Distinct Distances in Planar Point Sets with Forbidden 4-Point Patterns

Adrian Dumitrescu*

April 20, 2020

Abstract

A point set in a metric space is said to have property $\Phi(4,5)$ if every 4 elements determine at least 5 distinct distances. According to an old conjecture of Erdős (1986 or earlier), a set of n points in the Euclidean plane satisfying this restriction determines $\Omega(n^2)$ distinct distances. This property (restriction) is shown to be equivalent to forbidding eight 4-element patterns, π_i, $i = 1, \ldots, 8$ (described in Section 2, Lemma 1). The existence of n-element point sets without the three patterns π_1, π_2, π_3, that determine only $o(n^2)$ distinct distances was previously known. Here we exhibit n-element point sets without the seven patterns $\pi_1, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7, \pi_8$, that determine only $o(n^2)$ distinct distances. The existence of point sets missing all eight forbidden patterns and determining only $o(n^2)$ distinct distances remains open.

Keywords: distinct distances, distinct vectors, general position, probabilistic construction.

1 Introduction

In 1946, in his classical paper [4] published in the *American Mathematical Monthly*, Erdős raised the following inspiring question: What is the minimum number of distinct distances determined by n points in the plane? Denoting this number by $g(n)$, he proved that $g(n) = \Omega(\sqrt{n})$, and pointed out that the upper bound $g(n) = O(n/\sqrt{\log n})$ follows from estimating the number of distinct distances in a $\sqrt{n} \times \sqrt{n}$ section of the integer grid. He also went further to conjecture that the upper bound is the best possible, i.e., $g(n) = \Omega(n/\sqrt{\log n})$. After many successive improvements, in a breakthrough development, Guth and Katz [13] managed to bring the lower bound very close to the conjectured upper bound: specifically, they proved that $g(n) = \Omega(n/\log n)$. See also the many different surveys and articles dedicated to this research area, e.g., [19, 20] for some recent accounts. The problem of *distinct distances* posed in 1946 lead to many interesting variants, one of which is discussed here. Specifically, given a pair k, l of positive integers, with $l \leq \binom{k}{2}$, what is the minimum number of distinct distances determined by a set of n points in the plane, in which every k-element subset determines at least l distinct distances?

A set of points in the plane is said to be in *general position* if no three of them are collinear and no four of them are cocircular. If a point set determines only distinct vectors, it is called *parallelogram free*. A *kite* is a (convex or concave) quadrilateral, whose four sides can be grouped into two pairs of equal-length sides that are adjacent to each other; see Fig. 1 (right). In contrast, a parallelogram also has two pairs of equal-length sides, but they are opposite to each other rather than adjacent. (A rhombus is both a parallelogram and a kite.)

*Department of Computer Science, University of Wisconsin–Milwaukee, USA. Email: dumitres@uwm.edu

It is known [3] that there exist n-element point sets in the plane in general position, and parallelogram free, that determine only $O(n^2/\sqrt{\log n})$ distinct distances. A point set is said to have property $\Phi(k,l)$ if every k-element subset determines at least l distinct distances.

Erdős [8, p. 34] conjectured that every set of n points in the plane satisfying property $\Phi(4,5)$ determines $\Omega(n^2)$ distinct distances; see also [2, Conjecture 6, p. 204]; while he repeatedly asked the question over the years, see [5, p. 101], [6, p. 61], [7, p. 149], [9, p. 347]. A classification (in the forthcoming Lemma 1 in Section 2) shows that a planar point set has property $\Phi(4,5)$ if and only if it does not contain any of eight forbidden 4-element patterns, π_i, $i = 1,\ldots,8$: (i) an equilateral triangle and an arbitrary 4th point, (ii) a parallelogram, (iii) an isosceles trapezoid, (iv) a star with 3 edges of the same length, (v) the vertex set of a path with 3 edges of the same length, (vi) a kite, (vii) an isosceles triangle plus an edge incident to a base endpoint and whose length equals the length of the base, and (viii) an isosceles triangle plus an edge incident to the apex and whose length equals the length of the base.

The earlier construction described in [3] is a subset of size $\Theta(n)$ of the $n \times n$ section of $\mathbb{Z}^2$ that determines $O(n^2/\sqrt{\log n})$ distinct distances. Since $\mathbb{Z}^2$ does not determine any equilateral triangles, the constructed point set avoids pattern π_1; further, as shown in [3], it also avoids pattern π_2. Finally, since the construction in [3] has no four points on a circle, it avoids pattern π_3 as well; indeed, recall that every isosceles trapezoid can be inscribed in a circle. To summarize, the constructed point set avoids the three patterns π_1, π_2, π_3. Our main result in this paper is the following.

