

# A $\sqrt{17/2}$ Lower Bound for Depth Reduction in Explainable Planar $k$ -Center Clustering

Yuran Xiu

Institute for Interdisciplinary Information Sciences, Tsinghua University  
xiuyr23@mails.tsinghua.edu.cn

15 October 2026

## Abstract

We prove that the factor-2 lower bound for depth reduction in explainable planar  $k$ -center clustering is not tight. For every integer  $k \geq 2$ , an explicit set of  $4k$  distinct integer points has both ordinary and explainable  $k$ -center radius one, while every axis-parallel threshold tree with at most  $k$  nonempty leaves and depth less than  $k - 1$  has radius at least  $\sqrt{17/2}$ . Thus, for every fixed depth multiplier  $H$  and every  $C < \sqrt{17/2}$ , arbitrarily large instances admit no such tree of depth at most  $H \lceil \log_2(k + 1) \rceil$  and radius at most  $C$  times the optimal explainable radius. The point set is a rescaled tail of the ladder of Deng, Gavva, Karthik, Patel, and Srinivasan. The strengthened analysis uses a four-color packing bound and the exact cluster budget to prove rigidity of every sufficiently small-diameter partition, without an explainability assumption. The complete infinite-family theorem and its unbounded-size corollary have been checked in Lean. The existence of a larger universal constant for logarithmic-depth explanations remains unresolved.

## 1 Introduction

An explainable clustering assigns each data point to a leaf of a decision tree whose tests compare a single coordinate with a threshold. The leaf's root-to-leaf path supplies an explanation of the assignment. This model was studied for  $k$ -means and  $k$ -medians by Dasgupta et al. [1], and for  $k$ -center by Laber and Murtinho [2]. Requiring short explanations adds a depth constraint to the usual cluster budget. Laber, Murtinho, and Oliveira [3] studied depth-sensitive objectives and algorithms; the maximum-depth requirement considered here is a worst-case condition on every explanation.

Deng et al. [4] introduced the price of depth reduction: the cost of the best explanation of a prescribed depth, relative to the best explanation without a depth constraint. They proved unbounded losses in the plane for  $k$ -means and  $k$ -median, and a lower bound approaching 2 for  $k$ -center. The latter result is their Theorem 4.1. In the WADS 2023 open problem collection, Deng explicitly asked whether the factor-2 lower bound for explainable  $k$ -center depth reduction is tight [5, Open Problem 6]. Open Problem 1 of [4, Section 5.3] asks for a shallow explanation with factor 2, or at least with some constant independent of  $k$ .

We give a negative answer to the factor-2 question. The lower bound is

$$\gamma := \frac{\sqrt{34}}{2} = \sqrt{\frac{17}{2}} > 2.$$

Our construction comes directly from the published ladder: remove its first two exceptional clusters, rescale, and use the number of remaining clusters as the budget. The contribution is the stronger rigidity analysis of this tail and the resulting depth-reduction obstruction. We make no claim that the underlying ladder geometry is new. Nor does a lower bound of  $\gamma$  settle the broader question of whether a larger absolute constant is possible.

The proof has two parts. First, four well-separated color classes imply that every cluster of diameter below  $\sqrt{34}$  contains at most four points. With  $4k$  points and at most  $k$  clusters, every cluster must therefore contain exactly four points. Two boundary points then force the first diamond to remain intact. A coordinate swap and translation permit induction. Second, the projection intervals of the intact diamonds force every nontrivial threshold cut to remove an endpoint diamond. This requires depth  $k - 1$ . Both arguments apply to every  $k$ , rather than to a finite collection of computed instances.

## 2 Model and results

All distances are Euclidean distances in  $\mathbb{R}^2$ . An ordinary clustering of a nonempty finite set  $X$  into at most  $k$  clusters is a partition  $\mathcal{P}$  of  $X$  into at most  $k$  nonempty parts, with a center  $z_A \in \mathbb{R}^2$  assigned to each  $A \in \mathcal{P}$ . Its radius is

$$\text{cost}(\mathcal{P}, z; X) := \max_{A \in \mathcal{P}} \max_{p \in A} \|p - z_A\|_2.$$

Centers may lie anywhere in the plane, and different parts may use the same center. We write  $\text{OPT}(X, k)$  for the infimum of these radii.

