When LSH Breaks: Failure of Scaled Projections and Signed-Permutation Hashing in the ℓ_p-Sphere

Heon Lee **Manas Korimilli** **Christopher Jeong**
Department of Computer Science
Brown University
Providence, RI 02912
{heon_lee, manas_korimilli, christopher_jeong}@brown.edu

Abstract

In this work, we explore two natural generalizations of the cross-polytope locality-sensitive hashing scheme of Andoni *et al.* [2] from the Euclidean sphere to arbitrary ℓ_p-spheres: (1) a *scaled-rotation* approach that rotates points via random orthogonal matrices, reprojects them to the ℓ_p-sphere, and quantizes to the dominant coordinate; and (2) a *hyperoctahedral sampling* approach that replaces the orthogonal group by the symmetry group of the ℓ_p-sphere. We prove that both families fail for every $p \neq 2$: when $p < 2$ the scaled-rotation scheme allows arbitrarily close points to collide with vanishing probability, while for $p > 2$ it forces far-apart points to collide with probability above a threshold—so no nontrivial (r, cr)-sensitivity holds—and that the hyperoctahedral scheme induces exactly one partition (up to relabeling), making collision probabilities independent of distance. These negative results precisely delineate the limits of direct cross-polytope generalizations for ℓ_p-spheres with $p \neq 2$, providing insights into the challenges of hashing techniques in the ℓ_p-sphere.

1 Introduction

Approximate nearest-neighbor (ANN) search is a technique in large-scale machine learning, data mining, and information retrieval. In high dimensions, *locality-sensitive hashing* (LSH) provides a powerful framework: one designs a family of hash functions so that "nearby" points collide with higher probability than "far apart" points, and then uses these collisions to achieve sublinear-time queries [6]. LSH is well understood for Hamming distance and the Euclidean (ℓ_2) norm: for angular distance on the unit sphere, the cross-polytope scheme of Andoni *et al.* [2] achieves the asymptotically optimal exponent

$$\rho = \frac{1}{2c^2 - 1},$$

matching known lower bounds [1]. Moreover, it is practical—empirical evaluations show substantial speedups over classic hyperplane-based hashes in real-world datasets. Given its strong theoretical guarantees and empirical performance on the Euclidean sphere, one might reasonably expect that simply swapping in an ℓ_p-projection or replacing $O(n, \mathbb{R})$ with the appropriate symmetry group for the general ℓ_p-sphere would preserve locality-sensitive behavior. Surprisingly, as this paper shows, these most natural extensions break down completely for every $p \neq 2$, revealing deeper geometric barriers to non-Euclidean LSH.

In this work we investigate two *natural* attempts to extend the cross-polytope LSH from the Euclidean sphere $\mathbb{S}_2^{n-1}$ to the general ℓ_p-sphere $\mathbb{S}_p^{n-1}$:

1. **Scaled-rotation:** apply a random orthogonal matrix $A \in O(n, \mathbb{R})$ to $x \in \mathbb{S}_p^{n-1}$, renormalize back to the ℓ_p-sphere, and then quantize to the largest-coordinate corner of the ℓ_1 ball;

2. **Hyperoctahedral sampling:** replace $O(n, \mathbb{R})$ by the symmetry group of $\mathbb{S}_p^{n-1}$, and quantize directly without any continuous reprojection.

Both approaches mirror the Euclidean construction, preserve the spirit of random mixing, and respect the natural symmetries of the ℓ_p sphere. However, in contrast to the $p = 2$ case, we prove that *neither* family can satisfy any nontrivial LSH sensitivity for *any* $p \neq 2$.

Our contributions. We establish three impossibility results:

- For $p < 2$, the scaled-rotation scheme fails to bring nearby points into collision with non-vanishing probability; in fact, one can exhibit pairs at arbitrarily small ℓ_p distance whose collision probability tends to zero as the dimension grows.
- For $p > 2$, the same construction makes arbitrarily distant points collide with probability above a threshold —again precluding any valid (r, cr)-sensitivity trade-off.
- The hyperoctahedral scheme is even more degenerate: every signed-permutation map induces the same partition of $\mathbb{S}_p^{n-1}$ (up to relabeling), so collision probabilities become independent of distance (either 0 or 1).