Theorem 1. *There exist n-element point sets in the plane without the seven patterns $\pi_1, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7, \pi_8$, that determine only $O(n^2/\sqrt{\log n})$ distinct distances.*

Notation. For two points $a, b \in \mathbb{R}^2$, let $\ell(a,b)$ denote the line incident to a and b. Let $G_n = \{0,1,\ldots,n-1\} \times \{0,1,\ldots,n-1\}$ be the (standard) $n \times n$ section of the integer lattice $\mathbb{Z}^2$; and let $G_{\sqrt{n}}$ denote a $m \times m$ section of $\mathbb{Z}^2$, where $m = \lfloor \sqrt{n} \rfloor$. For a point-set S, its *distance graph* is the complete graph on vertex set S, where edges are colored so that two edges receive the same color if and only if they have the same Euclidean length.

2 Preliminaries

We will apply a classic result of Szemerédi and Trotter [23] on the number of point-lines incidences in the plane (the result holds in arbitrary dimensions); see also [17] for several applications of incidence bounds. The Szemerédi-Trotter bound comes in two equivalent formulations. Given a point set S in $\mathbb{R}^2$, for any integer $k \geq 2$, a line is called *k-rich* if it is incident to at least k points of S. We denote by $L_{\geq k}$ the set of k-rich lines.

Theorem 2. (Szemerédi-Trotter [23]). *Given n points in $\mathbb{R}^2$, the number of k-rich lines, $k \geq 2$, is*

$$|L_{\geq k}| = O\left(n^2/k^3 + n/k\right).$$

Theorem 3. (Szemerédi-Trotter [23]). *The number of point-line incidences among n points and ℓ lines in $\mathbb{R}^2$ is*

$$I(n,\ell) = O(n^{2/3}\ell^{2/3} + n + \ell).$$

Previous related work. It is known there exist suitable subsets of size $\Theta(n)$ of G_n that have no 3 collinear points, no 4 cocircular points, and no 4 points that make a parallelogram [3].

For a prime p, and $x \in \mathbb{Z}$, let $\widehat{x} := x \mod p$ (we view $\widehat{x}$ as an element of $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$). Let n be a prime and consider the n-element point set $E_n \subset G_n$:

$$E_n = \{(i, \widehat{i^2}) \mid i = 0, 1, \ldots, n-1\}.$$

Erdős showed that E_n has no three collinear points. Let

$$S_n = \{(i, \widehat{i^2}) \mid i = 0, 1, \ldots, (n-1)/4\}.$$

Recall that a $\sqrt{n} \times \sqrt{n}$ section of the integer lattice determines $O(n/\sqrt{\log n})$ distances and this leads to the upper bound $g(n) = O(n/\sqrt{\log n})$ [4]. As such, G_n determines $O(n^2/\sqrt{\log n})$ distinct distances; obviously, this upper bound also holds for any subset of G_n, E_n or S_n in particular. Clearly S_n has no three collinear points (as a subset of E_n).

Thiele [24] showed that S_n has no four cocircular points. This was in the context of finding large subsets of the $n \times n$ grid without four cocircular points in response to a problem raised by Erdős and Purdy; see [2, pp. 418]. Thiele's result was rediscovered in [3] (with a different proof), where it was also shown that S_n determines no parallelogram. Consequently, there exist n-element point sets in the plane with no 3 collinear points, no 4 cocircular points, and no 4 points that make a parallelogram, that determine only $O(n^2/\sqrt{\log n})$ distinct distances.

Distinct distances and the property $\Phi(k,l)$. Let S be a set of points in the plane. For integers $k, l \geq 2$ with $l \leq \binom{k}{2}$, let $\Phi(k,l)$ be the property that any k points from S determine at least l distinct distances; and let $\phi(n,k,l)$ denote the *minimum* number of distinct distances determined by a planar n-element point set S with the property $\Phi(k,l)$. Trivially, $\phi(n,k,\binom{k}{2}) = \binom{n}{2}$ for every $k \geq 4$, since no two distances are the same. Most questions about estimating $\phi(n,k,l)$ are notoriously hard and essentially unsolved; for instance, determining whether $\phi(n,3,3) = O(n)$ is currently an outstanding open problem; see also the survey [20] for a listing of current bounds. On the other hand, the famous Guth-Katz result [13] immediately implies that $\phi(n,k,l) = \Omega(n/\log n)$ for every k and l.

The current best bounds for a few relevant combinations (k,l), with $k \geq 4$, as implied by results of Erdős and Gyárfás [10], Fox, Pach, and Suk [12], and Pohoata and Sheffer [18], are listed in Table 1.

