deng

Johns Hopkins University

Sept 9th, 2026

When Geometry Went Beyond Euclid

Euclid's 5 Postulates



Derive the 5th
from the first 4?

The “birth” of hyperbolic geometry

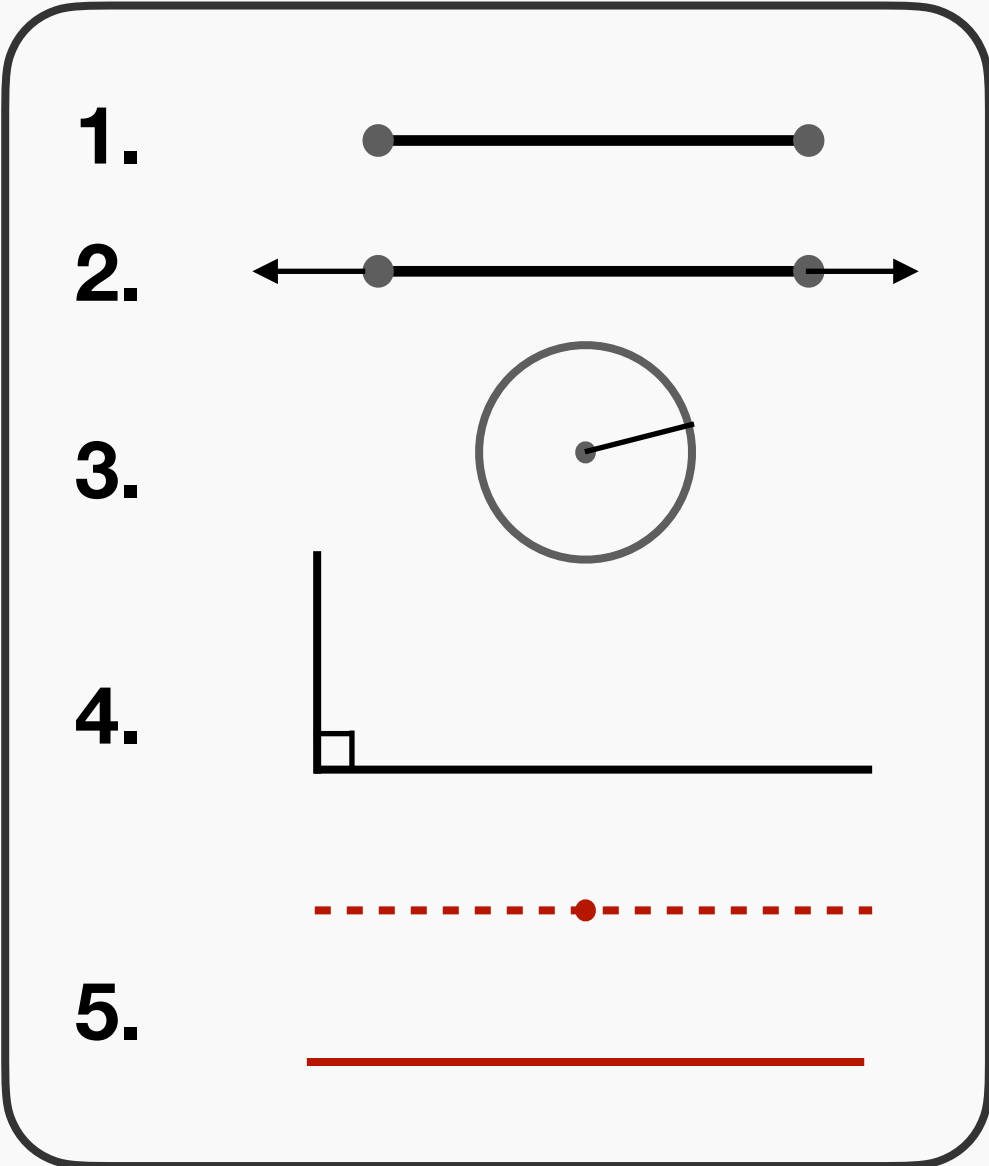
- Saccheri assumed P5 does not hold but he rejected himself later.
- Lobachevsky & Bolyai independently developed a geometry without P5
- Beltrami built a 2D model; Klein unified the name as hyperbolic geometry

300 BCE

1733, 19th century

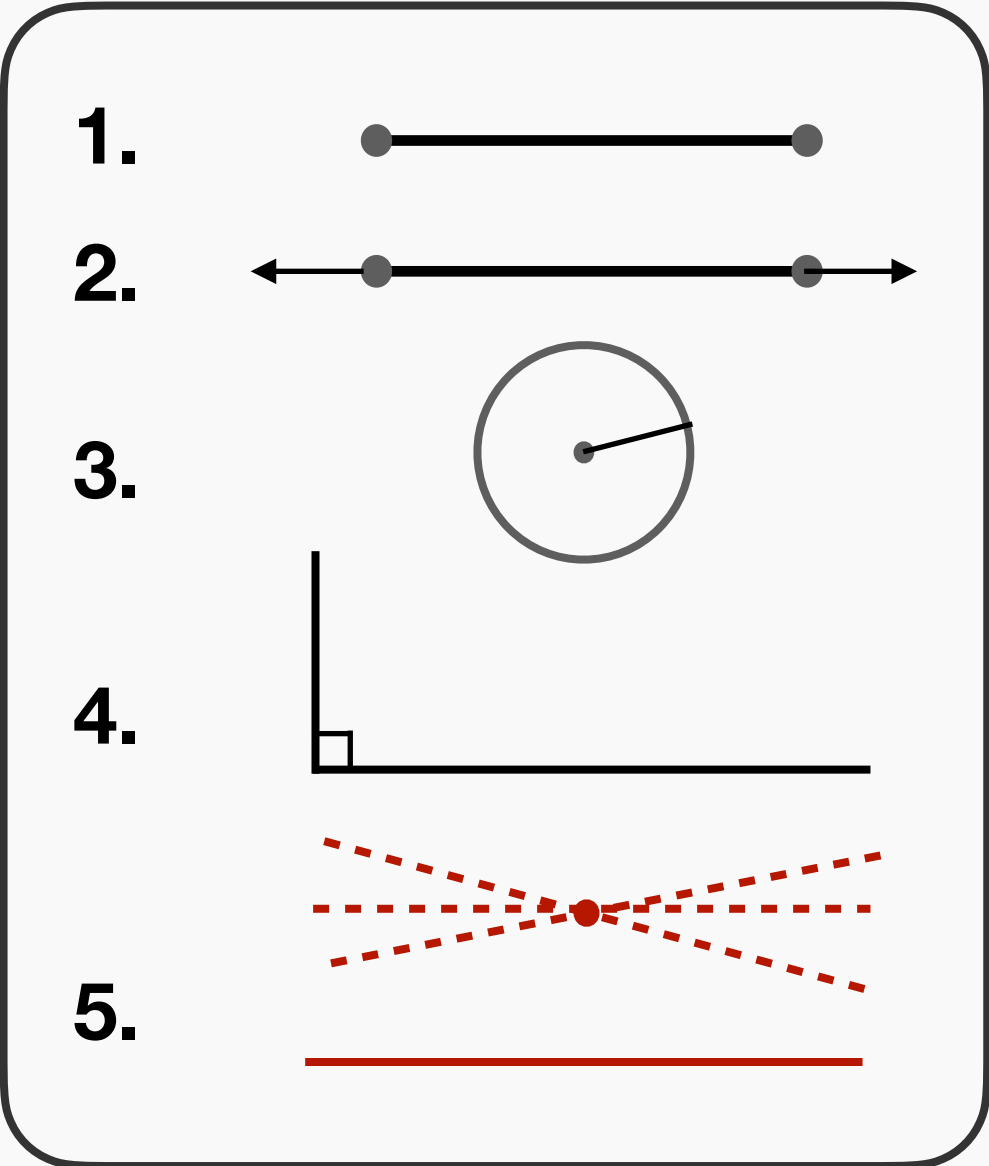
When Geometry Went Beyond Euclid

Euclid's 5 Postulates



Derive the 5th
from the first 4?

The “birth” of hyperbolic geometry



Metric Space

- Fréchet introduced abstract metric space, keeping only the distances, i.e. no coordinates required.
- (X, d) , s.t.
 $d(x, z) \leq d(x, y) + d(y, z)$
 $d(x, x) = 0, d(x, y) = d(y, x)$

300 BCE

1733, 19th century

1906

✓ Distance, geodesics, angles, convexity

✗ Parallelism, curvature, packing, etc.

✓ Distance, triangle inequality

✗ Coordinates, geodesics, angles, etc.

Some **math definitions** survive, some don't. How about **algorithms**?

Classical Algorithmic Ideas in Euclidean Geometry

Clustering

NP-hard, PTAS via
coresets and grids

Dimension Reduction

n points, $(1 + \varepsilon)$ -approx
 $\Theta(\varepsilon^{-2} \log n)$ dimension,
via random projection

TSP

NP-hard, PTAS via
randomly shifted quad-tree

Given a set of n points in $\mathbb{R}^d$...

Closest Pair

$\tilde{\Theta}(n)$ time via divide and
conquer and packing

Spanners

$(1 + \varepsilon)$ -approx with linear
size via well-separated pair
decomposition and packing

Optimal Transport

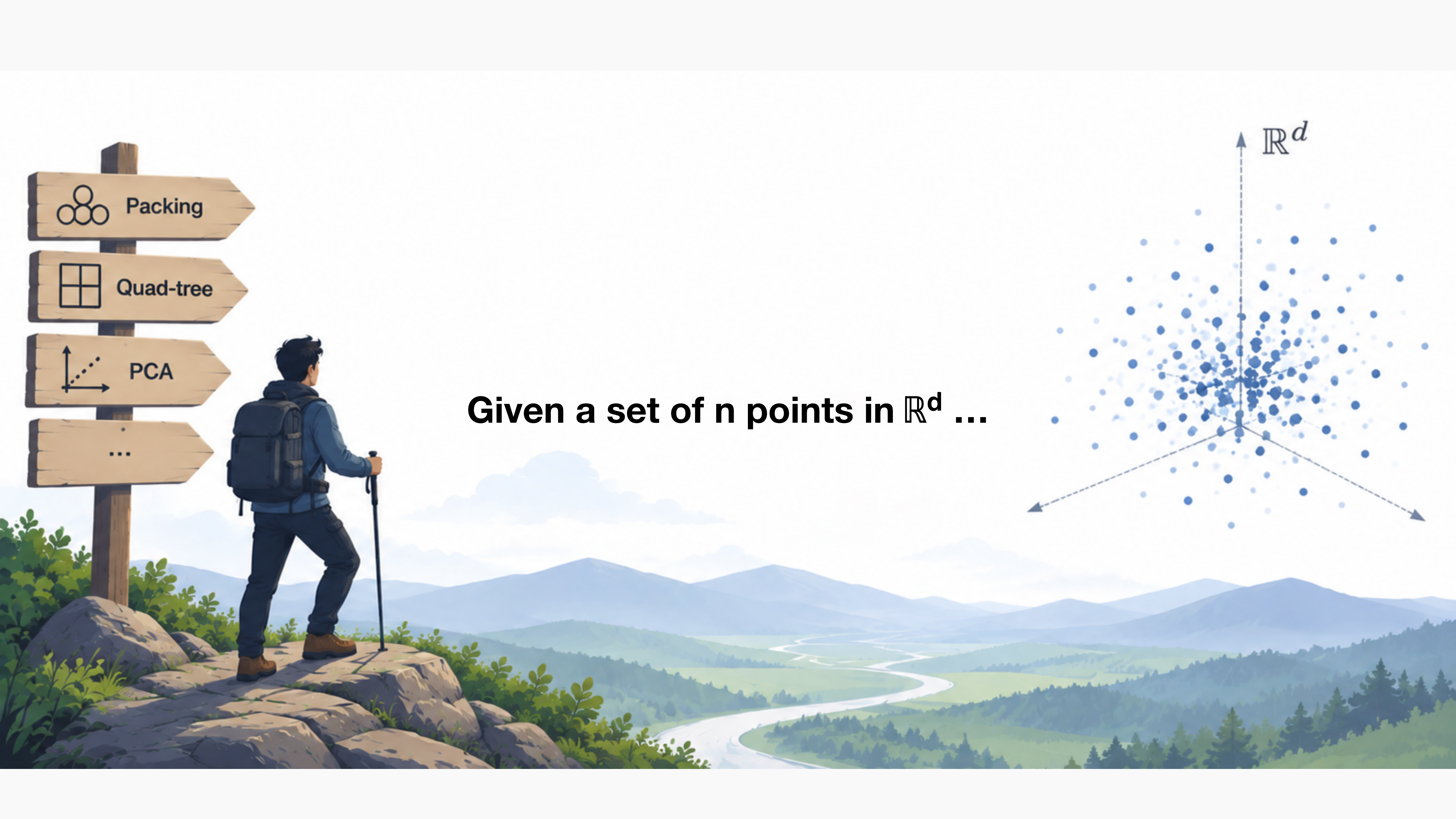
$(1 + \varepsilon)$ -approx in near-
linear time via randomly
shifted quad-tree

Range Searching

Sublinear query time via range
trees/k-d trees

Nearest Neighbor Search

Sublinear query time via locality-
sensitive hashing (grids)



Given a set of n points in $\mathbb{R}^d$...