An axis-parallel threshold tree is a finite rooted binary tree. Each internal node tests either  $x < \theta$  or  $y < \theta$ , for an arbitrary real  $\theta$ ; points satisfying the inequality go left, and all other points go right. Let  $\ell_T(p)$  be the leaf reached by  $p$ . The nonempty leaf sets form a partition of  $X$ . Given arbitrary centers  $z_\ell \in \mathbb{R}^2$ , one for each leaf, define

$$\text{cost}(T, z; X) := \max_{p \in X} \|p - z_{\ell_T(p)}\|_2.$$

Only leaves reached by at least one input point count toward the budget. The depth  $\text{depth}(T)$  is the largest number of edges on a root-to-leaf path, with depth zero for a single leaf. We allow empty branches and redundant tests; Section 5 explains why they cannot improve the lower bound. If unary internal nodes are allowed syntactically, they can be contracted without affecting the assignments and without increasing depth.

Let  $\text{OPT}_{\text{exp}}(X, k)$  be the infimum of  $\text{cost}(T, z; X)$  over trees with at most  $k$  nonempty leaves, with no depth restriction. This is the denominator for the price of depth reduction. In particular it is not, by definition, the ordinary optimum  $\text{OPT}(X, k)$ . Our instances have both optima equal to one, and both infima are attained.

For  $j \geq 0$ , put  $e_j = j \bmod 2$  and define

$$c_j = (2j + 4e_j, 2j - 4e_j), \quad Q_j = c_j + \{(1, 0), (-1, 0), (0, 1), (0, -1)\}, \quad X_k = \bigcup_{j=0}^{k-1} Q_j. \quad (1)$$

**Theorem 1** (Planar depth-reduction obstruction). *For every integer  $k \geq 2$ , the set  $X_k$  in (1) has  $4k$  distinct integer points and satisfies:*

- (i)  $\text{OPT}(X_k, k) = \text{OPT}_{\text{exp}}(X_k, k) = 1$ ;
- (ii) a threshold tree with  $k$  nonempty leaves and depth  $k - 1$  attains radius one;

(iii) every threshold tree  $T$  with at most  $k$  nonempty leaves and  $\text{depth}(T) < k - 1$ , with any assigned leaf centers  $z$ , satisfies

$$\text{cost}(T, z; X_k) \geq \gamma = \sqrt{17/2}.$$

The optimum equalities and the radius-one witness also hold for  $k = 1$ .

To state the logarithmic consequence with explicit quantifiers, let

$$L(n) := \min\{h \in \mathbb{N} : n \leq 2^h\} \quad (n \geq 1).$$

Thus  $L(n) = \lceil \log_2 n \rceil$ .

**Corollary 2** (Arbitrarily large logarithmic-depth obstructions). *For every  $H \in \mathbb{R}$ , every  $C < \gamma$ , and every  $K \in \mathbb{N}$ , there is an integer  $k \geq \max\{2, K\}$  such that  $|X_k| = 4k$ ,  $\text{OPT}(X_k, k) = \text{OPT}_{\text{exp}}(X_k, k) = 1$ , and no threshold tree with at most  $k$  nonempty leaves satisfies both*

$$\text{depth}(T) \leq HL(k + 1) \quad \text{and} \quad \text{cost}(T, z; X_k) \leq C \text{OPT}_{\text{exp}}(X_k, k).$$

In particular, no universal factor-2 guarantee is possible with any fixed logarithmic-depth multiplier.

The conclusion concerns the exact leaf budget. It gives no lower bound for a model that is allowed additional nonempty leaves. It also gives no impossibility result for a factor  $C \geq \gamma$ .

### 3 The ladder and its separated color classes

Write

$$R = (1, 0), \quad L = (-1, 0), \quad U = (0, 1), \quad D = (0, -1),$$

and use  $R_j, L_j, U_j, D_j$  for the corresponding points of  $Q_j$ . The first three centers are  $(0, 0)$ ,  $(6, -2)$ , and  $(4, 4)$ . Every  $Q_j$  has enclosing radius exactly one: the center  $c_j$  gives radius one, and the antipodal pair  $R_j, L_j$  has distance two.