Variant	Lower bound	Upper bound
$\phi(n,4,5)$	$\Omega(n)$	$O(n^2)$
$\phi(n,5,9)$	$\Omega(n)$	$O(n^2)$
$\phi(n,9,33)$ [12]	$\Omega(n^{8/7-\varepsilon})$	$O(n^2)$
$\phi(n,7,19)$ [10]	$\Omega(n^{4/3})$	$O(n^2)$
$\phi(n,6,14)$ [18]	$\Omega(n^{3/2})$	$O(n^2)$
$\phi(n,7,20)$ [12]	$\Omega(n^2)$	$O(n^2)$
$\phi(n,8,26)$ [12]	$\Omega(n^2)$	$O(n^2)$

Table 1: The entries are listed in lower bound order, from linear to quadratic.

Forbidden 4-point patterns and the property $\Phi(4,5)$. Consider the property $\Phi(4,5)$; we first show that the following makes a complete list of forbidden patterns in $G_{\sqrt{n}}$; if *none* of these patterns is present, the point set has the property $\Phi(4,5)$.

Lemma 1. *Assume that $Q = \{a, b, c, d\}$ determines at most 4 distinct distances. Then Q determines one of the following 8 patterns:*

- π_1: *an equilateral triangle plus an arbitrary vertex.*
- π_2: *a parallelogram.*
- π_3: *an isosceles trapezoid (four points on a line, a, b, c, d, where $ab = cd$, form a degenerate isosceles trapezoid).*
- π_4: *a star with 3 edges of the same length.*
- π_5: *a path with 3 edges of the same length.*
- π_6: *a kite.*
- π_7: *an isosceles triangle plus an edge incident to a base endpoint, and whose length equals the length of the base.*
- π_8: *an isosceles triangle plus an edge incident to the apex, and whose length equals the length of the base.*

Proof. Assume that $Q = \{a, b, c, d\}$ determines at most 4 distinct distances; see Fig. 1 for an illustration. Then either some distance occurs (at least) 3 times, or there are two distinct distances x, y each occurring exactly 2 times. We distinguish two cases (occasionally we refer to the distance graph and its coloring in the analysis):

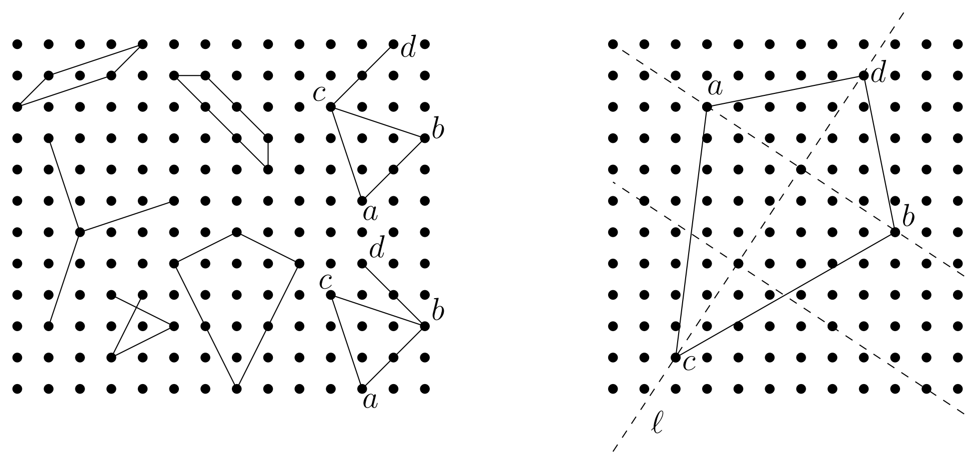

Figure 1: Left: seven forbidden patterns in $G_{\sqrt{n}}$. Right: a kite with axis $\ell(cd)$.

Case 1: x occurs (at least) 3 times. Then the edges of length x form either a monochromatic triangle, or a monochromatic star with 3 edges, or a monochromatic path with 3 edges; i.e., one of the patterns π_1, π_4, π_5 occurs.

Case 2: x and y occur exactly 2 times each.

Case 2.1: Neither the two edges of length x nor the two edges of length y share a common endpoint. Then the edges of length x and y form either a parallelogram or an isosceles trapezoid; note that in the latter case, the two non-parallel edges of the trapezoid have length x and the two diagonals of the trapezoid have length y. That is, one of the patterns π_2, π_3 occurs.