The Euclidean Ideas

A few properties of Euclidean geometry (may not be [Euclidean-unique](#))

Property	Euclidean Toolkit
Scale Invariance	Grid, random shifting, quad-tree
Positive Definiteness	Random projection (JL), spectral decomposition (PCA, MDS)
Packing	r -nets, well-separated pair decomposition
Convexity	Voronoi diagram, LP-type methods
...	...

Many Euclidean algorithms are a consequence of combining a few ideas based on a few properties.

The Euclidean Ideas

A few properties of Euclidean geometry (may not be Euclidean-unique)

Property

Euclidean Toolkit

Scale Invariance

Grid, random shifting, quad-tree

Positive Definiteness

Random projection (JL),
spectral decomposition (PCA, MDS)

Packing

r -nets, well-separated pair decomposition

Convexity

Voronoi diagram, LP-type methods

...

...

- Euclidean LSH

Random shifted grids + Random projection

The Euclidean Ideas

A few properties of Euclidean geometry (may not be Euclidean-unique)

Property	Euclidean Toolkit
Scale Invariance	Grid, random shifting, quad-tree
Positive Definiteness	Random projection (JL), spectral decomposition (PCA, MDS)
Packing	r -nets, well-separated pair decomposition
Convexity	Voronoi diagram, LP-type methods
...	...

- Euclidean k -means PTAS

Randomized shifted grid + bounded local density of separated centers

The Euclidean Ideas

A few properties of Euclidean geometry (may not be Euclidean-unique)

Property

Euclidean Toolkit

Scale Invariance

Grid, random shifting, quad-tree

Positive Definiteness

Random projection (JL),
spectral decomposition (PCA, MDS)

Packing

r -nets, well-separated pair decomposition

Convexity

Voronoi diagram, LP-type methods

...

...

- Euclidean $(1 + \varepsilon)$ -Spanner

Convex cones + inner-product angle bounds +
constant number of angular sectors

Beyond Euclidean Geometry

Utilize the survived properties

- **Packing property** holds for **doubling metrics**

With doubling dimension δ , an r -separated subset of an R -ball has size roughly $(R/r)^{O(\delta)}$

In $\mathbb{R}^d$, an r -separated subset of an R -ball has size at most $(1 + 2R/r)^d$

The most well-developed direction:

TSP PTAS; k -clustering, facility location PTAS; Spanners; Locality sensitive ordering...

- **Scale invariance** holds for all **normed spaces**
- **Convexity** holds for **hyperbolic** spaces and other **CAT(0) metric spaces**

Beyond Euclidean Geometry

Find a connection to Euclidean, with some compromise

- Bourgain's Theorem

Every n -point metric can be embedded into Euclidean space of dimension $O(\log^2 n)$ with **multiplicative distortion** $\Theta(\log n)$

Alg gives
coordinates

- Riemannian optimization methods solve Euclidean subproblems in tangent spaces (not today)
- [DGLX'24] Additive error is also possible

Every size- n symmetric hollow matrix can be written as a Euclidean distance matrix minus a constant e_n^2 on the off-diagonal entries

Run MDS, get
coordinates

Beyond Euclidean Geometry

Explore new properties

- Structural decomposition
 - Tree — heavy-light decomposition, recursive decomposition, etc.
 - Planar and minor-free graphs — separators, Baker's technique, etc.
- Geometry-induced structures
 - Ultrametrics — $d(x, z) \leq \max\{d(x, y), d(y, z)\}$, gives a hierarchical structure
 - Hyperbolic geometry — different structures at different scales

Quad-tree, spanner [KW'24]

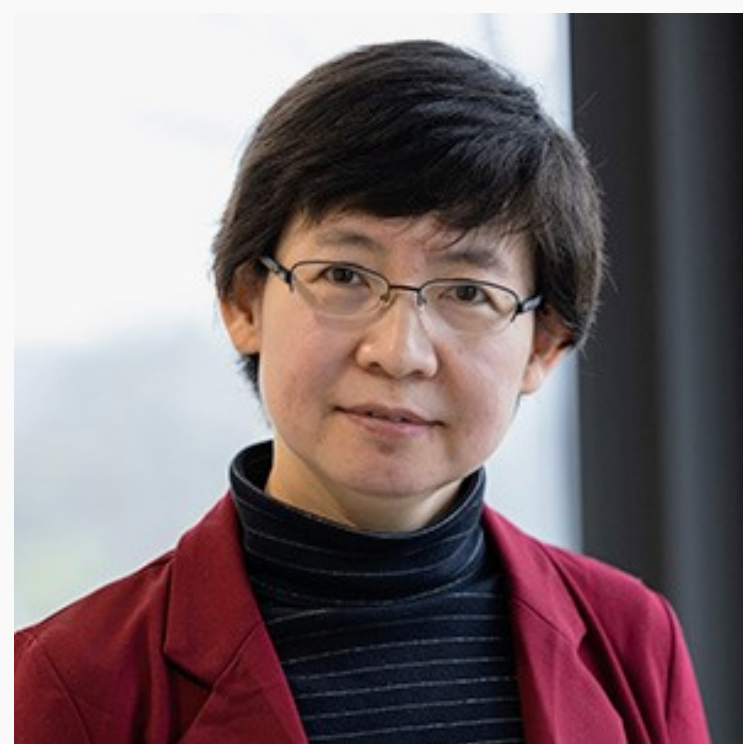
Navigable graphs [DGLLLWX'26]

- Locally, “Euclidean-like”, apply Euclidean techniques
- Globally, “tree-like”, use other techniques to accommodate the exponential volume growth
- Combine two parts

References and Collaborators

- Neuc-MDS: Non-Euclidean Multi-dimensional Scaling Through Bilinear Forms, *Neurips 2024*
- Johnson-Lindenstrauss Lemma Beyond Euclidean Geometry, *NeurIPS 2025*
- Locality Sensitive Hashing in Hyperbolic Space, *SoCG 2026*
- Navigable Graph in Low-dimensional Hyperbolic Space, *In Submission*

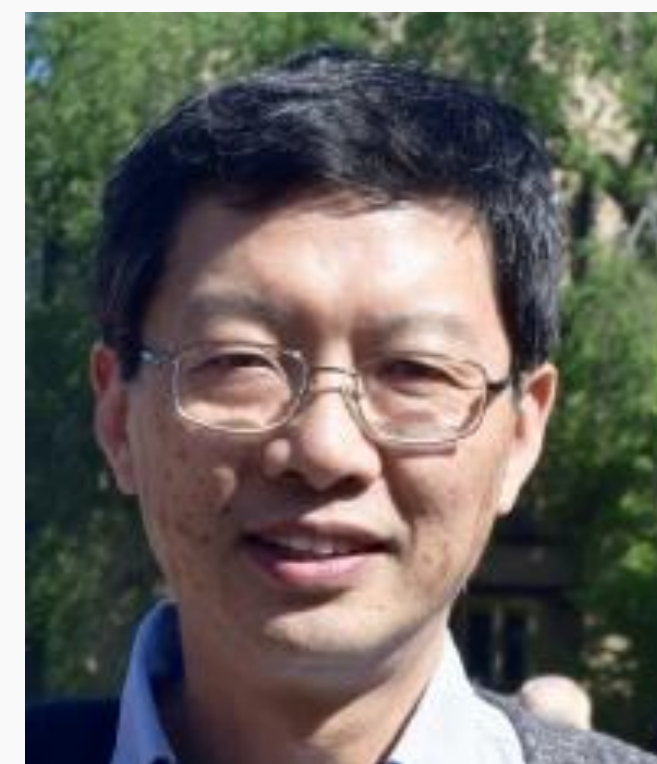
Joint work with



Jie Gao



Kevin Lu



Feng Luo



Cheng Xin

Dimension Reduction

Johnson-Lindenstrauss Lemma [JL'84]

Proposition: Euclidean data of size n can be projected to $\Theta(\log n/\epsilon^2)$ dimension, with pairwise distances preserved up to $1 \pm \epsilon$ multiplicative factor.

The projection is a random linear map, which makes JL transform also very practical

- Significant dimension reduction, irrelevant to the original dimension
- Easy to implement
- Data-independent

In practice JL transform is applied regardless of the data geometry

😞 Lower bound results: ℓ_1, ℓ_p , dimension $\Omega_d(\sqrt{n})$ for constant approximation

😊 Still holds for Hilbert spaces (but trivial)

JL Lemma in Non-Euclidean Geometry

Setup

- Two assumptions on the dissimilarity measure: $d(x, x) = 0, d(x, y) = d(y, x)$
- No coordinates available
- The input dissimilarity matrix is a symmetric hollow matrix

Find one connection to Euclidean — Generalized power distance

Key lemma: Every symmetric hollow matrix with squared dissimilarities can be written as a squared Euclidean distance matrix minus a constant e_n^2