The construction's relation to [4, Section 4.1] is exact. Take the source construction with  $k + 2$  clusters, retain  $C_3, \dots, C_{k+2}$ , and apply

$$(x, y) \mapsto (2x, 2y - 5). \tag{2}$$

The first two retained centers become  $(0, 0)$  and  $(6, -2)$ ; the source's two-step translation becomes  $(4, 4)$ . Its half-unit diamonds therefore become exactly  $Q_0, \dots, Q_{k-1}$ . The budget in this paper is  $k$ , not  $k + 2$ . We analyze the retained tail, and do not assert the strengthened bound for the original unmodified source instance.

**Lemma 3** (Distinctness). *The sets  $Q_j$  are pairwise disjoint and each has four points. Consequently  $|X_k| = 4k$ .*

*Proof.* For  $i < j$ , let  $m = j - i$ . If  $m$  is even, then  $c_j - c_i = (2m, 2m)$ , and  $m \geq 2$ , so each center-coordinate difference is at least four. If  $m$  is odd, one center-coordinate difference is  $2m + 4 \geq 6$ . The offsets of two diamond points can change any coordinate difference by at most two. Thus a point of  $Q_i$  cannot equal a point of  $Q_j$ . The four offsets within a diamond are distinct.  $\square$

Number the types  $R, L, U, D$  by  $s = 0, 1, 2, 3$  and color the point of type  $s$  in  $Q_j$  by

$$\chi(j, s) = (j + s) \bmod 4. \tag{3}$$

**Lemma 4** (Color separation). *Distinct points with the same color have squared distance at least 34.*

*Proof.* Points in one diamond have different colors. Consider indices  $i < j$  with  $m = j - i$ . Equal colors require the type at index  $j$  to be  $s - m$  modulo four when the type at index  $i$  is  $s$ .

For  $m = 1$ , the center displacement is  $(6, -2)$  if  $i$  is even, and  $(-2, 6)$  if  $i$  is odd. Direct substitution of the four offsets gives the following squared distances:

| Type in $Q_i$     | $R$ | $L$ | $U$ | $D$ |
|-------------------|-----|-----|-----|-----|
| $i$ even, $m = 1$ | 34  | 68  | 34  | 36  |
| $i$ odd, $m = 1$  | 34  | 36  | 34  | 68  |
| $m = 2$           | 34  | 34  | 34  | 34  |

The third row follows from center displacement  $(4, 4)$  and type difference two. For example, its first entry is  $\|(4, 4) + U - R\|_2^2 = 3^2 + 5^2 = 34$ .

For even  $m \geq 4$ , both point-coordinate differences are at least  $2m - 2 \geq 6$ , so the squared distance is at least 72. For odd  $m \geq 3$ , one center-coordinate difference is  $2m + 4$ , and the corresponding point-coordinate difference is at least  $2m + 2 \geq 8$ . Its square alone exceeds 34. This covers every pair of indices.  $\square$

The suffix of the ladder reproduces the whole family. The isometry

$$\Phi(x, y) = (y + 2, x - 6) \quad (4)$$

satisfies  $\Phi(c_{j+1}) = c_j$  and permutes the four offsets by swapping coordinates. Hence

$$\Phi(Q_{j+1}) = Q_j, \quad \Phi(X_k \setminus Q_0) = X_{k-1}. \quad (5)$$

This isometry preserves Euclidean distances and takes axis-parallel threshold tests to axis-parallel threshold tests.

## 4 Rigidity of all small-diameter partitions

**Theorem 5** (Partition rigidity). *For every  $k \geq 1$ , every partition of  $X_k$  into at most  $k$  nonempty parts of diameter strictly less than  $\sqrt{34}$  is  $\{Q_0, \dots, Q_{k-1}\}$ , up to the order of its parts.*

*Proof.* By Lemma 4, a part contains at most one point of each color, and hence at most four points. There are  $4k$  points and at most  $k$  parts. It follows that there are exactly  $k$  parts and every part has exactly four points.

