POINT-SETS WITH MANY SIMILAR COPIES OF A PATTERN: WITHOUT A FIXED NUMBER OF COLLINEAR POINTS OR IN PARALLELOGRAM-FREE POSITION

BERNARDO M. ÁBREGO AND SILVIA FERNÁNDEZ-MERCHANT

Department of Mathematics
California State University, Northridge
18111 Nordhoff St, Northridge, CA 91330-8313.
email:{bernardo.abrego,silvia.fernandez}@csun.edu

Abstract. Let P be a finite *pattern*, that is, a finite set of points in the plane. We consider the problem of maximizing the number of similar copies of P over all sets of n points in the plane under two general position restrictions: (1) Over all sets of n points with no m points on a line. We call this maximum $S_P(n,m)$. (2) Over all sets of n points with no collinear triples and not containing the 4 vertices of any parallelogram. These sets are called *parallelogram-free* and the maximum is denoted by $S_P^{\nparallel}(n)$. We prove that $S_P(n,m) \geq n^{2-\varepsilon}$ whenever $m(n) \to \infty$ as $n \to \infty$ and that $\Omega(n \log n) \leq S_P^{\nparallel}(n) \leq O(n^{3/2})$.

1. Introduction

All sets considered in this paper are finite subsets of the plane, which we identify with the set of complex numbers $\mathbb{C}$. We say that the sets A and B are *similar*, denoted by $A \sim B$, if there exist complex numbers w and $z \neq 0$ such that $B = zA + w$. Here, $zA = \{za : a \in A\}$ and $A + w = \{a + w : a \in A\}$.

Let P be a fixed set of at least 3 points, P is called a *pattern.* For any set Q, we denote by $S_P(Q)$ the number of similar copies of P contained in Q, that is,

$$S_P\left(Q\right) = \left|\{P' \subseteq Q : P' \sim P\}\right|.$$

2000 *Mathematics Subject Classification.* Primary 52C10, Secondary 05C35.
Key words and phrases. Similar copy, pattern, geometric pattern, general position, parallelogram-free, Minkovski Sum.

We are interested in the maximum of $S_P(Q)$ over all n-point sets Q satisfying certain restrictions.

Erdős and Purdy [8]-[10] considered $S_P(n) = \max S_P(Q)$ where the maximum is taken over all n-point sets Q. Elekes and Erdős proved [6] that this function is quadratic or close to quadratic (depending on the pattern P); and later Elekes and the authors [4] showed that the optimal sets, when $S_P(n)$ is quadratic, contain many collinear points. In [3] we considered the maximum $S'_P(n) = \max S_P(Q)$ over all n-point sets in general position, that is, with no three collinear points. In the first part of this paper, we relax the general position condition to sets Q without $m \geq 4$ collinear points and consider the corresponding maximum

$$S_P(n,m) = \max\{S_P(Q) : |Q| = n \text{ and no } m \text{ collinear points in } Q\}.$$

Note that $S'_P(n) = S_P(n,3)$ and

$$S'_P(n) \leq S_P(n,m_1) \leq S_P(n,m_2) \leq S_P(n)$$

whenever $m_1 \leq m_2$. Moreover, if m is fixed, then by Corollary 1 in [4] the function $S_P(n,m)$ is subquadratic, i.e., $\lim_{n\to\infty} S_P(n,m)/n^2 = 0$.

The general lower bound for $S'_P(n)$ in Theorem 1 of [3] still holds for $S_P(n,m)$ as long as no m points in P are collinear. More precisely, if $\mathrm{Iso}^+(P)$ denotes the group of orientation-preserving isometries of the pattern P, also known as the *proper symmetry group* of P; and the *index* of a point set A with respect to the pattern P, denoted by $i_P(A)$, is equal to $i_P(A) = \log(|\mathrm{Iso}^+(P)|S_P(A) + |A|)/\log|A|$, then

Theorem 1. *For any A and P finite sets in the plane with at most $m-1$ collinear points, there is a constant $c = c(P,A)$ such that, for n large enough,*

$$S_P(n,m) \geq cn^{i_P(A)}.$$

The proof is the same as that in [3] because Lemma 2 in that paper was actually proved in greater generality for sets with no m points on a line.