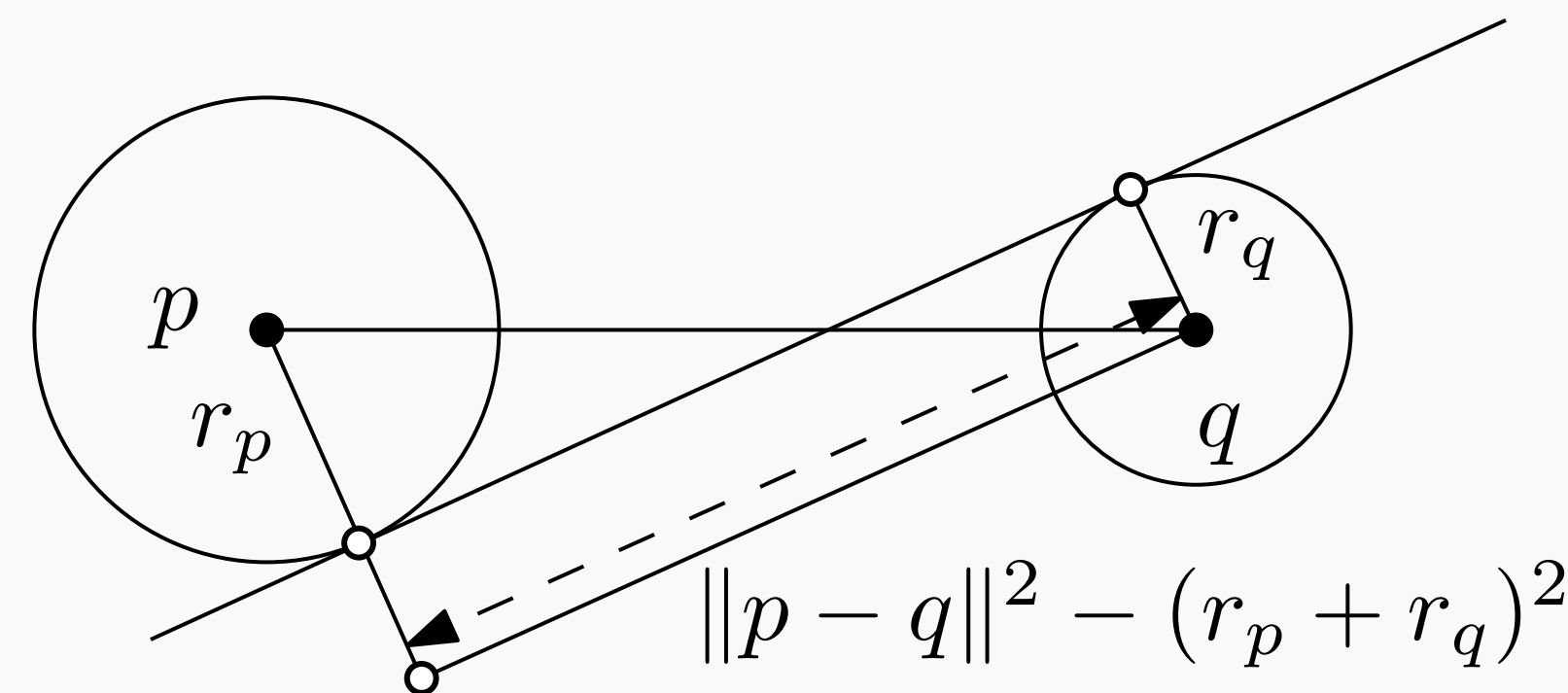
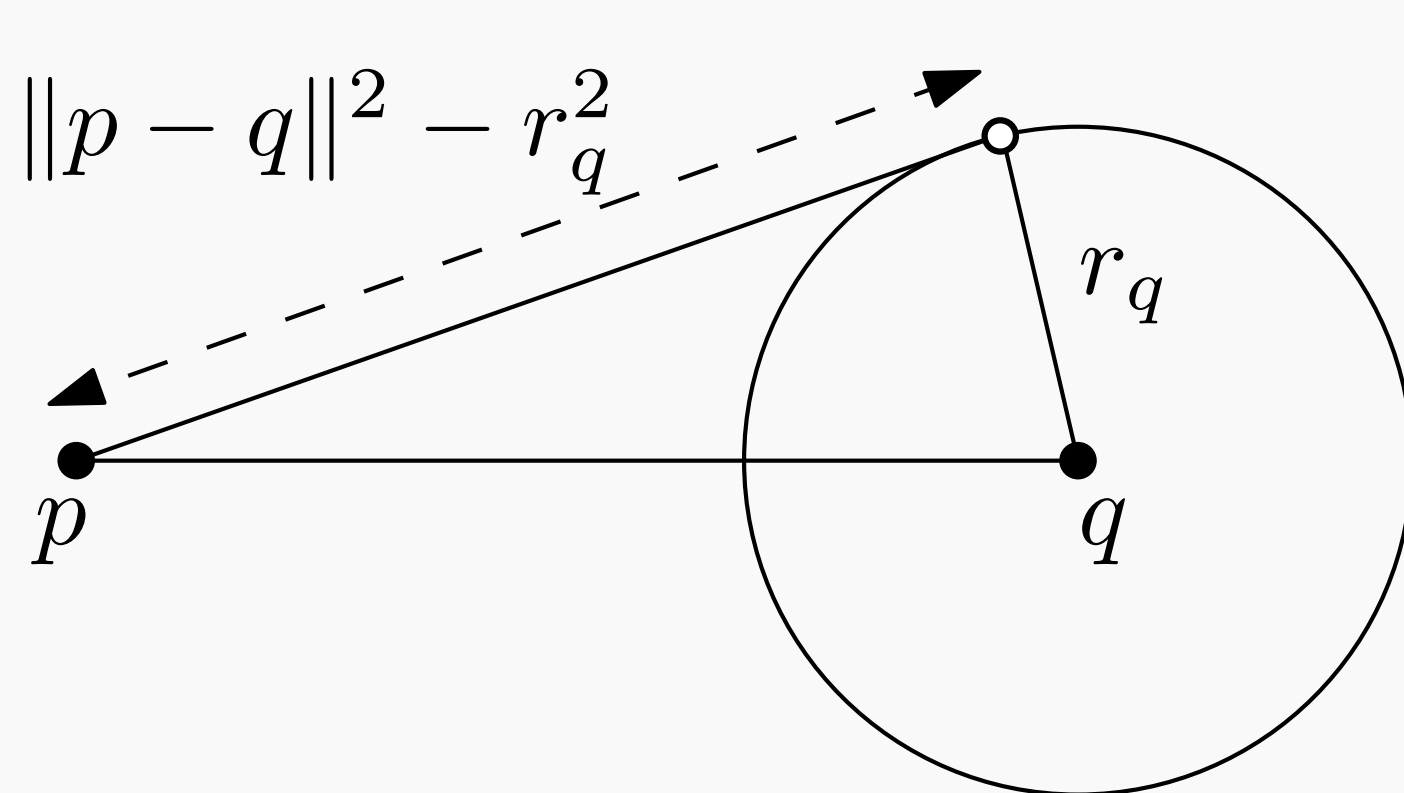
Theorem: Given $D, \varepsilon \in (0, 1)$, there exists a map to dimension $O(\log n / \varepsilon^2)$, s.t. every D_{ij} is preserved within $(1 \pm \varepsilon)$ multiplicative factor and $2\varepsilon e_n^2$ additive factor, where e_n is the smallest eigenvalue of the Gram matrix of D

JL Lemma in Non-Euclidean Geometry

Generalized power distance

Given two balls centered at $p, q \in \mathbb{R}^d$, with radii r_p, r_q , the generalized power distance between p, q is

$$\|p - q\|^2 - (r_p + r_q)^2$$



For the key lemma, we show there exists $r = r_p = r_q = e_n$

Parameter r can be interpreted as measure of **deviation from Euclidean geometry**

JL Lemma in Non-Euclidean Geometry

The algorithm

1. Input D , obtain Euclidean matrix $E = D + e_n^2(J - I)$
2. Apply **classical multi-dimensional scaling** to E *Retrieve coordinates from a Euclidean distance matrix* $k = n$
3. Obtain n points (the centers)
4. Apply standard JL on the centers

Multi-dimensional Scaling (MDS)

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \quad \Lambda' = \text{Diag}(\lambda_1, \dots, \lambda_k, 0, \dots, 0)$$

WLOG $\lambda_1 \geq \lambda_2 \dots \geq \lambda_n \geq 0$

- Given a squared dissimilarity matrix D and target dimension k
- Apply Eigen Decomposition to the gram matrix of D , $\text{Gram}(D) = U\Lambda U^T$
- **Pick $\lambda_1, \dots, \lambda_k$** , throw the rest eigenvalues away
- The dimension- k embedding is $X = \sqrt{\Lambda'}U^T$
- The projected distance matrix $\hat{D} = X^T X$

JL Lemma in Non-Euclidean Geometry

Find a second connection to Euclidean — Pseudo-Euclidean Space

Definition: A Pseudo-Euclidean space of signature (p, q) is a n -dimensional

vector space equipped with (p, q) bilinear form $\langle u, v \rangle_{p,q} = \sum_{i=1}^p u_i v_i - \sum_{i=p+1}^{p+q} u_i v_i$

- Standard inner product: $p = n, q = 0$
- Some other special cases of (p, q) are recognized
 - The Lorentzian n -space has signature $(n - 1, 1)$
 - 4-dim Lorentzian space is the Minkowski space, which forms the basis of special relativity:
 $\Delta s^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$
 - Vectors in $\mathbb{R}^{n+1}$ with Lorentzian inner product of -1, captures the **hyperboloid model of Hyperbolic n -space**

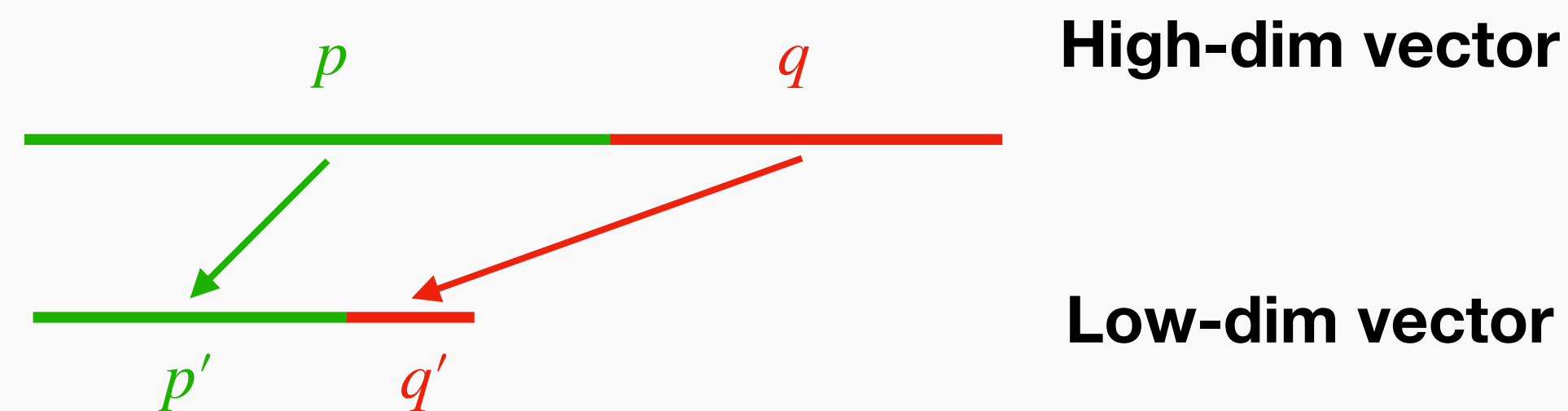
JL Lemma in Non-Euclidean Geometry

Find a second connection to Euclidean — Pseudo-Euclidean Space

Proposition: For any symmetric hollow matrix D , there is a symmetric bilinear form $\langle \cdot, \cdot \rangle_{p,q}$ and vectors $x_1, \dots, x_n$ such that

$$D_{ij} = \langle x_i - x_j, x_i - x_j \rangle_{p,q} \quad (D_{ij}^E)^2 = \langle x_i - x_j, x_i - x_j \rangle$$

- This says, D is a distance matrix of some vectors in Pseudo-Euclidean space with a certain signature (p, q)
- The JL algorithm:
 - Apply standard JL separately to the p -subspace and q -subspace



JL Lemma in Non-Euclidean Geometry

Find a second connection to Euclidean — Pseudo-Euclidean Space

Theorem: Given $D, \varepsilon \in (0,1)$, there exists a map to dimension $O(\log n/\varepsilon^2)$, s.t. every D_{ij} is preserved within $(1 \pm \varepsilon \cdot C_{ij})$ multiplicative factor, where C_{ij} is the **ratio** between squared Euclidean distance and D_{ij}

Parameter C_{ij} can be interpreted as measure of **deviation from Euclidean geometry**

Performance evaluated on the relative error $|D_{ij} - \hat{D}_{ij}|/D_{ij}$

Method	Simplex	Ball	Brain	Breast	Renal	MNIST	CIFAR10	Email	Facebook	MOOC
JL Max	6.47e5	5.97e5	1.02e12	7.09e10	3.42e12	inf	inf	3.85e4	1.18e6	2.94e5
JL-PE Max	1.11e6	5.31e5	8,21e4	262.20	2.37e4	8.47e5	2.24e6	3.24e6	2.51e7	1.35e7
JL-Power Max	109.18	12.102	3.11e7	1.92e6	1.17e8	85.76	55.74	781.45	82.74	24.22
JL Ave	442.65	43.78	1.11e9	9.86e7	1.50e9	inf	inf	15.64	12.62	39.83
JL-PE Ave	47.59	23.71	8.16	1.15	6.60	15.93	18.24	13.79	52.52	63.44
JL-Power Ave	13.52	1.002	3.73e4	3.12e3	5.57e4	1.47	1.61	1.60	1.32	2.04
JL Median	6.36	6.68	2.69e12	1.95e11	1.18e12	1.78	2.29	7.00	5.08	6.13
JL-PE Median	2.57	2.78	1284.06	174.74	3160.58	2.11	2.04	2.06	5.83	6.03
JL-Power Median	12.79	1.00	1.01e7	8.20e6	7.57e7	1.30	1.35	1.35	1.16	1.99