These impossibility results rule out only the most direct extensions of cross-polytope LSH to ℓ_p spheres for $p \neq 2$, and leave open the prospect that more sophisticated or data-adaptive hashing schemes may still succeed in this setting.

Paper organization. In Section 2 we review LSH basics and the Euclidean cross-polytope construction. Section 3 defines the scaled-rotation and hyperoctahedral hash families on $\mathbb{S}_p^{n-1}$, and presents our counterexample constructions to prove the three impossibility theorems. We conclude in Section 4 with a discussion of open directions toward non-Euclidean LSH schemes.

2 Background and Related Work

Locality-Sensitive Hashing (LSH). Locality-sensitive hashing (LSH) circumvents the curse of dimensionality in nearest-neighbor search by trading off space for query time. For a metric space $(\mathbb{R}^n, d)$ and parameters $r > 0$, $c > 1$, a hash family $\mathcal{H}$ is (r, cr, p_1, p_2)-*sensitive* if for all $x, y \in \mathbb{R}^n$:

$$\begin{aligned} \mathbb{P}_{h\sim\mathcal{H}}(h(x) = h(y)) \;&\geq\; p_1 \quad \text{whenever } d(x, y) \leq r, \\ \mathbb{P}_{h\sim\mathcal{H}}(h(x) = h(y)) \;&\leq\; p_2 \quad \text{whenever } d(x, y) \geq cr, \end{aligned}$$

with $p_1 > p_2$. Such a family yields a data structure for (c, r)-ANN with sublinear query time [6]. Classic LSH constructions exist for Hamming space [6], Jaccard similarity [3], the Euclidean (ℓ_2) norm [2], the cosine similarity and earthmover distance via the hyperplane LSH [4], and the ℓ_p norm via p-stable projections [5].

LSH for Cosine Similarity. Charikar's *hyperplane LSH* [4] hashes by random sign of dot-products, giving collision probabilities tied to cosine similarity. More recent cosine similarity LSH schemes—most notably the *cross-polytope* family of Andoni *et al.* [2]—rotate points randomly and quantize to the nearest vertex of an ℓ_1-sphere, achieving the asymptotically optimal exponent for angular distance. Multiprobe and fast-rotation variants further improve practical performance.

Our contributions. This work shows that two of the most *natural* extensions of Euclidean cross-polytope LSH to the ℓ_p sphere *both fail* for every $p \neq 2$. We analyze:

- **Scaled-rotation:** randomly rotate any ℓ_p-unit vector, project back to the ℓ_p sphere, then quantize to the dominant coordinate.
- **Hyperoctahedral sampling:** replace continuous rotations by the finite signed-permutation group (the hyperoctahedral group) and quantize directly.

We prove that for $p < 2$ the scaled-rotation scheme yields vanishing collision probability on arbitrarily close points, and for $p > 2$ it forces distant points to collide with probability above a threshold; and that the hyperoctahedral scheme induces a single degenerate partition independent of distance. These negative results rule out only the simplest generalizations of cross-polytope hashing and leave open the search for more sophisticated or data-adaptive LSH constructions in non-Euclidean norms.

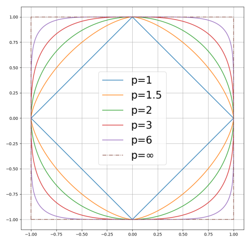

Figure 1: **Unit ℓ_p-spheres for $p \in \{1, 1.5, 2, 3, 6, \infty\}$ in $\mathbb{R}^2$.**

3 Generalized Cross-Polytope Hash Families

Let $\mathbb{S}_p^{n-1}$ be the ℓ_p-sphere in n-dimension. Examples of $\mathbb{S}_p^1$ for varying values of p can be found in Figure 1. Andoni *et al.* [2] propose a cross-polytope LSH scheme on the Euclidean sphere $\mathbb{S}_2^{n-1}$. Given a random orthogonal matrix $A \in O(n, \mathbb{R})$, they define the hash function

$$h_A := \mu \circ r_A|_{\mathbb{S}_2^{n-1}} : \mathbb{S}_2^{n-1} \to \mathbb{S}_1^{n-1}$$

where $r_A : \mathbb{R}^n \to \mathbb{R}^n$ applies A to the vector $x \in \mathbb{R}^n$ and the quantizer $\mu : \mathbb{R}^n \setminus \{0\} \to \mathbb{S}_1^{n-1}$ is given by

$$\mu(y) = \operatorname{sgn}(y_j) e_j, \quad j = \operatorname*{argmax}_{1 \le k \le n} |y_k|,$$

and e_j is the j-th standard basis vector. Then we have that

$$\mathcal{H} := \{h_A : A \in O(n, \mathbb{R})\}$$

defines a family of hash functions.