Case 2.2: Both the two edges of length x and the two edges of length y share common endpoints. Assume that Δcab is an isosceles triangle, with $|ca| = |cb| = x$. There are two possibilities: (i) the two edges of length y are adjacent at d, i.e., $|da| = |db| = y$, or (ii) they are adjacent at a (or b, this case is symmetric), i.e., $|ab| = |ad| = y$. In the former case ($|ca| = |cb| = x$ and $|da| = |db| = y$), we have a kite (pattern π_6); in the latter case we have an isosceles triangle plus an edge incident to a base endpoint, and whose length equals the length of the base (pattern π_7).

Case 2.3: The two edges of length x share a common endpoint while the two edges of length y do not share a common endpoint, i.e., $|ca| = |cb| = x$ and $|ab| = |cd| = y$; and we have an isosceles triangle plus an edge incident to the apex c, and whose length equals the length of the base (pattern π_8). $\square$

3 Forbidden patterns in the $\sqrt{n} \times \sqrt{n}$ grid section

Our proof of Theorem 1 is based on a probabilistic construction of a subset of a square section of the integer lattice. To analyze this construction we estimate the frequency of each of the forbidden patterns in such a section.

We begin with a simple lemma on the geometry of orthogonal line intersections inside a square.

Lemma 2. *Let U be a square and ℓ and h be two orthogonal lines that intersect inside U; refer to Fig. 2. Consider an arbitrary orientation of ℓ and let ℓ^- and ℓ^+ be the two closed halfplanes determined by ℓ, respectively. Then*

$$\min\left(|h \cap \ell^- \cap U|, |h \cap \ell^+ \cap U|\right) \leq |\ell \cap U|.$$

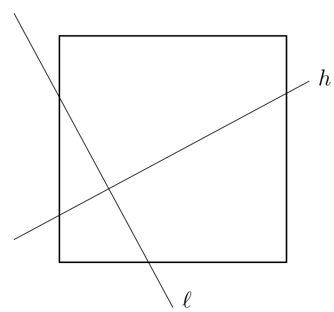

Figure 2: A square and two orthogonal lines.

Proof. We may assume that U is a unit square. We distinguish two cases.

Case 1. ℓ intersects two adjacent sides of U, say, $\ell^- \cap U$ is a right triangle Δ with orthogonal sides of lengths x and y, respectively. Then $\min\left(|h \cap \ell^- \cap U|, |h \cap \ell^+ \cap U|\right)$ is at most the length of the height of Δ, i.e.,

$$\min\left(|h \cap \ell^- \cap U|, |h \cap \ell^+ \cap U|\right) \leq \frac{|xy|}{|\ell \cap U|} \leq |x| \leq |\ell \cap U|.$$

Case 2. ℓ intersects two opposite sides of U; as such, $|\ell \cap U| \geq 1$. On the other hand, since $|h \cap U| \leq \texttt{diam}(U) = \sqrt{2}$, we have

$$\min\left(|h \cap \ell^- \cap U|, |h \cap \ell^+ \cap U|\right) \leq |h \cap U|/2 \leq \sqrt{2}/2 \leq 1 \leq |\ell \cap U|.$$

It is easy to see that apart from a small constant factor, the inequality in the lemma cannot be improved. $\square$

We continue with a sequence of lemmas concerning the number of forbidden patterns of each type that are present in the $\sqrt{n} \times \sqrt{n}$ section $G_{\sqrt{n}}$; specifically, let $F_i(n)$ denote the number of (forbidden) patterns π_i present in $G_{\sqrt{n}}$, for $i = 1, \ldots, 8$.

Lemma 3. *Let P denote the number of parallelograms in $G_{\sqrt{n}}$; i.e., $P = F_2(n)$. Then $P = \Theta(n^3)$.*

Proof. Given any three elements of $a, b, c \in G_{\sqrt{n}}$, they can be completed to a parallelogram in at most 3 ways; thus $P = O(n^3)$. It is easy to see that this bound is also attainable. □

For an isosceles triangle Δcab, where $|ca| = |cb|$, c is incident to the perpendicular bisector, say, ℓ, of segment ab; we say that Δcab is an isosceles triangle with axis ℓ.