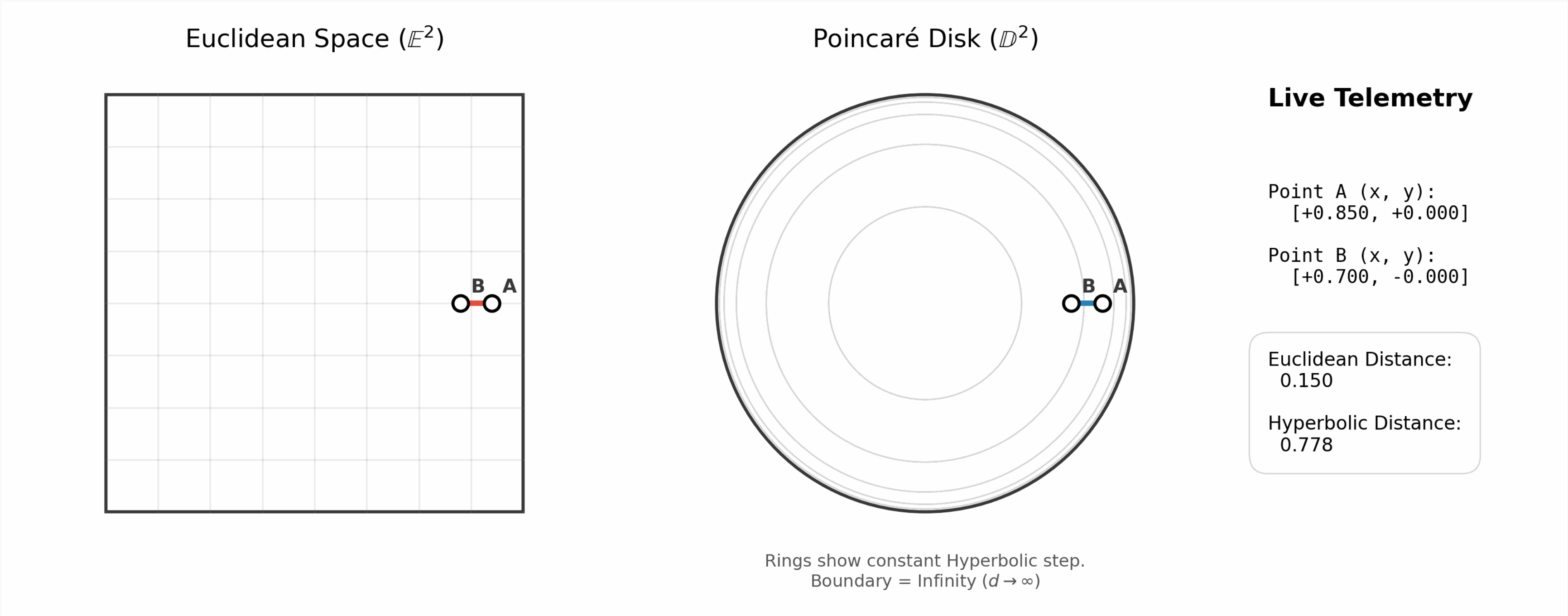
Hyperbolic Space

Hyperbolic space $\mathbb{H}^d$

A simply connected, negatively curved d -dimensional Riemannian manifold

“ We understand new concepts by relating them to familiar ones.”

— Henri Poincaré



Geodesics in Euclidean
and Hyperbolic space

Hyperbolic Space

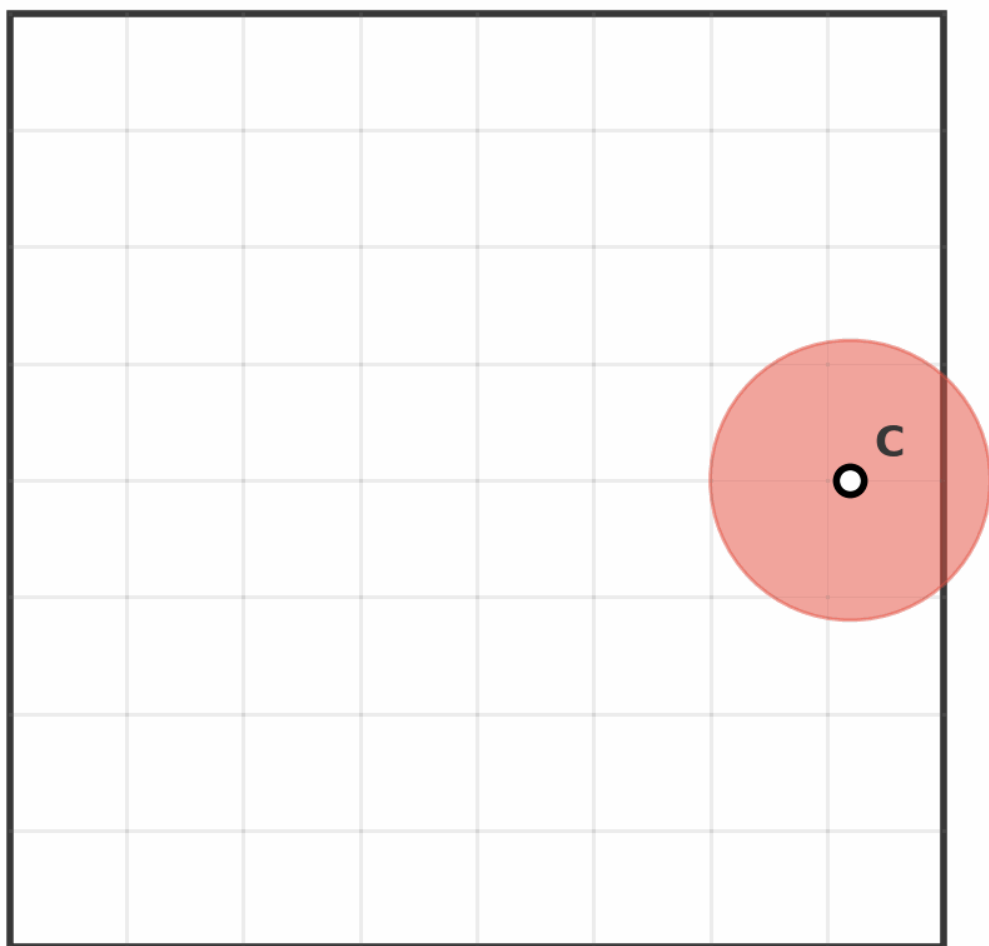
Hyperbolic space $\mathbb{H}^d$

A simply connected, negatively curved d -dimensional Riemannian manifold

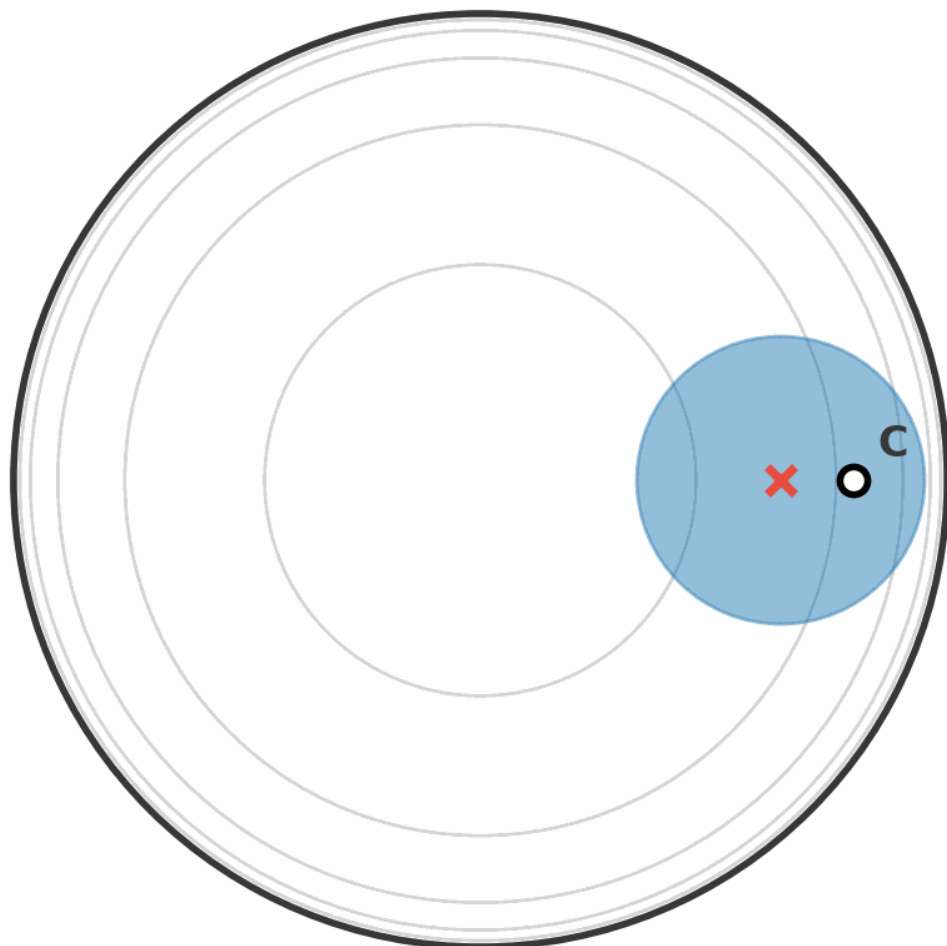
“ We understand new concepts by relating them to familiar ones.”

—Henri Poincaré

Euclidean Ball ($\mathbb{E}^2$)



Hyperbolic Ball in $\mathbb{D}^2$



Constant Hyperbolic Radius ($R_H = 1.5$)

Live Telemetry

Center 'C' (x, y):
[+0.800, +0.000]

Euclidean Offset:
0.157

Hyperbolic Radius:
1.500 (Constant)

Apparent Euclidean:
0.308 (Shrinking)