Intuition Behind h_A

The process behind h_A can be understood in three simple steps:

1. Random Rotation: Multiplying by A "spins" the entire sphere. Nearby points stay close, but their coordinates get mixed up independently of the data.
2. Dominant Axis Selection: After rotation, inspect the resulting vector $y = Ax$. Identify which coordinate y_j has the largest absolute value. Geometrically, this finds the single axis along which the rotated point sticks out the most.
3. Snapping to a Corner: Record the sign of that dominant coordinate (positive or negative) and map y exactly to the corresponding corner of the cross-polytope, namely $\pm e_j$. Figure 2 shows such a cross-polytope for $n = 4$. This collapses the continuous sphere onto the $2n$ vertices of the ℓ_1 unit sphere.

Two points with a small angle between them will, after the same random rotation, typically share both the index and sign of their dominant coordinate, and hence collide under h_A. Points far apart on the sphere are unlikely to do so. Repeating this with multiple independent A's yields the usual LSH guarantees: high collision probability for similar points, low for dissimilar.

3.1 Generalization to the ℓ_p-Sphere

To extend this to the ℓ_p-sphere $\mathbb{S}_p^{n-1}$ for $p \neq 2$, one can modify either the rotation or the quantization step in $h_A = \mu \circ r_A$:

1. Scaled Rotation. Apply $A \in O(n, \mathbb{R})$ to $x \in \mathbb{S}_p^{n-1}$, then project back to the ℓ_p-sphere via the scaling map
$$\pi_p(y) = \frac{y}{\|y\|_p},$$

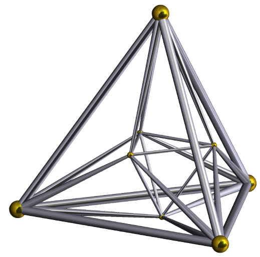

Figure 2: Cross-polytope (i.e. ℓ_1-sphere) for $n = 4$, by Robert Webb (Stella software), Wikimedia Commons (CC BY-SA 3.0) commons.wikimedia.org/wiki/File:Schlegel_wireframe_16-cell.png.

and finally quantize with μ:

$$\tilde{h}_A = \mu \circ \pi_p \circ r_A.$$

As we show in Proposition A.1, π_p is exactly the nearest-point projection onto $\mathbb{S}_p^{n-1}$ in the ℓ_p norm.

2. Hyperoctahedral Sampling. Replace the full orthogonal group by the symmetry group G_p of $\mathbb{S}_p^{n-1}$ since the orthogonal group is the symmetry group of $\mathbb{S}_2^{n-1}$. Then

$$\bar{h}_A = \mu \circ r_A$$

for $A \in G_p$.

3.2 Scaled Hash Family

Lemma 3.1. *For any $p \geq 1$,*

$$\mu \circ \pi_p = \mu.$$

Proof. It suffices to observe that scaling does not change the maximum absolute coordinate. □

Theorem 3.2. *Let $p \in [1, 2)$. Then for any $r \in (0, 2)$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ one can find $x_1, \ldots, x_m \in \mathbb{S}_p^{n-1}$ with $\|x_i - x_j\|_p \leq r$ but*

$$\mathbb{P}_{A\sim\mathrm{Uniform}(O(n,\mathbb{R}))}\big(\tilde{h}_A(x_i) = \tilde{h}_A(x_j)\big) \;\leq\; n^{-\frac{r^2}{4-r^2}+o(1)},$$

implying that

$$\mathbb{P}_{A\sim\mathrm{Uniform}(O(n,\mathbb{R}))}\big(\tilde{h}_A(x_i) = \tilde{h}_A(x_j)\big) \to 0 \text{ as } n \to \infty.$$

In particular, no fixed $p_1 > 0$ can serve as a lower bound on the collision probability for all such "nearby" pairs.