We first force  $Q_0$  to be a part. The squared distances from  $L_0 = (-1, 0)$  and  $D_0 = (0, -1)$  to the next two diamonds are:

|       | $R_1$ | $L_1$ | $U_1$ | $D_1$ | $R_2$ | $L_2$ | $U_2$ | $D_2$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $L_0$ | 68    | 40    | 50    | 58    | 52    | 32    | 50    | 34    |
| $D_0$ | 50    | 26    | 36    | 40    | 50    | 34    | 52    | 32    |

Columns whose index is at least  $k$  are simply absent. These are the only outside candidates at squared distance below 34. Indeed, if  $j \geq 3$  is odd, every point in  $Q_j$  has  $x$ -coordinate at least  $2j + 3 \geq 9$ , so its  $x$ -coordinate distance from either  $L_0$  or  $D_0$  is at least nine. If  $j \geq 3$  is even, then  $j \geq 4$  and every point in  $Q_j$  has both coordinates at least  $2j - 1 \geq 7$ . Its  $x$ -coordinate distance

from  $L_0$  is at least eight, and its  $y$ -coordinate distance from  $D_0$  is at least eight. Thus the complete lists of outside points at squared distance strictly below 34 are

$$\begin{aligned} L_0 : & \quad L_2, \\ D_0 : & \quad L_1, D_2, \end{aligned} \tag{6}$$

again omitting nonexistent indices.

Let  $A$  be the four-point part containing  $L_0$ . If  $A$  has no point outside  $Q_0$ , then  $A = Q_0$ . Otherwise (6) forces its only possible outside point to be  $L_2$ . Among the three other points of  $Q_0$ , the squared distances from  $L_2$  to  $R_0, U_0, D_0$  are respectively 20, 18, 34. Since  $|A| = 4$  and its diameter is strictly below  $\sqrt{34}$ , necessarily

$$A = \{L_0, R_0, U_0, L_2\}.$$

The point  $D_0$  lies in another part. Its three possible partners in  $Q_0$  are already used by  $A$ , and its only possible outside partners are  $L_1, D_2$ . Its part therefore has at most three points, contradicting the four-point conclusion. We have proved that  $Q_0$  is a part. This also handles  $k = 1, 2$ , when some or all of the outside candidates are absent.

For  $k = 1$  the conclusion follows immediately. For  $k \geq 2$ , remove the part  $Q_0$  and apply  $\Phi$  from (4). The remaining  $k - 1$  parts partition  $X_{k-1}$ , preserve their diameters, and have the required budget. Induction implies that they are its original diamonds. Pulling them back under  $\Phi$  completes the proof.  $\square$

**Corollary 6** (Radius rigidity). *Every ordinary clustering of  $X_k$  into at most  $k$  nonempty parts with radius less than  $\gamma$  has parts  $Q_0, \dots, Q_{k-1}$ .*

*Proof.* If a part has center  $z$  and radius at most  $r < \gamma$ , then for any two points  $p, q$  in it,  $\|p - q\|_2 \leq \|p - z\|_2 + \|q - z\|_2 \leq 2r < \sqrt{34}$ . Apply Theorem 5.  $\square$

## 5 The depth forced by intact diamonds

Call a threshold cut *clean* if it does not split any  $Q_j$ . It is *nontrivial* on  $X_k$  if both sides receive input points. The closed projection intervals of  $Q_j$  are

$$I_j^x = [2j + 4e_j - 1, 2j + 4e_j + 1], \quad I_j^y = [2j - 4e_j - 1, 2j - 4e_j + 1]. \tag{7}$$

Both endpoints of each interval are attained by points of that diamond.

**Lemma 7** (Clean cuts remove an endpoint). *For  $k \geq 2$ , every nontrivial clean axis-parallel threshold cut of  $X_k$  separates either  $Q_0$  from all the other diamonds or  $Q_{k-1}$  from all the other diamonds.*

*Proof.* Two projection intervals that touch or overlap cannot be put wholly on opposite sides of a threshold. More explicitly, if one cluster is wholly on the  $u < \theta$  side and another wholly on the  $u \geq \theta$  side, then their attained endpoints satisfy  $\max I < \theta \leq \min J$ , so  $\max I < \min J$ . Reversing the roles gives the same necessary strict gap in the reverse order. This argument applies to all real thresholds, including thresholds equal to input coordinates.