Balls in Euclidean and
Hyperbolic space

Hyperbolic Space

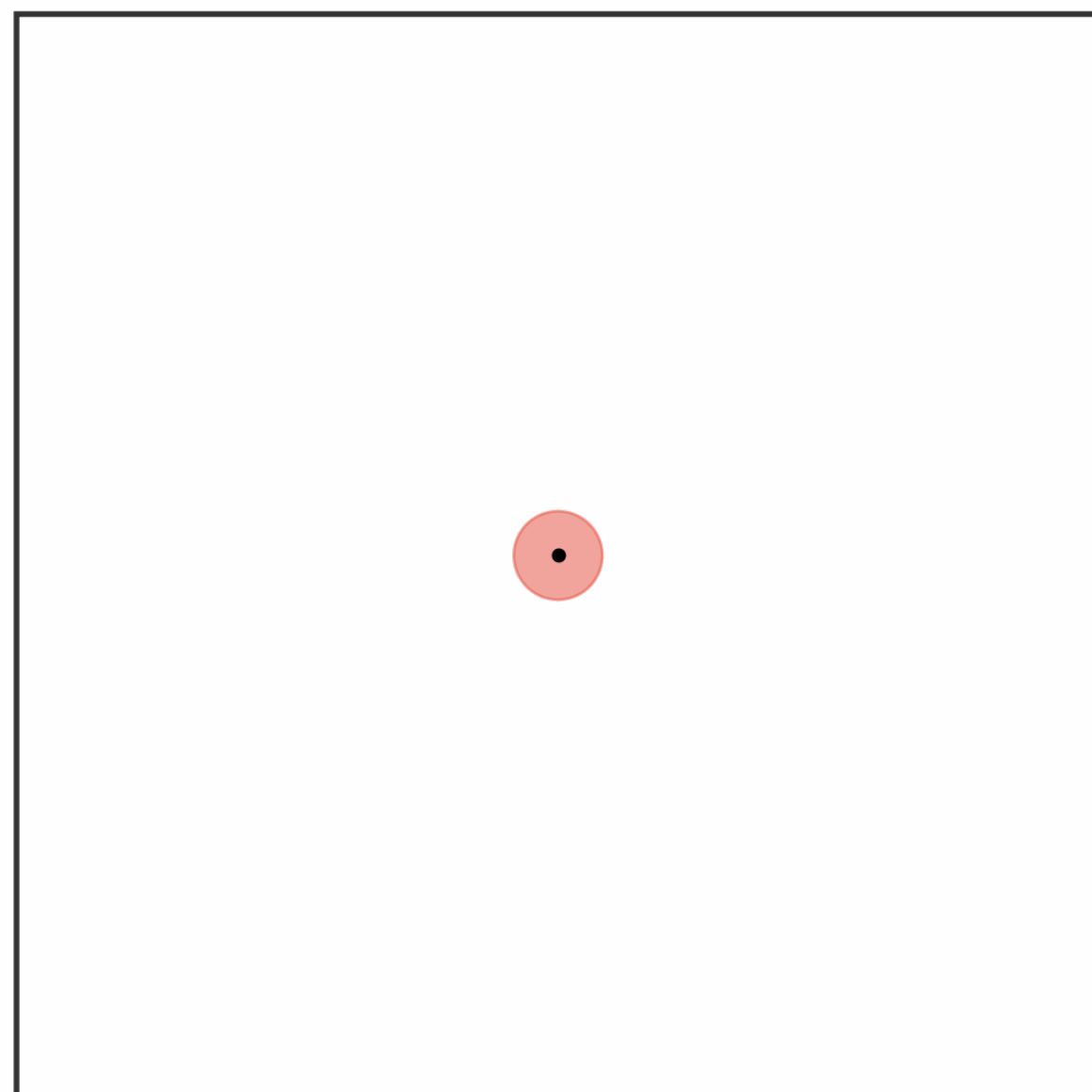
Hyperbolic space $\mathbb{H}^d$

A simply connected, negatively curved d -dimensional Riemannian manifold

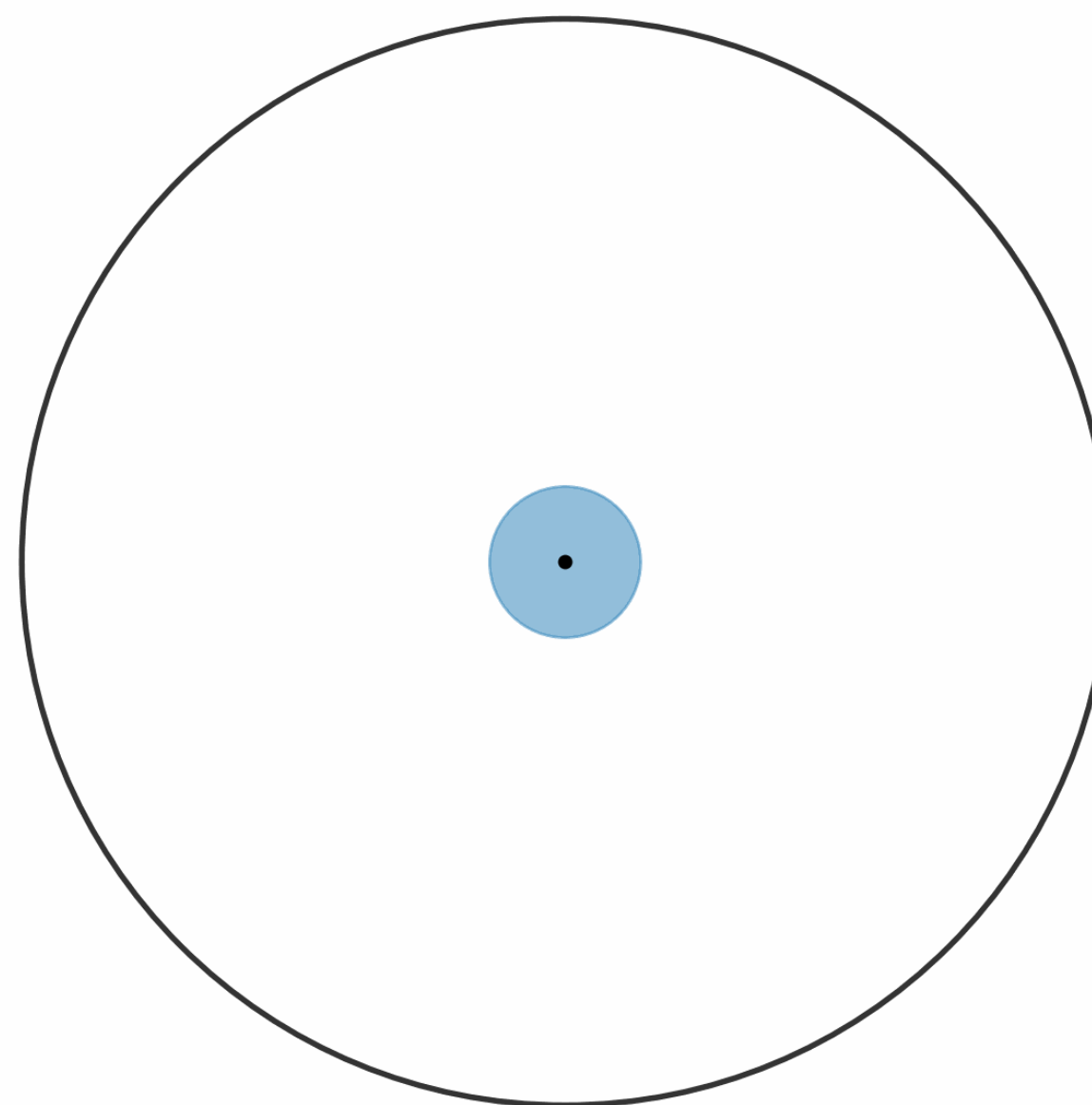
“ We understand new concepts by relating them to familiar ones.”

— Henri Poincaré

Euclidean $\mathbb{E}^2$
Polynomial Space ($V \sim r^2$)



Hyperbolic $\mathbb{D}^2$
Exponential Space ($V \sim e^r$)



Volume growth and tree
packing in Euclidean
and Hyperbolic space

Dimension Reduction in Hyperbolic Space

Good news: JL Lemma holds with 1 extra dimension [DGLLX'26]

Theorem: For a point set P in $\mathbb{H}^d$ of size n , there exists a projection to P' in $\mathbb{H}^k$ where $k = \Theta(\log n/\varepsilon^2) + 1$, with pairwise distances preserved up to $1 \pm \varepsilon$ multiplicative factor.

The projection is based on the upper half-space model

- Each point in $\mathbb{H}^d$ is written as (z, x) where $z \in \mathbb{R}^+$ and $x \in \mathbb{R}^{d-1}$
- x coordinates are Euclidean “subspace”, z coordinate is the “depth”, infinity at $z \rightarrow 0$
- Projection: keep z coordinate, apply JL transform to x coordinates
- Proposed by prior work [BM'07], we refine the analysis

Locality Sensitive Hashing in Hyperbolic Space

Nearest Neighbor Search

- **Definition:** Given a metric space (X, d) , a point set P of size n , a query input $q \in X$, find $p \in P$ such that $d(p, q) = \min_{x \in P} d(x, q)$
- The major goal: Fast query time and small storage
- Exact solution does not allow for efficient algorithms

Approximate Nearest Neighbor Search

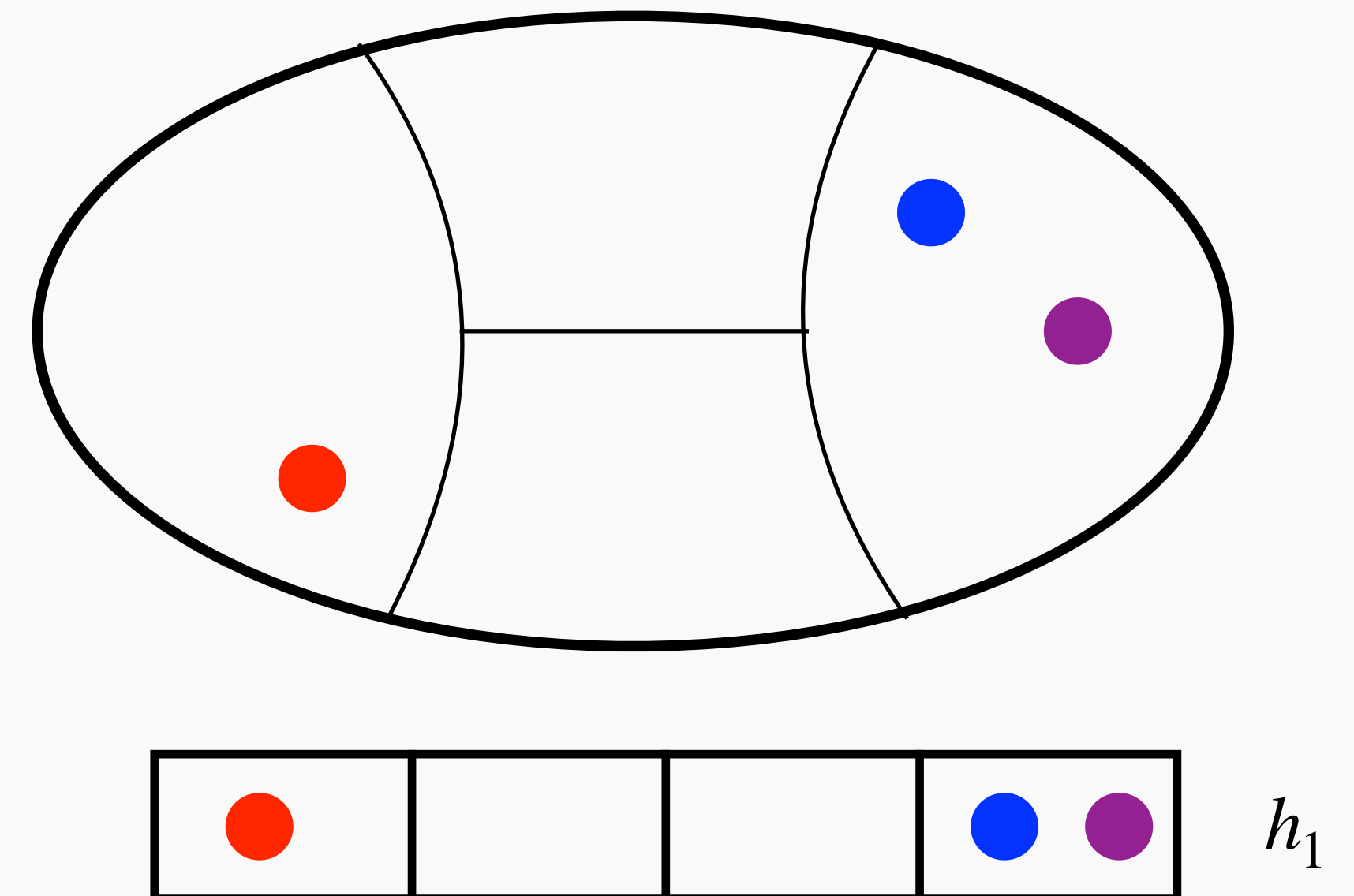
- **Definition:** Given a metric space (X, d) , a point set P of size n , a query input $q \in X$, find $p \in P$ such that $d(p, q) \leq c \cdot \min_{x \in P} d(x, q)$
- Locality Sensitive Hashing (LSH) is one popular technique to solve the ANNS problem