Proof. It suffices to find $x, y \in \mathbb{S}_p^{n-1}$ such that $\|x - y\|_p \leq r$ and $\mathbb{P}_{A\sim U(O(n,\mathbb{R}))}(\tilde{h}_A(x) = \tilde{h}_A(y)) \leq n^{-\frac{r^2}{4-r^2}+o(1)}$. Consider $v := \left(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right) \in \mathbb{R}^n$ and $a := (\epsilon, -\epsilon, 0, \ldots, 0) \in \mathbb{R}^n$ for $\epsilon > 0$. Then we know that

$$N_p := \|v - a\|_p = \|v + a\|_p = \left(\left|\frac{1}{n} - \epsilon\right|^p + \left|\frac{1}{n} + \epsilon\right|^p + \frac{n-2}{n^{p/2}}\right)^{1/p}.$$

Let

$$u_p^{\pm} := \frac{v \pm a}{N_p} \in \mathbb{S}_p^{n-1}.$$

Because $u_p^+ - u_p^- = \frac{v+a}{N_p} - \frac{v-a}{N_p} = \frac{2a}{N_p} = \frac{(2\epsilon, -2\epsilon, 0, \ldots, 0)}{N_p}$, then

$$\|u_p^+ - u_p^-\|_p = \left\|\frac{2a}{N_p}\right\|_p = \frac{2^{(p+1)/p}\epsilon}{\left(\left|\frac{1}{\sqrt{n}} - \epsilon\right|^p + \left|\frac{1}{\sqrt{n}} + \epsilon\right|^p + \frac{n-2}{n^{p/2}}\right)^{1/p}}.$$

For $p = 2$,

$$||u_2^+ - u_2^-||_2 = \left\| \frac{2a}{\sqrt{1+2\epsilon^2}} \right\|_2 = \frac{2\sqrt{2}\epsilon}{\sqrt{1+2\epsilon^2}}.$$

Given $r \in (0, 2)$ and $n > 1$, with $p < 2$. We want to show there exists ϵ such that $\|u_p^+ - u_p^-\|_p \le r$ and $\|u_2^+ - u_2^-\|_2 > r$. By continuity of $f_p(\epsilon) := \|u_p^+ - u_p^-\|_p$ and because $f_p(\epsilon) < f_2(\epsilon)$, we know that there exists some $\epsilon > 0$ such that $\|u_p^+ - u_p^-\|_p \le r$ and $\|u_2^+ - u_2^-\|_2 > r$.

By Lemma 3.1, we know that $\tilde{h}_A = \mu \circ \pi_p \circ r_A = \mu \circ r_A = h_A$. Then applying Theorem 1 of [2] on cross-polytope LSH gives

$$\mathbb{P}_{A\sim U(O(n,\mathbb{R}))}\big(\tilde{h}_A(u_p^+) = \tilde{h}_A(u_p^-)\big) = \mathbb{P}\big(\mu(Au_2^+) = \mu(Au_2^-)\big) \le n^{-\frac{r^2}{4-r^2}+o(1)}.$$

Since $r \in (0, 2)$ is a fixed constant, the exponent $\frac{r^2}{4-r^2}$ is positive, then

$$\Pr\big(\tilde{h}_A(u_p^+) = \tilde{h}_A(u_p^-)\big) \le n^{-\frac{r^2}{4-r^2}+o(1)} \longrightarrow 0 \quad (n \to \infty).$$

Thus, no lower bound $p_1 > 0$ can hold for all such "nearby" pairs, and the scaled-rotation scheme fails to be (r, cr)-sensitive on S_p^{n-1} when $p < 2$. $\square$

Theorem 3.3. *Let $p \in (2, \infty]$. Then for any $r \in (0, 2)$ and any $c > 1$, there exists $N \in \mathbb{N}$ such that for all $n \ge N$ one can find $x_1, \dots, x_m \in \mathbb{S}_p^{n-1}$ with $\|x_i - x_j\|_p \ge cr$ but*

$$\mathbb{P}_{A\sim \mathrm{Uniform}(O(n,\mathbb{R}))}\big(\tilde{h}_A(x_i) = \tilde{h}_A(x_j)\big) \;\ge\; 1 - n^{-\frac{r^2}{4-r^2}+o(1)},$$

implying that

$$\mathbb{P}_{A\sim \mathrm{Uniform}(O(n,\mathbb{R}))}\big(\tilde{h}_A(x_i) = \tilde{h}_A(x_j)\big) \to 1 \text{ as } n \to \infty.$$

Hence, no fixed $p_2 < 1$ can serve as an upper bound on the collision probability for all such "far apart" pairs.