Lemma 4. *Let ℓ be a line determined by $G_{\sqrt{n}}$, that is incident to j points of $G_{\sqrt{n}}$. Then the number of isosceles triangles with axis ℓ is $O(j^3)$.*

Proof. Let Δcab be an isosceles triangle with axis ℓ. The points in $G_{\sqrt{n}}$ that are incident to ℓ are evenly spaced, say at distance δ. Since $\ell(a, b)$ is orthogonal to ℓ, the points in $G_{\sqrt{n}}$ that are incident to $\ell(a, b)$ are also evenly spaced with the same distance δ. Let m_L and m_R be the numbers of points of $G_{\sqrt{n}} \cap \ell(a, b)$ that are left and right of ℓ, respectively. By Lemma 2 we have $\min(m_L, m_R) = O(j)$. It follows that the number of isosceles triangles Δcab, where $\ell(a, b)$ is a fixed line, and $c \in \ell$ is a fixed point, is $O(j)$. Moreover, the lines $\ell(a, b)$ are also evenly spaced with the same distance δ, thus the total number of lines $\ell(a, b)$ where Δcab is an isosceles triangle with axis ℓ is also $O(j)$. Altogether there are $O(j^2)$ choices for a, b (fixing one of them determines the other) and $O(j)$ choices for c; consequently, there are $O(j^3)$ isosceles triangles with axis ℓ. □

Lemma 5. *Let ℓ be a line determined by $G_{\sqrt{n}}$, that is incident to j points of $G_{\sqrt{n}}$. Then the number of kites with axis ℓ is $O(j^4)$.*

Proof. Let $\{a, b, c, d\}$ be a kite with axis $\ell = \ell(c, d)$, for some $c, d \in G_n$; see Fig. 1 (right). Observe that Δcab and Δdab are both isosceles triangles with a common axis ℓ and sharing the same pair of vertices a, b, and so we can apply the arguments in the proof of Lemma 4. We only include a summary of the findings. The number of kites $\{a, b, c, d\}$, where $\ell(a, b)$ is a fixed line, and $c, d \in \ell$ are fixed, is $O(j)$. Moreover, the total number of lines $\ell(a, b)$, where $\{a, b, c, d\}$ is a kite with axis ℓ is also $O(j)$. Altogether there are $O(j^2)$ choices for a, b and $O(j^2)$ choices for c and d; consequently, there are $O(j^4)$ kites with axis ℓ, as required. □

Lemma 6. *Let K denote the number of kites in $G_{\sqrt{n}}$; i.e., $K = F_6(n)$. Then $K = \Theta(n^{5/2})$.*

Proof. Let $S = G_{\sqrt{n}}$. Observe that every line determined by S is incident to at most $\sqrt{n}$ points of S. Denote by L_i the set of lines incident to at least 2^i but fewer than 2^{i+1} points of S, for $i = 1, \ldots, \log \sqrt{n}$.

By Theorem 2, we have $|L_i| = O\left(n^2/2^{3i}\right)$ in this range (since the first term dominates the second term in the upper bound). Summing the number of kites with axis ℓ over all lines ℓ determined by S yields

$$K = \sum_{i=1}^{\log \sqrt{n}} O\left(\frac{n^2}{2^{3i}} \cdot 2^{4i}\right) = \sum_{i=1}^{\log \sqrt{n}} O\left(n^2 2^i\right) = O(n^{5/2}),$$

as required.

It is easy to check that this upper bound is asymptotically tight, since there are $\Omega(n^{5/2})$ kites whose axes are axis-aligned: there are about $\sqrt{n}$ choices for $\ell(c, d)$, $\binom{\sqrt{n}}{2}$ choices for c, d on this line, $\sqrt{n}$ choices for $\ell(a, b)$, and $\sqrt{n}$ choices for the pair a, b on this line. □

For an isosceles trapezoid $\tau = \{a,b,c,d\}$, where $|ad| = |bc|$, let ℓ denote the common perpendicular bisector of ab and cd; we say that τ is an isosceles trapezoid with axis ℓ.

Lemma 7. *Let ℓ be a line determined by $G_{\sqrt{n}}$, that is incident to j points of $G_{\sqrt{n}}$. Then the number of isosceles trapezoids with axis ℓ is $O(j^4)$.*

Proof. Let $\{a,b,c,d\}$ be an isosceles trapezoid with axis ℓ, where $\ell(a,b) \parallel \ell(c,d)$; see Fig. 3. The arguments are similar to those in the proofs of Lemmas 4 and 5. In particular, if the points in $G_{\sqrt{n}} \cap \ell$ are at distance δ from each other, the points in $G_{\sqrt{n}} \cap \ell(a,b)$ are also at distance δ from each other. Let m_L and m_R be the numbers of points of $G_{\sqrt{n}} \cap \ell(a,b)$ that are left and right of ℓ, respectively. By Lemma 2 we have $\min(m_L, m_R) = O(j)$. Similarly, if p_L and p_R are the numbers of points of $G_{\sqrt{n}} \cap \ell(c,d)$ that are left and right of ℓ, respectively, we have $\min(p_L, p_R) = O(j)$.