Locality Sensitive Hashing in Hyperbolic Space

Locality Sensitive Hashing

Given a point set P in a metric space, LSH maps each point into a bucket

1. Close points are likely in the same bucket
2. Far points are likely in the different buckets

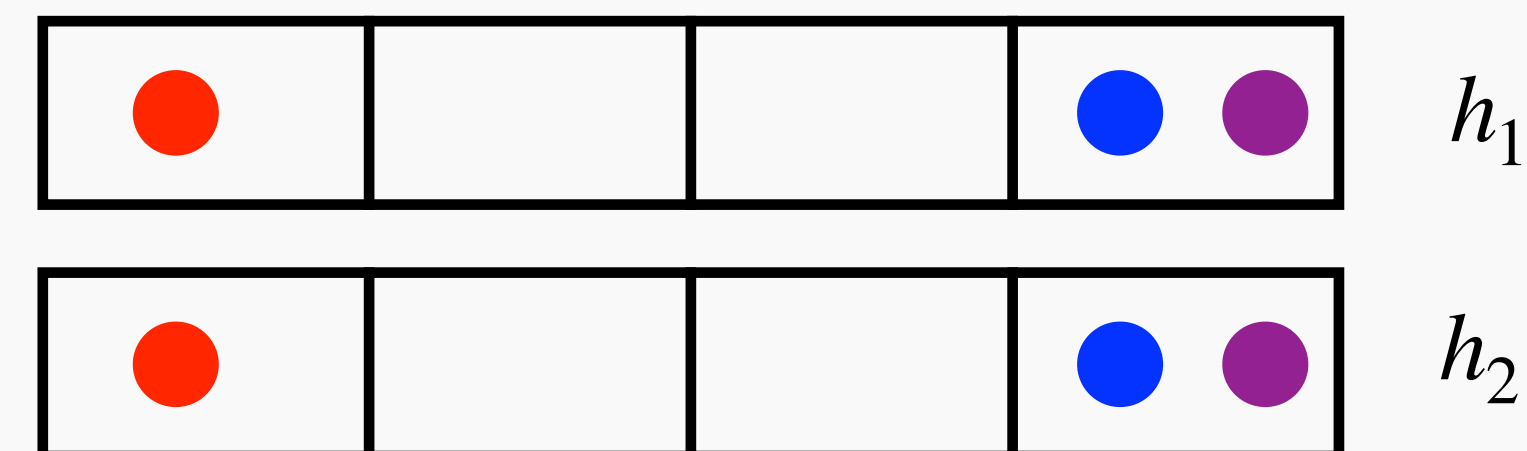
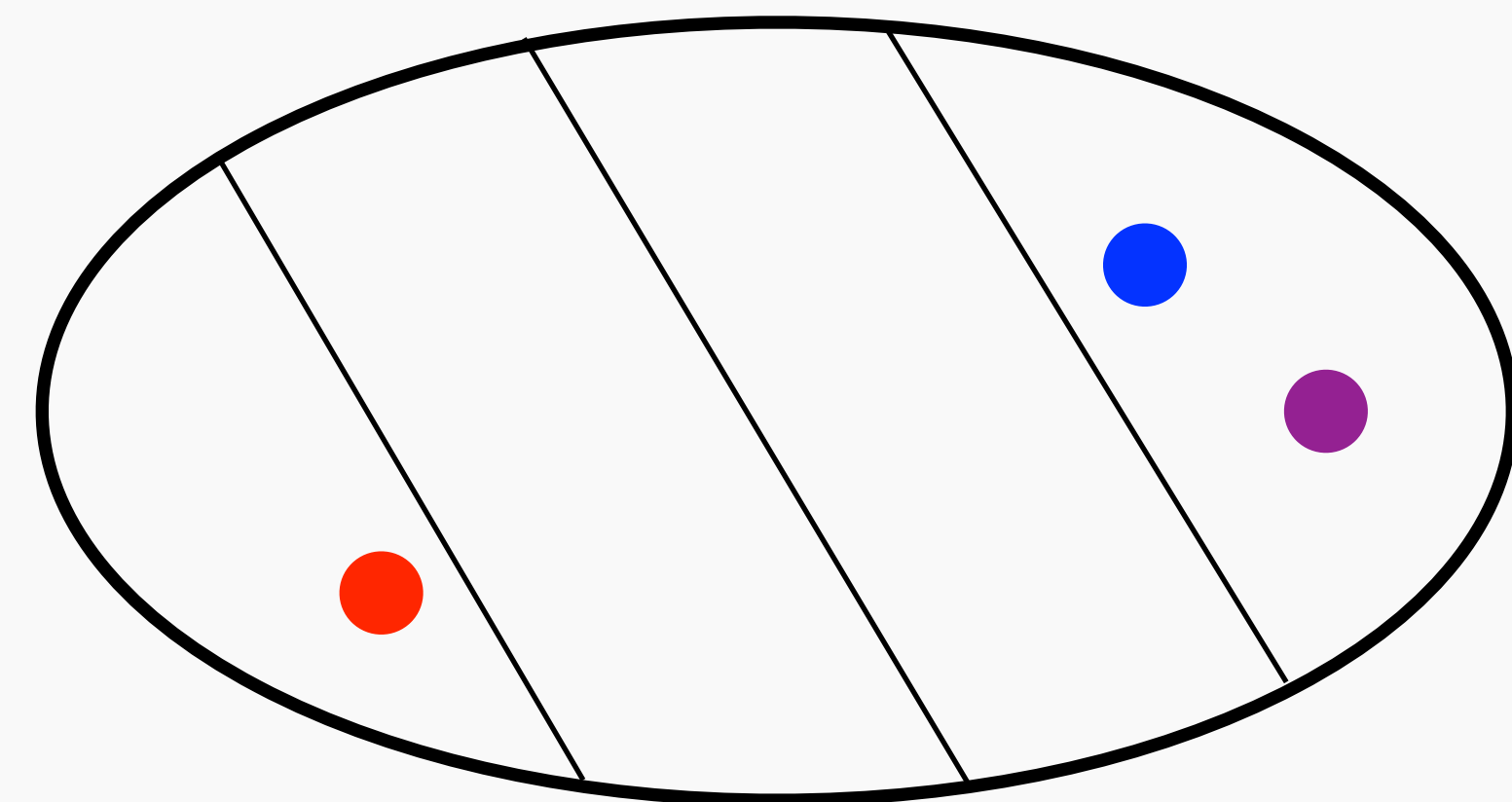


Locality Sensitive Hashing in Hyperbolic Space

Locality Sensitive Hashing

Given a point set P in a metric space, LSH maps each point into a bucket

1. Close points are likely in the same bucket
2. Far points are likely in the different buckets

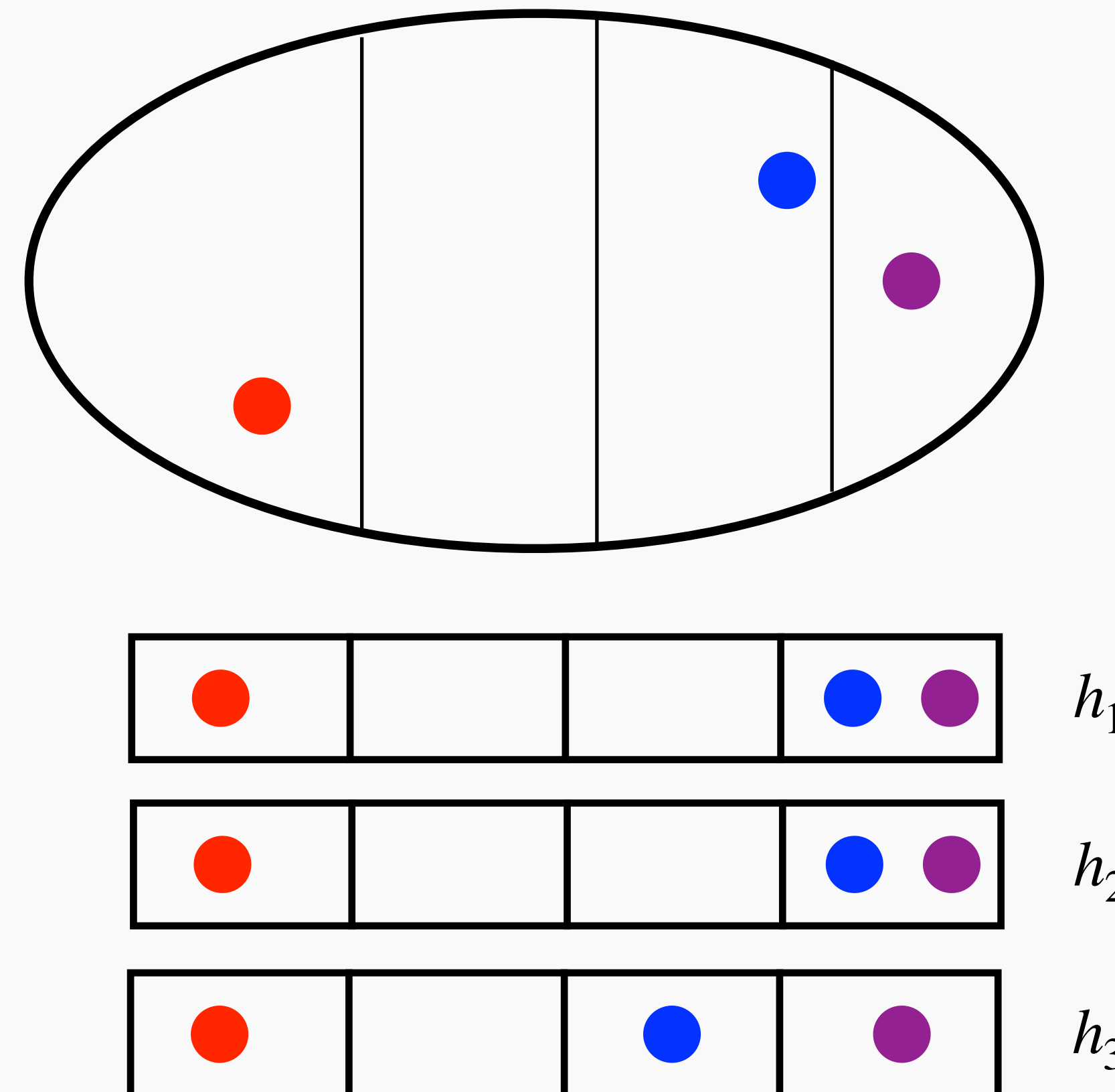


Locality Sensitive Hashing in Hyperbolic Space

Locality Sensitive Hashing

Given a point set P in a metric space, LSH maps each point into a bucket

1. Close points are likely in the same bucket
 2. Far points are likely in the different buckets
- LSH is essentially a random space partition scheme