Proof. The theorem follows with the same construction as Theorem 3.2, with the difference being to use the lower bound in the cross-polytope LSH instead. $\square$

As a corollary of Theorems 3.2 and 3.3, we can informally state that there are no good scaled hash families for $p \neq 2$.

Corollary 3.4. *Let $p \ge 1$ with $p \neq 2$. Then the "scaled-rotation" hash family*

$$\tilde{\mathcal{H}} \;=\; \big\{\tilde{h}_A = \mu \circ \pi_p \circ r_A : A \in O(n, \mathbb{R})\big\}$$

on the ℓ_p–sphere $\mathbb{S}_p^{n-1}$ cannot be (r, cr, p_1, p_2)–sensitive for any fixed constants $p_1, p_2 \in (0, 1)$:

- *If $p < 2$, then for every $r \in (0, 2)$ there is no choice of $p_1 > 0$ such that*

$$\mathbb{P}_A\big(\tilde{h}_A(x) = \tilde{h}_A(y)\big) \;\ge\; p_1 \quad \textit{whenever } \|x - y\|_p \le r.$$

- *If $p > 2$, then for every $r \in (0, 2)$ and every $c > 1$ there is no choice of $p_2 < 1$ such that*

$$\mathbb{P}_A\big(\tilde{h}_A(x) = \tilde{h}_A(y)\big) \;\le\; p_2 \quad \textit{whenever } \|x - y\|_p \ge c\,r.$$

In other words, no nontrivial LSH guarantees can hold for the scaled-rotation scheme on $\mathbb{S}_p^{n-1}$ unless $p = 2$.

3.3 Hyperoctahedral Hash Family

In this section, we formally define our signed-permutation hashing scheme for points on the ℓ_p-sphere.

We wish to construct a hash family $\bar{\mathcal{H}}$ with the locality-sensitive property: points that are closer in ℓ_p should have a higher probability of colliding under the hash.

3.3.1 Group Actions

Definition 3.5. *(Group Action) Given a group $(G, *)$ and a set S, a* **group action** *of G on S is a function $\cdot : G \times S \to S$, denoted by $g \cdot s$ for $g \in G$ and $s \in S$, that satisfies the following two properties:*

1. *for the identity element $e \in G$, $e \cdot s = s$ for all $s \in S$.*
2. *for any $g, h \in G$ and $s \in S$, $(g * h) \cdot s = g \cdot (h \cdot s)$.*

Definition 3.6. *(Transitive Group Action) Let G be a group acting on a set S. The action of G on S is said to be* **transitive** *if for every pair of elements $s_1, s_2 \in S$, there exists at least one element $g \in G$ such that $g \cdot s_1 = s_2$.*

3.3.2 Hyperoctahedral Group Review

The hyperoctahedral group B_n consists of all *signed permutations*. In this way, B_n can be viewed as a subgroup of the orthogonal group $O(n, \mathbb{R})$. Matrices naturally act on $\mathbb{R}^n$ by left multiplication, so there is a well-defined group action of B_n of $\mathbb{R}^n$. We are interested in B_n because it is the group of symmetries of the ℓ_p-sphere for $p \neq 2$.

Each element A of B_n can be decomposed as $A = SP$ where S and P are signature and permutation matrices, respectively. Because there are 2^n signature matrices, $n!$ permutation matrices, and $(S_1, P_1) \neq (S_2, P_2)$ implies $S_1 P_1 \neq S_2 P_2$, there are $|B_n| = 2^n \cdot n!$ possible such transformations. Because B_n is finite (unlike $O(d, \mathbb{R})$), then B_n is not transitive. In fact, this lack of a continuous transformation, which is exhibited for the symmetries of the Euclidean sphere, will be why a hash family scheme using the symmetries of the ℓ_p-sphere for $p \neq 2$ will fail.