The projection centers divided by two are

$$a_j = j + 2e_j \quad \text{on the } x\text{-axis}, \quad b_j = j - 2e_j \quad \text{on the } y\text{-axis}.$$

Listing the even and odd indices separately gives the following sorted sets, where  $\{a, \dots, b\}$  is empty if  $a > b$ :

$$\begin{array}{c|c|c}
& \{a_j : 0 \leq j < k\} & \{b_j : 0 \leq j < k\} \\
\hline
k = 2m & \{0\} \cup \{2, \dots, 2m-1\} \cup \{2m+1\} & \{-1, \dots, 2m-2\} \\
k = 2m+1 & \{0\} \cup \{2, \dots, 2m+1\} & \{-1, \dots, 2m-2\} \cup \{2m\}.
\end{array} \tag{8}$$

The first row holds for  $m \geq 1$  and the second for  $m \geq 1$ . Consecutive integers in these sets represent actual centers separated by two; their unit-half-width projection intervals touch. Therefore no clean cut can separate diamonds within such a consecutive run.

The value 0 isolated on the left of the  $x$  list belongs to  $Q_0$ . For even  $k$ , the isolated rightmost  $x$  value  $2m+1$  belongs to  $Q_{k-1}$ ; for odd  $k$ , the isolated rightmost  $y$  value  $2m$  belongs to  $Q_{k-1}$ . These are the only gaps. A threshold is monotone in the tested coordinate, so any nontrivial clean cut must occur at one of these gaps and has the claimed form. For  $k=2$ , the  $x$  list is  $\{0, 3\}$  and the  $y$  intervals touch; the same conclusion holds.  $\square$

**Lemma 8** (Exact depth for the diamond partition). *For every  $k \geq 1$ , the partition  $\{Q_0, \dots, Q_{k-1}\}$  is realized by a threshold tree with  $k$  nonempty leaves and depth  $k-1$ . Every threshold tree whose nonempty leaf sets are these diamonds has depth at least  $k-1$ .*

*Proof.* For existence, the case  $k=1$  is one leaf. For  $k \geq 2$ , the cut  $x < 2$  separates  $Q_0$ , whose  $x$ -coordinates are at most one, from every other diamond. The latter coordinates are at least three (and at least five if only  $Q_1$  remains). By (5), their image under  $\Phi$  is  $X_{k-1}$ . Pull back an inductively constructed tree on that image. The left child is a leaf, the right subtree has depth  $k-2$ , and the resulting tree has  $k$  nonempty leaves and depth  $k-1$ .

For the lower bound, empty subtrees and trivial tests may first be pruned. At any node receiving input in only one child, replace the node by that child's subtree; discard its empty sibling. Repeating this operation terminates because the tree is finite. The nonempty leaf partition is unchanged, and depth does not increase. It suffices to lower-bound the depth of the resulting tree, in which every internal node reached by input splits its input nontrivially.

We induct on  $k$ . The claim is immediate for  $k=1$ . For  $k \geq 2$ , the root is a nontrivial clean cut: splitting a diamond at the root would prevent it from being a single leaf set. Lemma 7 shows that one child receives either  $Q_0$  or  $Q_{k-1}$  alone, and the other receives the remaining  $k-1$  diamonds. Removing  $Q_{k-1}$  leaves  $X_{k-1}$  exactly; removing  $Q_0$  leaves a set mapped to  $X_{k-1}$  by  $\Phi$ . In either case, the other child's tree realizes the diamond partition of a congruent copy of  $X_{k-1}$ , so its depth is at least  $k-2$ . The full tree has depth at least  $k-1$ .  $\square$

*Proof of Theorem 1.* Distinctness and cardinality are Lemma 3. The tree in Lemma 8, with center  $c_j$  at the leaf containing  $Q_j$ , attains radius one. This is also a feasible ordinary clustering.

Conversely, suppose an ordinary clustering with at most  $k$  nonempty parts has radius  $r < 1$ . Since  $1 < \gamma$ , Corollary 6 forces its parts to be the  $Q_j$ . Their antipodal pairs have distance two, so no assigned center can have radius below one on any of these parts. This contradiction proves the ordinary lower bound one, and hence the explainable lower bound one as well. The explicit witness proves attainment of both optima.