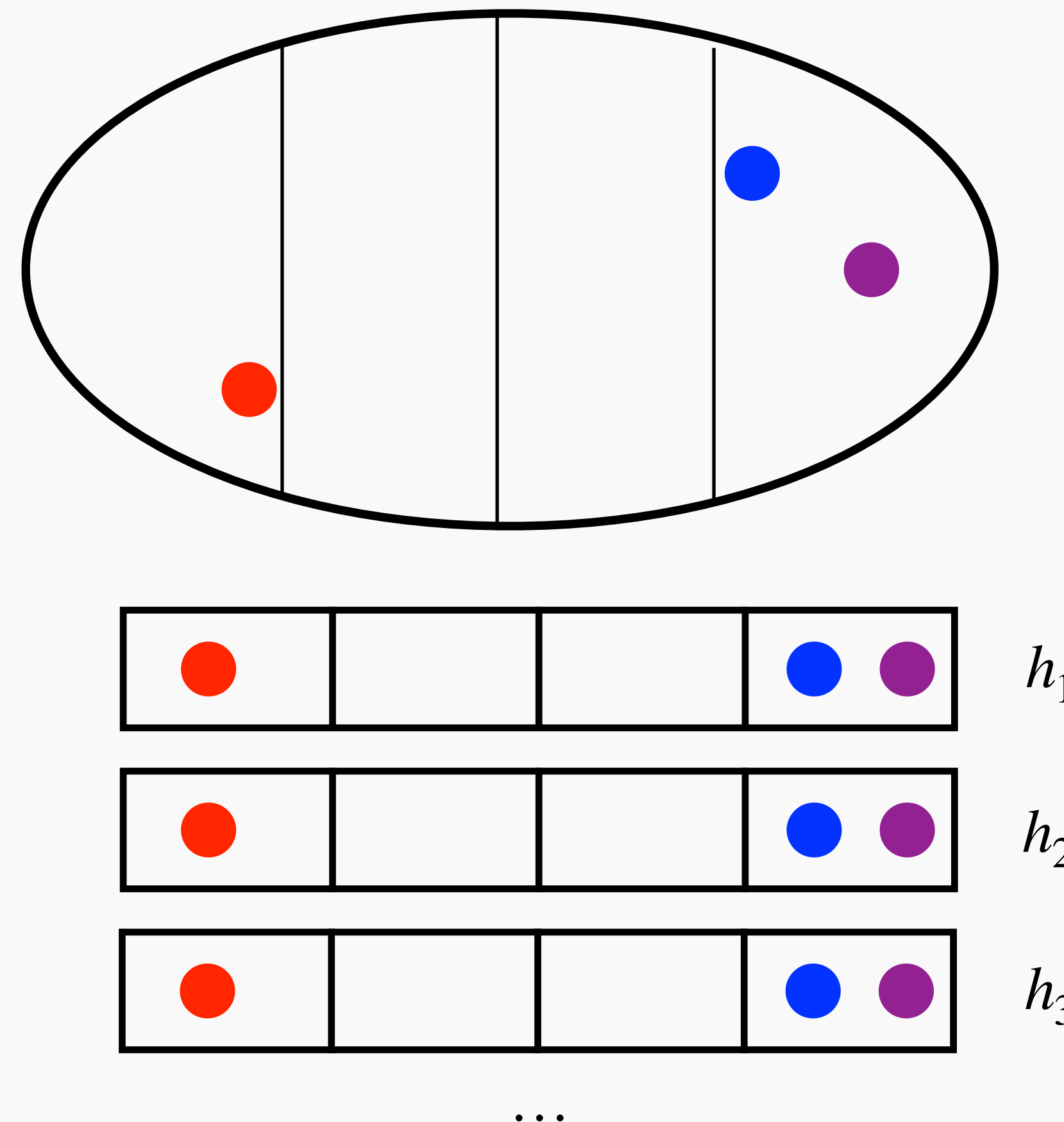


Locality Sensitive Hashing in Hyperbolic Space

Locality Sensitive Hashing

Given a point set P in a metric space, LSH maps each point into a bucket

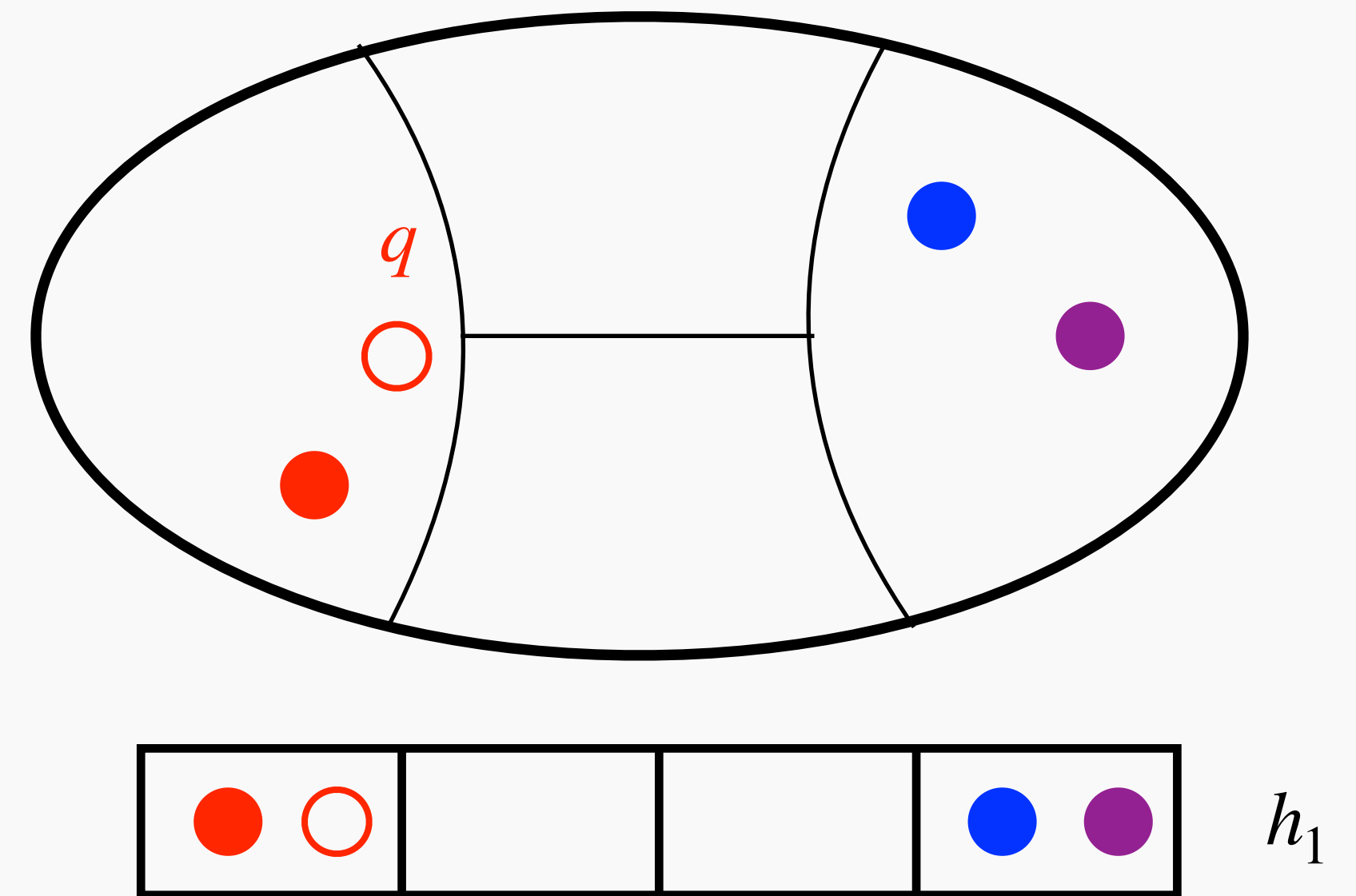
1. Close points are likely in the same bucket 😊
 2. Far points are likely in the different buckets
- LSH is essentially a random space partition scheme
 - Repetition of hashing helps
 - We should construct a family of hashing functions



Locality Sensitive Hashing in Hyperbolic Space

Given a metric space (X, d) , let U be a countable set and $H : X \rightarrow U$ be a family of functions. We say H is a (r, cr, p_1, p_2) -sensitive family, where $r > 0$, $c > 1$, and $0 < p_2 < p_1 < 1$, such that for any $x, y \in \mathcal{X}$:

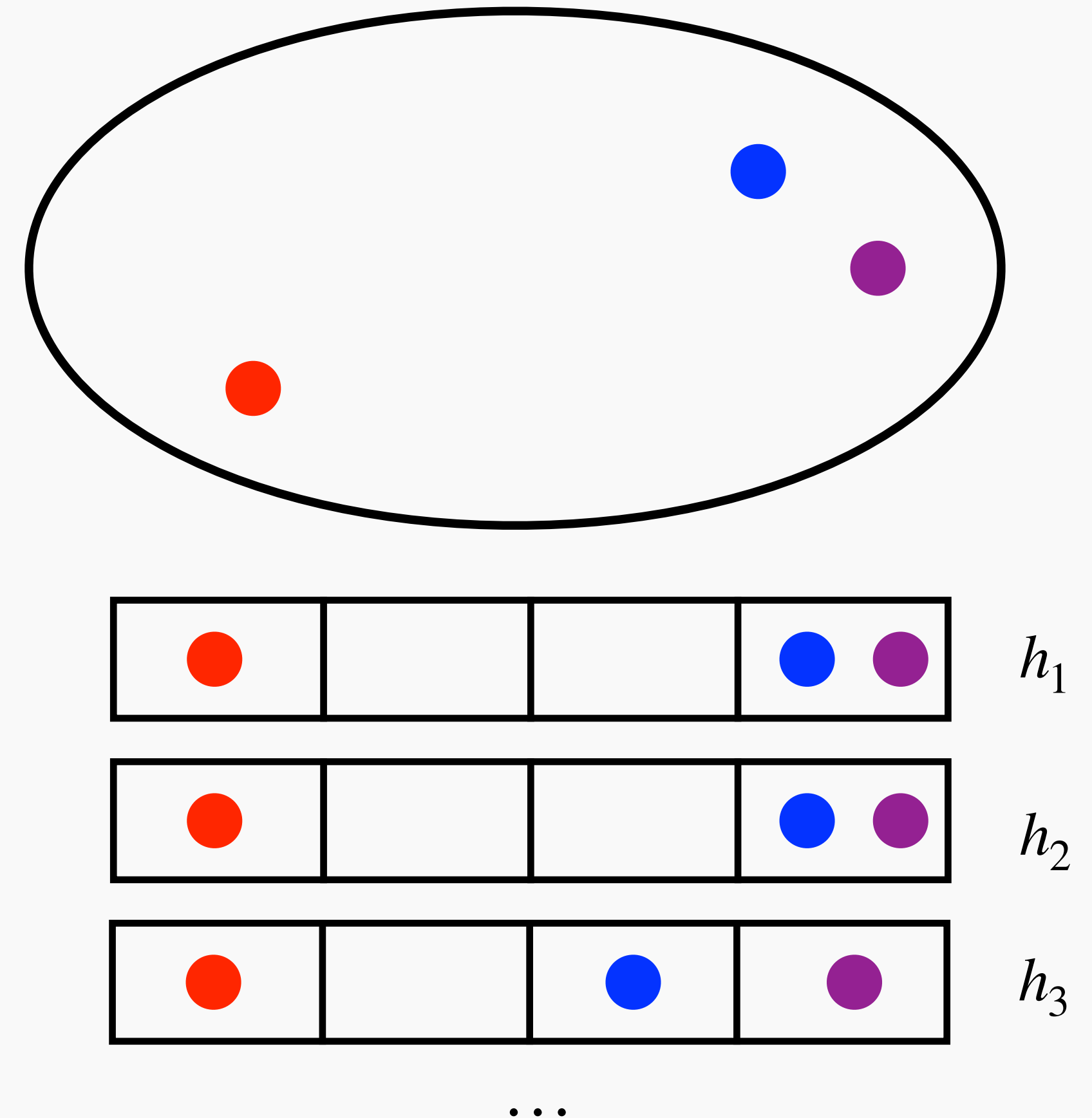
1. If $d(x, y) \leq r$, $\Pr_{h \sim H} [h(x) = h(y)] \geq p_1$
 2. If $d(x, y) \geq cr$, $\Pr_{h \sim H} [h(x) = h(y)] \leq p_2$
- LSH solves the ANNS problem, with $p_1 \approx 1$, $p_2 \approx 0$



Locality Sensitive Hashing in Hyperbolic Space

Given a metric space (X, d) , let U be a countable set and $H : X \rightarrow U$ be a family of functions. We say H is a (r, cr, p_1, p_2) -sensitive family, where $r > 0$, $c > 1$, and $0 < p_2 < p_1 < 1$, such that for any $x, y \in \mathcal{X}$:

1. If $d(x, y) \leq r$, $\Pr_{h \sim H} [h(x) = h(y)] \geq p_1$
 2. If $d(x, y) \geq cr$, $\Pr_{h \sim H} [h(x) = h(y)] \leq p_2$
- LSH solves the ANNS problem, with $p_1 \approx 1$, $p_2 \approx 0$
 - How many hash tables are needed?
 $O(n^\rho)$, where $\rho = \log p_1 / \log p_2$
 - Query time: $O(d \cdot n^\rho)$; Space: $O(n^{1+\rho} + dn)$
 ρ is the **quality measure** of LSH, as function of c