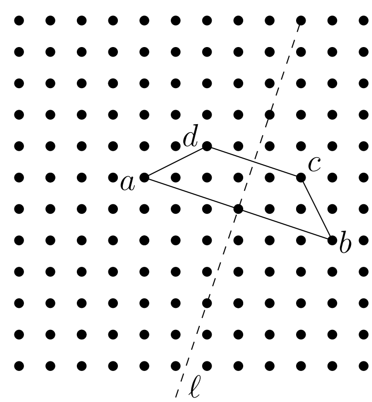

Figure 3: An isosceles trapezoid with axis ℓ.

It follows that the number of isosceles trapezoids $\{a,b,c,d\}$, where $\ell(a,b)$ and $\ell(c,d)$ are fixed lines is $O(j^2)$. There are $O(j^2)$ choices for $\ell(a,b)$ and $\ell(c,d)$, and thus $O(j^4)$ isosceles trapezoids with axis ℓ altogether. $\square$

Lemma 8. *Let T denote the number of isosceles trapezoids in $G_{\sqrt{n}}$; i.e., $T = F_3(n)$. Then $T = \Theta(n^{5/2})$.*

Proof. The upper bound calculation is the same with that in the proof of Lemma 6. Summing the number of isosceles trapezoids with with axis ℓ over all lines ℓ determined by S yields

$$T = \sum_{i=1}^{\log \sqrt{n}} O\left(\frac{n^2}{2^{3i}} \cdot 2^{4i}\right) = \sum_{i=1}^{\log \sqrt{n}} O\left(n^2 2^i\right) = O(n^{5/2}).$$

The lower bound argument is also similar to that in the proof of Lemma 6. There are $\Omega(n^{5/2})$ isosceles trapezoids whose axes are axis-aligned: there are about $\binom{\sqrt{n}}{2}$ choices for $\ell(a,b)$ and $\ell(c,d)$, $\sqrt{n}$ choices for the perpendicular bisector ℓ, $\sqrt{n}$ choices for a on $\ell(a,b)$ left of ℓ, and $\sqrt{n}$ choices for d on $\ell(c,d)$ left of ℓ. $\square$

Let I denote the number of isosceles triangles with vertices in $G_{\sqrt{n}}$. It is known [11] that $I = \Omega(n^2 \log n)$; see also [2, p. 276]. From the other direction, we have the following tight upper bound.

Lemma 9. $I = O(n^2 \log n)$.

Proof. Similarly to the proof of Lemma 6, we now have

$$I = \sum_{i=1}^{\log \sqrt{n}} O\left(\frac{n^2}{2^{3i}} \cdot 2^{3i}\right) = O\left(n^2 \sum_{i=1}^{\log \sqrt{n}} 1\right) = O\left(n^2 \log n\right). \qquad \square$$

Let $f(n)$ denote the maximum number of occurrences of the same distance, as determined by pairs of points in $G_{\sqrt{n}}$. It was pointed out by Erdős [4] that $f(n) \geq n^{1+c_1/\log\log n}$, where $c_1 > 0$ is a suitable constant—without providing complete proof details; see [15, pp. 52–53], or [16, pp. 143–144] for such arguments. On the other hand, the current best upper bound stands at $O(n^{4/3})$ [21], as established by Spencer et al. [21] in the early 1980s; see also [22] for a short proof of this bound; or [2, Ch. 5.1] for a general account. The above upper bound[1] gives the following.

Lemma 10. *Let $\delta > 0$ and v be any lattice point in $G_{\sqrt{n}}$. Then v is at distance δ from $O(n^{1/3})$ points in $G_{\sqrt{n}}$.*

Proof. Assume that v is at distance δ from x points in $G_{\sqrt{n}}$. Then one could find $\Omega(n)$ points in the (slightly larger) lattice section $G_{2\sqrt{n}}$ that are each at distance δ from at least x points; and thus obtain $\Omega(xn)$ repeated distances in this section. By the upper bound previously mentioned, this number is $O((4n)^{4/3}) = O(n^{4/3})$, hence $x = O(n^{1/3})$, as claimed. $\square$

Lemma 11. *Let $F_i(n)$ denote the number of patterns π_i present in $G_{\sqrt{n}}$, for $i = 1, \ldots, 8$. Then $F_i(n) = o(n^{5/2})$ for $i = 4, 5, 7, 8$.*

Proof. Let I denote the number of isosceles triangles with vertices in $G_{\sqrt{n}}$. By Lemma 9, $I = O(n^2 \log n)$. Observe that any 3-star or pattern π_8 can be obtained by appending an edge of a given length to the apex of an isosceles triangle. Similarly, observe that any 3-path or pattern π_7 can be obtained by appending an edge of a given length to one of the endpoints of the base of an isosceles triangle. Lemma 10 yields that

$$F_i(n) \leq 2I \cdot O(n^{1/3}) = O(n^{7/3} \log n) = o(n^{5/2}), \text{ for } i = 4, 5, 7, 8. \qquad \square$$