Finally, if a threshold tree with at most  $k$  nonempty leaves had radius less than  $\gamma$ , its leaf partition would be the original diamonds by Corollary 6. Lemma 8 would then force depth at least  $k-1$ . Taking the contrapositive yields the non-strict radius lower bound in part (iii). All arguments for the optimum and its witness also apply to  $k=1$ .  $\square$

## 6 Logarithmic depth and the scope of the obstruction

**Lemma 9.** *For every real  $H$  and natural number  $K$ , there exists  $k \geq \max\{2, K\}$  with  $HL(k+1) < k-1$ .*

*Proof.* Choose an integer  $N > \max\{H, K\}$ , put  $m = N + 4$ , and let  $k = 2^m - 1$ . For every integer  $m \geq 4$ , one has  $2^m \geq m^2$ . The base case is equality at four; the induction step follows from  $2m^2 \geq (m+1)^2$  for  $m \geq 4$ . In particular, our  $k$  is at least  $\max\{2, K\}$ , and  $L(k+1) = L(2^m) = m$ . Since  $H < m - 4$  and  $m \geq 4$ ,

$$HL(k+1) = Hm < m(m-4) < m^2 - 2 \leq 2^m - 2 = k - 1.$$

□

*Proof of Corollary 2.* Choose  $k$  from Lemma 9. Every tree obeying the stated depth bound then has depth strictly below  $k-1$ . Theorem 1 gives radius at least  $\gamma$ . Since  $\text{OPT}_{\text{exp}}(X_k, k) = 1$  and  $C < \gamma$ , it cannot also have radius at most  $C \text{OPT}_{\text{exp}}(X_k, k)$ . The parameter  $K$  was arbitrary, so the counterexamples have unbounded size. Finally  $2 < \gamma$  follows from  $4 < 17/2$ . □

The comparison uses the explainable optimum of the same instance and the same cluster budget, as in the price of depth reduction of [4]. The equality of ordinary and explainable optima is a property proved for this family, not a change of denominator. The source uses  $k$ -leaf trees; a lower bound for at most  $k$  nonempty leaves also applies to that requirement. On these instances the denominator is one under either convention because the witness already has exactly  $k$  nonempty leaves. No equivalence between exact and at-most leaf budgets under a fixed depth bound is needed.

The strict inequality in the rigidity hypothesis is essential to the argument: equal-color pairs can have squared distance exactly 34, and the boundary pair  $D_0, L_2$  has squared distance 34. We have proved a non-strict lower bound  $\gamma$ , not rigidity at radius  $\gamma$  or an impossibility result for that factor itself. We also do not restrict centers to input points; that restriction changes the optimum and is not the numerical model analyzed here. Repeated coordinate values are intentional and are covered by the threshold argument with attained interval endpoints. The input points themselves are distinct.

Thus the factor-2 sharpness question has a negative answer in the plane, even for unweighted integer points. Whether there exist universal constants  $C, H$  for which every planar instance has an explanation with at most  $k$  nonempty leaves, depth at most  $HL(k+1)$ , and radius at most  $C \text{OPT}_{\text{exp}}(X, k)$  remains unresolved by this result.

## 7 Lean verification and reproducibility

The associated formal development proves the complete infinite family, including arbitrary real thresholds, arbitrary Euclidean centers, nonempty-leaf counting, both optimum equalities, and arbitrarily large logarithmic-depth obstructions. It consists of seven modules under `Ladder/` and the entry point `Ladder.lean`. The following correspondence identifies the principal declarations; every name is in the `Ladder` namespace.

| Mathematical statement           | Lean declaration(s)                                                                                      |
|----------------------------------|----------------------------------------------------------------------------------------------------------|
| Distinct points; $ X_k  = 4k$    | <code>point_injective</code> , <code>card_points</code>                                                  |
| Color separation                 | <code>same_color_separated</code>                                                                        |
| Partition rigidity               | <code>rigidity</code>                                                                                    |
| Clean cuts and depth lower bound | <code>clean_choice</code> , <code>Tree.explains_depth</code>                                             |
| Explicit radius-one tree         | <code>Tree.optimal_depth</code> , <code>Tree.optimal_budget</code> ,<br><code>Tree.optimal_radius</code> |
| Arbitrary-partition lower bound  | <code>assignment_cost_ge_one</code> , <code>point_set_cost_ge_one</code>                                 |
| Both attained optima             | <code>ladder_both_optima</code>                                                                          |
| Theorem 1                        | <code>ladder_theorem</code> , <code>ladder_both_optima</code>                                            |
| Lemma 9                          | <code>logarithmic_lt_linear_unbounded</code>                                                             |
| Corollary 2                      | <code>ladder_obstruction_unbounded</code> ,<br><code>no_factor_below_unbounded</code>                    |
| Universal factor-2 refutation    | <code>no_factor_two</code>                                                                               |

The plane is represented by a pair of real coordinates, and its distance is defined explicitly as  $\sqrt{(x - x')^2 + (y - y')^2}$ ; the product max-metric is not used. The formal point enumeration swaps offset coordinates at odd indices. If  $s$  is the type number used in (3), its formal type number is  $s + 2e_j$  modulo four. The formal color  $3j + (s + 2e_j)$  modulo four therefore agrees with  $j + s$  modulo four. This enumeration change preserves both the underlying set and the coloring used here.