Locality Sensitive Hashing in Hyperbolic Space

Our result [DGLX SoCG'26]

Theorem: There exists an LSH family in $\mathbb{H}^2$ that achieves $\rho \leq 1/c$; Further, there exists an LSH family in $\mathbb{H}^d$ that achieves $\rho \leq 1.59/c$

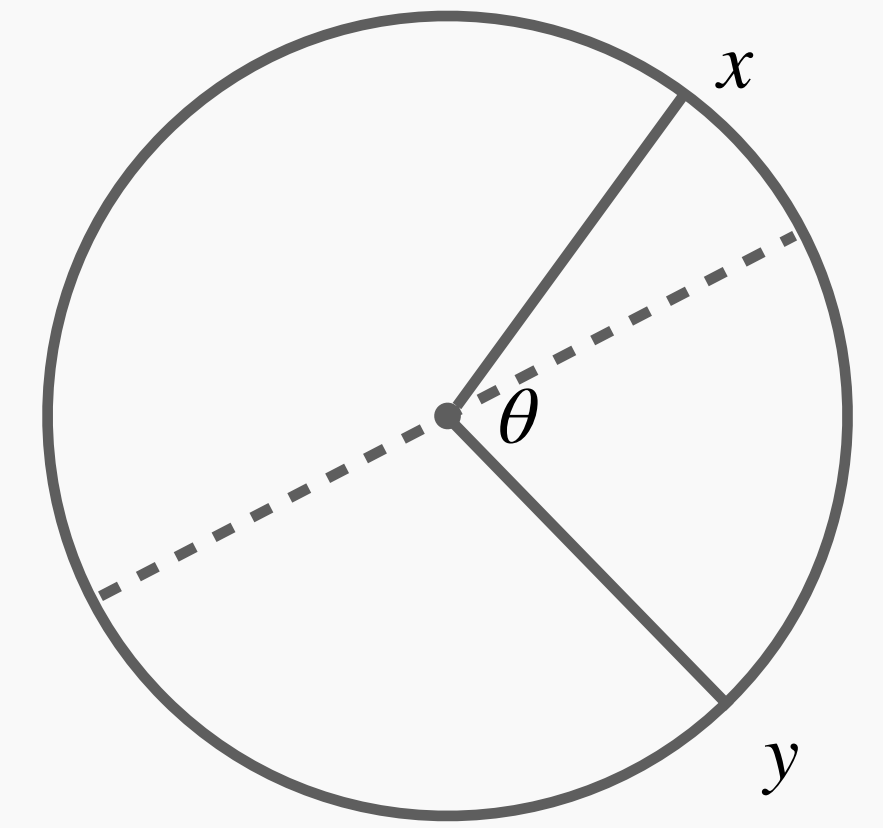
Remarks

- The lower bound follows the Euclidean hardness, $\rho \geq 1/c^2$
- The technical novelty appears in $\mathbb{H}^2$, where we use the **random hyperplane rounding**
- For high-dimension, we apply dimension reduction to $\mathbb{H}^2$ first
- The dimension reduction imposes a constant distortion, therefore leads to slightly worse bound on ρ
- We did not **explore “new” properties**; We try to find a natural geodesic separator that **respects** the negative curvature and exponential radial growth of hyperbolic geometry

Locality Sensitive Hashing in $\mathbb{H}^2$

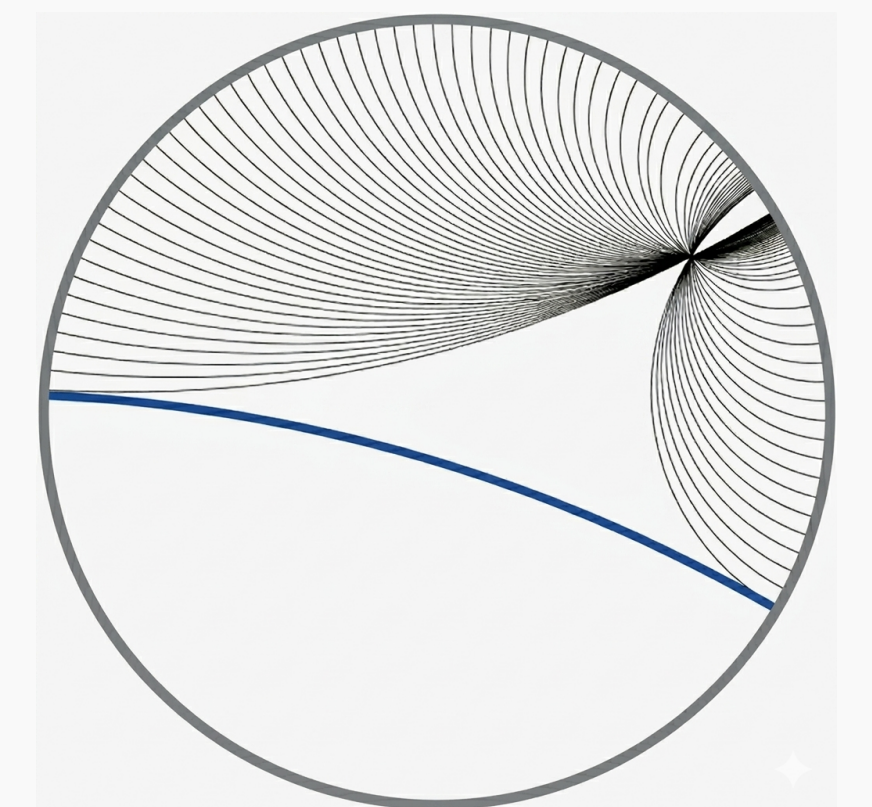
Random hyperplane rounding

- Uniformly at random sample a **straight line** passing through the origin
- If x, y lie on different sides, assign $h(x) = 1, h(y) = 0$
- Otherwise, assign $h(x) = h(y)$
- Studied before for the max cut problem, LSH for angular distance, etc.



We consider the Poincaré disk model

- Defined on $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$
- Straight lines are **circular arcs orthogonal** to the unit circle
- Intuition: if x, y are far in $\mathbb{H}^2$, a random geodesic likely separates them



Locality Sensitive Hashing in $\mathbb{H}^2$

Proposed LSH function

- Uniformly at random sample a random geodesic γ
- Hash a point x as $h_\gamma(x) \in \{+1, -1\}$ according to which side of γ it lies on

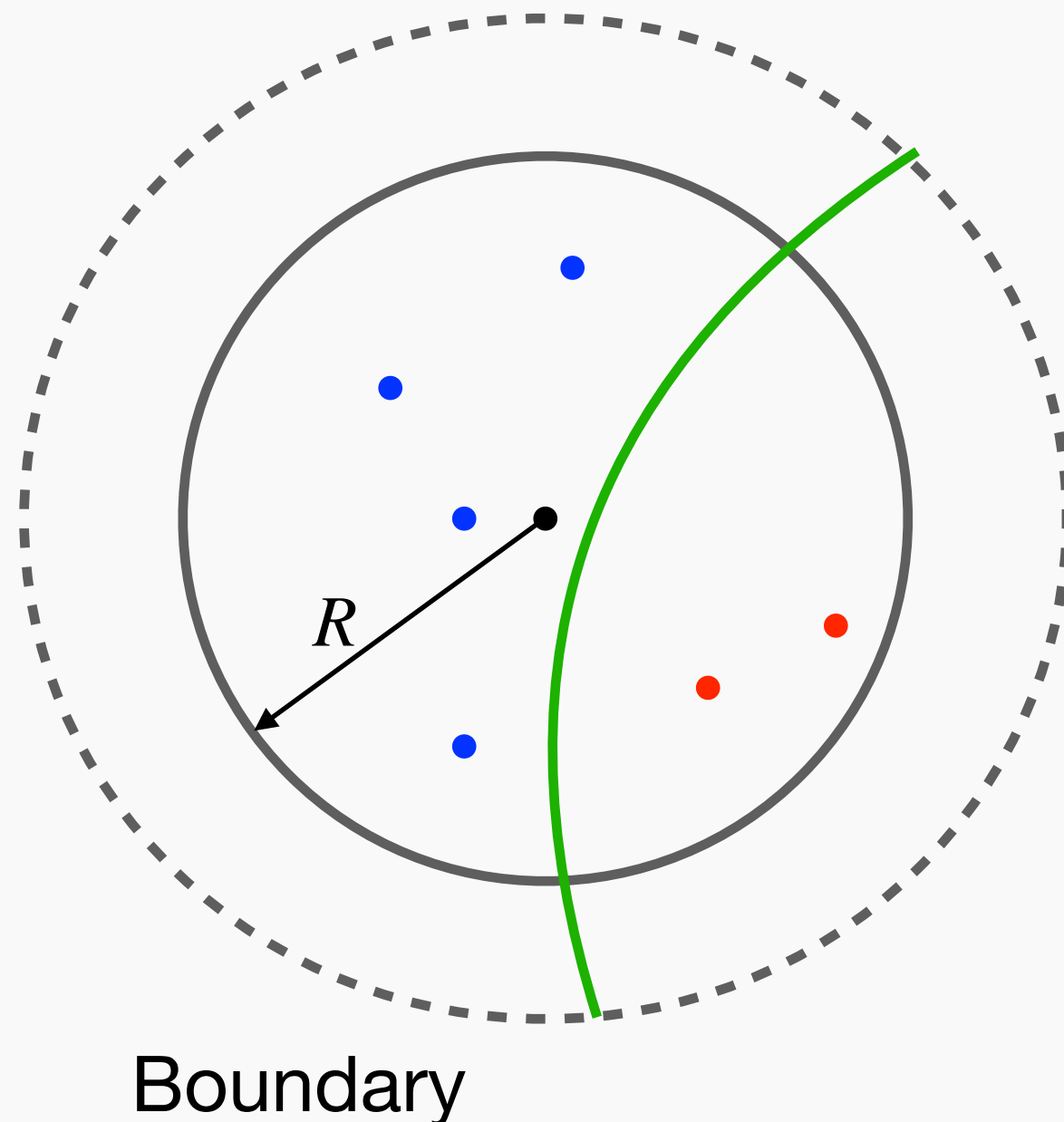
⚠ Is there something off?

1. To sample, we need a **isometry-invariant measure** on the set of all geodesics
 - We do have one, kinematic measure $d\mu(t, \theta) = \cosh(t) dt d\theta$
2. The space of all geodesics has **infinite total measure**
 - We use a trick by assuming the data is bounded by radius R
 - R is a constant and gets cancelled later
 - Analysis relies on the **Crofton's formula** from integral geometry

Locality Sensitive Hashing in $\mathbb{H}^2$

Proposed LSH function

- Let R be a radius such that the point set is contained in the ball $B(0, R)$.
- Let μ_R be the **normalized** kinematic measure of the set of geodesics intersecting $B(0, R)$
- Uniformly at random sample a random geodesic $\gamma \sim \mu_R$
- Hash a point x as $h_\gamma(x) \in \{+1, -1\}$ according to which side of γ it lies on



Key lemma (collision probability):

$$\Pr[h(x) = h(y) \mid d_{\mathbb{H}}(x, y) = r] = 1 - \frac{r}{\pi \sinh(R)}$$

$$\bullet p_1 = 1 - \frac{r}{\pi \sinh(R)}, \quad p_2 = 1 - \frac{cr}{\pi \sinh(R)}$$

$$\bullet \rho = \frac{\log p_1}{\log p_2} \leq \frac{1}{c}$$

Summary

In the generic non-Euclidean setting

- Almost no properties left to exploit
- Our way, basically reduces to finding a connection to Euclidean
 - Generalized power distances
 - Pseudo-Euclidean space with (p, q) signature

In the hyperbolic geometry

- We can still find some Euclidean techniques useful
- We can at least respect the known hyperbolic structure while designing algorithms
- Recent works show it is promising to have **different tools at different scales**
- In general, we are short of tools for negatively-curved geometry

Thank you!