3.3.3 Theoretical Analysis

In this section, we show that the signed-permutation hash family fails to exhibit the locality-sensitive property under the ℓ_1-norm. Specifically, we prove that we can construct arbitrarily bad examples.

In Section 3, we constructed the hash function using the hyperoctahedral group. Specifically, for any $A \in B_n$, we defined

$$\bar{h}_A := \mu \circ r_A.$$

Lemma 3.7. *Let $A_1, A_2 \in B_n$ be any two signed-permutation matrices, and set*

$$R := A_2\, A_1^{-1} \;\in\; B_n.$$

Then,

$$\bar{h}_{A_2} = R \circ \bar{h}_{A_1}.$$

In particular, up to a relabeling of the $2n$ output buckets by the signed-permutation R, the two hash functions coincide.

Proof. Recall that for any $A \in B_n$ and any $y \in \mathbb{R}^n$, one has

$$\mu(A\,y) \;=\; A\,\mu(y),$$

because A merely permutes and possibly flips the signs of the coordinates, and the quantizer μ picks out the coordinate of largest absolute value together with its sign. Concretely, if $\mu(y) = \operatorname{sgn}(y_j) e_j$ with $j = \operatorname{argmax}_k |y_k|$, then

$$(Ay)_i = \; s_i\, y_{\pi(i)}, \quad \text{so} \quad \underset{i}{\operatorname{argmax}}\, |(Ay)_i| = \; \pi^{-1}(j),$$

and hence

$$\mu(Ay) = \operatorname{sgn}\big((Ay)_{\pi^{-1}(j)}\big)\, e_{\pi^{-1}(j)} = s_{\pi^{-1}(j)} \operatorname{sgn}(y_j)\, e_{\pi^{-1}(j)} \;=\; A\big(\operatorname{sgn}(y_j) e_j\big) \;=\; A\,\mu(y).$$

It follows that for any $x \in S_p^{n-1}$,

$$\bar{h}_{A_2}(x) = \mu\big(A_2\, x\big) = \mu\big(R\, A_1\, x\big) = R\,\mu\big(A_1\, x\big) = R\big(\bar{h}_{A_1}(x)\big),$$

as desired. ∎

Applying Lemma 3.7 to any $A, I \in B_n$ where I is the identity matrix, then

$$\bar{h}_A = \bar{h}_I = \mu \circ r_I \equiv \mu.$$

In other words, the symmetry fails to apply any meaningful transformation.

Theorem 3.8. *Let*

$$\bar{\mathcal{H}} \; = \; \{\bar{h}_A : A \in B_n\}$$

be the hash family generated by the hyperoctahedral group on S_p^{n-1} (any $p \neq 2$). Then $\bar{\mathcal{H}}$ induces exactly one partition of the sphere: namely, the one given by $\bar{h}_I(x) = \mu(x)$. In particular, for any two points x, y and any choice of A,

$$\bar{h}_A(x) = \bar{h}_A(y) \quad \Longleftrightarrow \quad \mu(x) = \mu(y),$$

so the collision probability $\Pr_{A \sim \mathrm{Uniform}(B_n)}[\bar{h}_A(x) = \bar{h}_A(y)]$ is either 0 or 1, independent of the distance $\|x - y\|_p$. Consequently, $\mathcal{H}$ cannot be locality-sensitive under the ℓ_p–metric for any $p \neq 2$.

Proof. By Lemma 3.7, all $\bar{h}_A$ induce the same partition, namely that of $\bar{h}_I = \mu$. Hence for every fixed pair (x, y),

$$\Pr_{A \sim B_n}\left[\bar{h}_A(x) = \bar{h}_A(y)\right] = \begin{cases} 1, & \mu(x) = \mu(y), \\ 0, & \mu(x) \neq \mu(y). \end{cases}$$

But on S_p^{n-1} one can always find arbitrarily close points x, y with $\mu(x) \neq \mu(y)$ (or arbitrarily far points with $\mu(x) = \mu(y)$), so no nontrivial (r, cr, p_1, p_2)-sensitivity can hold. □

4 Conclusion and Future Work

In this paper, we have shown that two of the most *direct* attempts to generalize the Euclidean cross-polytope LSH to arbitrary ℓ_p-spheres ($p \neq 2$)—namely, (i) the *scaled-rotation* scheme and (ii) the *hyperoctahedral sampling* scheme—both fail to achieve any nontrivial (r, cr)-sensitivity. For $p < 2$, the scaled-rotation family can drive the collision probability of arbitrarily close points to zero, and for $p > 2$ it forces distant points to collide with overwhelming probability; while the hyperoctahedral family induces exactly one partition (up to relabeling), so collision events become independent of distance. These negative results demonstrate that any successful LSH for $p \neq 2$ must depart from these symmetry-based constructions.