Tree evaluation returns a complete branch path, so distinct leaves remain distinct even if their centers coincide. The budget counts the image of the actual input under this evaluation. The checked depth argument covers trivial tests directly, rather than assuming a pruned tree. The ordinary lower bound quantifies over an arbitrary label type and counts only labels used by the data; it has no tree premise. The two optimum definitions use infima, and their values are established using attained least feasible radii. Finally, `ceilLog` is `Nat.clog 2`; `ceilLog_characterization` proves  $L(n) \leq h \iff n \leq 2^h$ . This is the exact integer ceiling of the binary logarithm on the positive inputs used by the theorem.

The pinned toolchain is `leanprover/lean4:v4.27.0`; the pinned mathlib revision is

`a3a10db0e9d66acbebf76c5e6a135066525ac900`.

The project’s `lake-manifest.json` fixes the remaining dependency revisions. With this toolchain and the supplied manifest, a portable build procedure, run from the Lean project directory, is:

```
lake exe cache get
lake build
lake env lean Ladder.lean
```

The final command prints the transitive axiom dependencies as well as the main declaration signatures. The artifact metadata records the project’s repository-relative location, exact source hashes, and the archived successful verification logs; the commands above do not depend on the original machine’s absolute paths.

The accepted verification record dated 14 October 2026 contains a fresh compilation of all eight authored modules and an independent fresh reproduction from a separate source snapshot. All nineteen transitive `#print axioms` queries report only `propext`, `Classical.choice`, and `Quot.sound`. The authored development uses no `sorry`, `admit`, newly postulated mathematical axiom, `native_decide`, or external computational certificate. The finite arithmetic cases are checked proof terms, and the numerical experiments are not premises of the theorems. The import audit records source and object

hashes for 3,274 imported modules and the pinned dependency repositories. Existing dependency caches were reused; a complete rebuild of all library dependencies is not claimed. Formal verification establishes the stated mathematics, not bibliographic priority or a solution to the larger-constant question.

## References

- [1] Sanjoy Dasgupta, Nave Frost, Michal Moshkovitz, and Cyrus Rashtchian. Explainable  $k$ -means and  $k$ -medians clustering. In *Proceedings of the 37th International Conference on Machine Learning*, Proceedings of Machine Learning Research 119, pages 7055–7065, 2020. <https://proceedings.mlr.press/v119/moshkovitz20a.html>.
- [2] Eduardo S. Laber and Lucas Murtinho. On the price of explainability for some clustering problems. In *Proceedings of the 38th International Conference on Machine Learning*, Proceedings of Machine Learning Research 139, pages 5915–5925, 2021. <https://proceedings.mlr.press/v139/lab21a.html>.
- [3] Eduardo Laber, Lucas Murtinho, and Felipe Oliveira. Shallow decision trees for explainable  $k$ -means clustering. arXiv:2112.14718v2, 26 August 2022. <https://arxiv.org/abs/2112.14718v2>.
- [4] Chengyuan Deng, Surya Teja Gavva, Karthik C. S., Parth Patel, and Adarsh Srinivasan. Impossibility of depth reduction in explainable clustering. arXiv:2305.02850v2, 2 March 2026. <https://arxiv.org/abs/2305.02850v2>. The cited construction, theorem, and open problem are Section 4.1, Theorem 4.1, and Open Problem 1 of this version.
- [5] WADS 2023. *Open Problems from WADS 2023*, 3 August 2023. Open Problem 6, “Depth-Reduction for Explainable  $k$ -Center,” posed by Chengyuan Deng, page 3. [https://wadscg2023.ens.concordia.ca/assets/pdf/wads\\_open\\_problems.pdf](https://wadscg2023.ens.concordia.ca/assets/pdf/wads_open_problems.pdf).