4 Proof of Theorem 1

Construction. A 4-element subset of G_n is called *bad* if it corresponds to one of the seven forbidden patterns, π_i, $i = 1, 3, 4, 5, 6, 7, 8$. Since $\mathbb{Z}^2$ does not determine any equilateral triangles, G_n avoids pattern π_1. Let $B_1, \ldots, B_k$, denote all the bad 4-element subsets in G_n. By Lemmas 6,8,11, we have $k \leq cn^5$, for some positive constant c (and for n sufficiently large). We choose a point set by the following construction.

In the first step, choose a subset R of size λn from among the n^2 points in G_n uniformly at random; here $\lambda < 1$ is a suitable positive constant (to be determined). Let p be the probability that a given 4-element subset Q is present in the random sample of λn points. This probability equals the ratio between the number of λn-sets containing Q and the total number of λn-sets. In particular, p is the probability that a given bad 4-element subset appears in the random sample R of λn points. Let X be the random variable counting the number of bad 4-element subsets in the

[1] A sharper estimate on the number of repeated distances, $O(n^{1+\varepsilon})$ for any $\varepsilon > 0$, is known for points in $G_{\sqrt{n}}$; see [14, Sections 16.9 and 18.1]. On the other hand, a sharper bound is not needed in our derivation.

random sample. We have

$$\begin{aligned}
p = \frac{\binom{n^2-4}{\lambda n-4}}{\binom{n^2}{\lambda n}} &= \frac{(n^2-4)!}{(\lambda n-4)!(n^2-\lambda n)!} \cdot \frac{(\lambda n)!(n^2-\lambda n)!}{(n^2)!} \\
&= \frac{\lambda n\,(\lambda n-1)(\lambda n-2)(\lambda n-3)}{n^2\,(n^2-1)(n^2-2)(n^2-3)} \\
&\leq \frac{\lambda^4}{n^4},
\end{aligned}$$

for n sufficiently large (indeed, we have $\frac{\lambda n-i}{n^2-i} \leq \frac{\lambda n}{n^2} = \frac{\lambda}{n}$, for $i \leq 3$ and $\lambda n \geq 4$).

Consequently, the expected value of X is

$$\mathbb{E}[X] = \sum_{1=1}^{k} \text{Prob}(B_i \text{ appears in } R) \leq kp \leq cn^5 \frac{\lambda^4}{n^4} = c\lambda^4 n \leq \frac{\lambda n}{2},$$

provided that $2c\lambda^3 \leq 1$. We can now set $\lambda = (2c)^{-1/3}$, and obtain $\mathbb{E}[X] \leq \lambda n/2$. By the standard expectation argument, there exists $R \subset G_n$, $|R| = \lambda n$, with $|X| \leq \lambda n/2$. Since $R \subset G_n$, R determines $O(n^2/\sqrt{\log n})$ distinct distances.

In the second step, we employ the *deletion* method (see, e.g., [1, Ch. 3]), and delete one point of R from each bad 4-element subset determined by R. Since there are at most $\lambda n/2$ bad 4-element subsets in R, at most $\lambda n/2$ points (out of λn) are deleted, no bad 4-element subsets are introduced, and the remaining (at least) $\lambda n/2 = \Omega(n)$ points in R determine no bad 4-element subset. As such, there exists an $\Omega(n)$-element point set in the plane without the seven patterns $\pi_1, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7, \pi_8$, that determines $O(n^2/\sqrt{\log n})$ distinct distances. This concludes the proof of Theorem 1. □

5 Conclusion

A point set is said to have property $\Phi(4,5)$ if every 4 elements determine at least 5 distinct distances. An examination (Lemma1 in Section 2) indicates that a planar point set has property $\Phi(4,5)$ if and only if it does not contain any of eight forbidden 4-element patterns, π_i, $i = 1, \ldots, 8$. The existence of n-element point sets without the three patterns π_1, π_2, π_3, that determine only $o(n^2)$ distinct distances was previously known. Here we established the existence of n-element point sets without the seven patterns $\pi_1, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7, \pi_8$, and determining $o(n^2)$ distinct distances.