Looking ahead, several promising avenues remain open:

- **Partial or adaptive quantization.** Rather than snapping to a single dominant coordinate, might multi-coordinate or hierarchical quantizers yield genuine sensitivity on ℓ_p spheres?
- **Alternative group actions.** Beyond orthogonal or signed-permutation groups, are there other natural transformations of $\mathbb{R}^n$ whose induced partitions respect ℓ_p geometry in a locality-sensitive way?

We hope that this work—by clarifying which paths are *not* viable—will help focus future efforts on more nuanced, data-adaptive, or hybrid approaches to similarity search beyond the Euclidean realm. Understanding and harnessing the geometry of non-Euclidean norms remains an exciting challenge for both theory and practice.

References

[1] Alexandr Andoni, Piotr Indyk, Huy L. Nguyen, and Ilya Razenshteyn. Beyond locality-sensitive hashing. *Symposium on Discrete Algorithms (SODA)*, pages 1018–1028, 2014.

[2] Alexandr Andoni, Piotr Indyk, Thijs Laarhoven, Ilya Razenshteyn, and Ludwig Schmidt. Practical and optimal locality-sensitive hashing for angular distance. *Proceedings of the 29th International Conference on Neural Information Processing Systems*, pages 1225-1233, 2015.

[3] Andrei Z. Broder. On the resemblance and containment of documents. *Proceedings of the Compression and Complexity of Sequences (SEQUENCES)*, pages 21–29, 1997.

[4] Moses S. Charikar. Similarity estimation techniques from rounding algorithms. *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC)*, pages 380–388, 2002.

[5] Mayank Datar, Nicole Immorlica, Piotr Indyk, and Vahab Mirrokni. Locality-sensitive hashing scheme based on p-stable distributions. *Proceedings of the 20th Annual Symposium on Computational Geometry (SoCG)*, pages 253–262, 2004.

[6] Piotr Indyk and Rajeev Motwani. Approximate nearest neighbors: Towards removing the curse of dimensionality. *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC)*, pages 604–613, 1998.

A Appendix

Proposition A.1. *Let $x \in \mathbb{R}^n \setminus \{0\}$ and $p > 1$. Consider the problem*

$$\min_{y \in \mathbb{R}^n} \|y - x\|_p \quad \textit{s.t.} \quad \|y\|_p = 1.$$

Then the unique solution is obtained by

$$y^* \ = \ \frac{x}{\|x\|_p}.$$

Proof. Let y be any feasible point, so $\|y\|_p = 1$. By the triangle inequality,

$$\|x\|_p = \|(x-y) + y\|_p \leq \|x - y\|_p + \|y\|_p = \|x - y\|_p + 1,$$

which rearranges to

$$\|x - y\|_p \ \geq \ \|x\|_p - 1.$$

On the other hand, swapping x and y gives $\|y - x\|_p \geq 1 - \|x\|_p$, so altogether

$$\|y - x\|_p \ \geq \ \big|\|x\|_p - 1\big|.$$

This implies that no feasible y can achieve a smaller objective value than $\big|\|x\|_p - 1\big|$.

Next, let

$$x = \|x\|_p \, \hat{x}, \quad \text{where} \quad \hat{x} \ = \ \frac{x}{\|x\|_p},$$

so that $\|\hat{x}\|_p = 1$. Then

$$\big\|\hat{x} - x\big\|_p = \big\|\hat{x} - \|x\|_p \, \hat{x}\big\|_p = \big| \ 1 - \|x\|_p \big| \ \|\hat{x}\|_p = \big|\|x\|_p - 1\big|.$$

Hence $y = \hat{x}$ attains the lower bound, implying its optimality.

Finally, if $p > 1$, the function $y \mapsto \|y - x\|_p$ is strictly convex, so the minimizer is unique. $\square$