A challenge for our construction seems to be limiting the number of parallelograms; as shown in Lemma 3, $G_{\sqrt{n}}$ determines a large number, $\Theta(n^3)$. It is worth noting that while the existence of parallelograms appears as a serious obstacle for the randomized construction, their absence is guaranteed in the earlier deterministic construction (point set S_n from [3], also described in Section 2). The existence of point sets missing all eight forbidden patterns and determining only $o(n^2)$ distinct distances remains open; as such, the conjecture of Erdős currently remains unsettled. We have shed some light on the problem of distinct distances in point sets with property $\Phi(4,5)$ and perhaps made some partial progress towards its resolution. Three relevant questions are:

Problem 1. *Is there a $\Theta(n)$ size subset of G_n that avoids all eight forbidden patterns?*

Problem 2. *Is there a $\Theta(n)$ size subset of S_n that avoids all eight forbidden patterns?*

Problem 3. *Is $\phi(n,4,5)$ super-linear in n?*

It is conceivable that all these questions have positive answers. In any case, reducing the gaps between the upper and lower bounds in the first five entries of Table 1 remain as future challenges. This table should *not* by any means be interpreted as containing all interesting questions that one could formulate in this regard. In particular, the lack of entries with subquadratic upper bounds deserves attention.

References

[1] N. Alon and J. Spencer, *The Probabilistic Method*, fourth edition, Wiley, New York, 2016.

[2] P. Braß, W. Moser, and J. Pach, *Research Problems in Discrete Geometry*, Springer, New York, 2005.

[3] A. Dumitrescu, On distinct distances among points in general position and other related problems, *Periodica Mathematica Hungarica* **57(2)** (2008), 165–176.

[4] P. Erdős, On sets of distances of n points, *American Mathematical Monthly* **53** (1946), 248–250.

[5] P. Erdős, On some problems of elementary and combinatorial geometry, Annali di Matematica pura ed applicata **IV** Vol. CIII (1975), 99–108.

[6] P. Erdős, Extremal problems in number theory, combinatorics and geometry, in *Proceedings of the International Congress of Mathematicians* August 16-24, 1983, Warszawa, pp. 51–70.

[7] P. Erdős, On some metric and combinatorial geometric problems, *Discrete Mathematics* **60** (1986), 147–153.

[8] P. Erdős, Some old and new problems in combinatorial geometry, *Applications of Discrete Mathematics* (Clemson, SC, 1986), pp. 32–37, SIAM, Philadelphia, PA, 1988.

[9] P. Erdős, Some of my recent problems in combinatorial number theory, geometry and combinatorics, in *Graph Theory, Combinatorics, Algorithms and Applications, Vol. 1* (Y. Alavi et al., editors), Wiley 1995, 335–349.

[10] P. Erdős and A. Gyárfás, A variant of the classical Ramsey problem, *Combinatorica* **15** (1997), 459–467.

[11] P. Erdős and G. Purdy, Some extremal problems in geometry III, *Congressus Numerantium* **14** (Proc. 6th South-Eastern Conf. on Combinatorics, Graph Theory, and Computing), 1975, pp. 291–308.

[12] J. Fox, J. Pach, and A. Suk, More distinct distances under local conditions, *Combinatorica* **38(2)** (2018), 501–509.

[13] L. Guth and N. Katz, On the Erdős distinct distances problem in the plane, *Annals of Mathematics* **181** (2015), 155–190.

[14] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, fifth edition, Oxford University Press, 1979.

[15] J. Matoušek, *Lectures on Discrete Geometry*, Springer, 2002.

[16] J. Pach and P. K. Agarwal, *Combinatorial Geometry*, Wiley-Interscience, New York, 1995.

[17] J. Pach and M. Sharir, Geometric incidences, in *Towards a Theory of Geometric Graphs* (J. Pach, editor), Contemporary Mathematics, Vol. 342, American Mathematical Society, Providence, RI, 2004.

[18] C. Pohoata and A. Sheffer, Higher distance energies and expanders with structure, preprint, Sept. 2017; `arXiv:1709.06696`.

[19] M. Sharir, Computational Geometry Column 65, *SIGACT News Bulletin* **48(2)** (2017), 68–85.

[20] A. Sheffer, Distinct distances: open problems and current bounds, preprint, July 2018; `arXiv:1406.1949`.

[21] J. Spencer, E. Szemerédi and W. T. Trotter, Unit distances in the Euclidean plane, in *Graph Theory and Combinatorics* (B. Bollobás, editor), Academic Press, London, 1984, pp. 293–303.

[22] L. Székely, Crossing numbers and hard Erdős problems in discrete geometry, *Combinatorics, Probability and Computing* **6** (1997), 353–358.

[23] E. Szemerédi and W. T. Trotter, Extremal problems in discrete geometry, *Combinatorica* **3** (1983), 381–392.

[24] T. Thiele, The no-four-on-circle problem, *Journal of Combinatorial Theory Ser. A* **71** (1995), 332–334.