We believe that the true asymptotic value of $S_P(n,m)$ is close to quadratic for every pattern P with no m collinear points. In Section 2, we briefly explore the behavior of the function $S_P(n,m)$ for larger values of m and $P = \triangle$ the equilateral triangle. We then present general asymptotic results, for arbitrary patterns, when $m = m(n)$ is a function of n such that $m(n) \to \infty$ when $n \to \infty$. In this case we prove in Section 3 that $S_P(n,m) \geq n^{2-\varepsilon}$ for every n sufficiently large.

All constructions in [3] are obtained by applying Minkovski sums to suitable initial sets. The outcome is a set with multiple quadruples forming the vertex set of a parallelogram. In the second part of this paper, we strengthen the general position condition to *parallelogram-free* position, i.e., no three collinear points or parallelogram's vertex set. We are able to construct parallelogram-free n-sets Q with $\Omega(n \log n)$ copies of a parallelogram-free pattern P. We also show a non-trivial upper bound for these patterns, namely, we prove that at most $O(n^{3/2})$ copies of P are possible. These results are presented in Section 4.

We note that Erdős' unit distance problem [7] has also been studied under the parallelogram-free restriction [13] (see also [5, Section 5.5]).

2. $S_P\,(n,m)$ for an equilateral triangle P

When $m \geq 4$, the initial sets A_m with the largest indices we know are clusters of points of the equilateral triangle lattice in the shape of a circular disk.

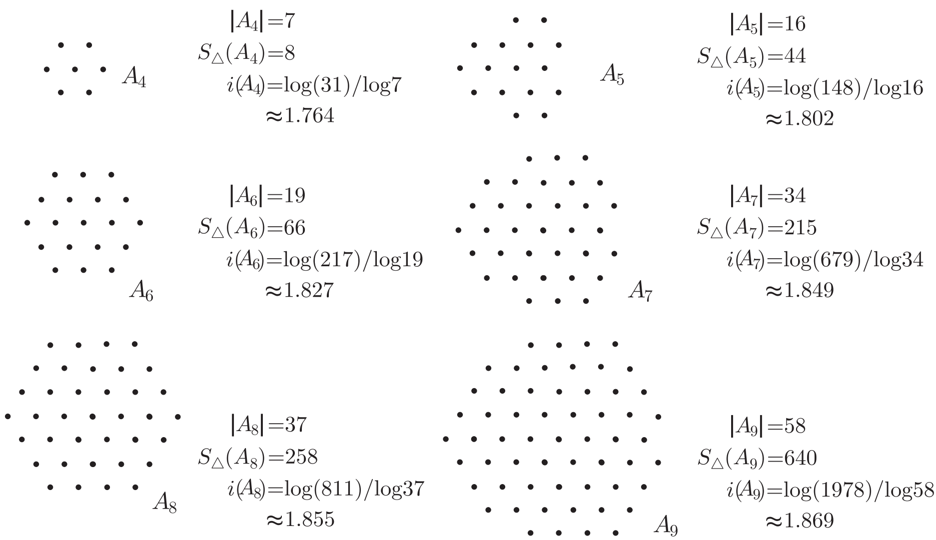

FIGURE 1. Best known constructions of initial sets A_m with many equilateral triangles and at most $m-1$ points on a line.

Theorem 2. *For $m \geq 4$ and $P = \triangle$ the equilateral triangle,*

$$S_\triangle(n,m) \geq \Omega\left(n^{i_\triangle(A_m)}\right)$$

where

$$\begin{aligned} i_\triangle(A_4) &\geq \tfrac{\log 31}{\log 7} \geq 1.764, & i_\triangle(A_5) &\geq \tfrac{\log 148}{\log 16} \geq 1.802, \\ i_\triangle(A_6) &\geq \tfrac{\log 217}{\log 19} \geq 1.827, & i_\triangle(A_7) &\geq \tfrac{\log 679}{\log 34} \geq 1.849, \\ i_\triangle(A_8) &\geq \tfrac{\log 811}{\log 37} \geq 1.855, & i_\triangle(A_9) &\geq \tfrac{\log 1978}{\log 58} \geq 1.869, \end{aligned}$$

and in general, if m is even, then

$$i_\triangle(A_m) = \frac{\log(21m^4 - 84m^3 + 156m^2 - 144m + 64) - \log 64}{\log(3m^2 - 6m + 4) - \log 4}$$

Proof. For the first part refer to Figure 1 where the sets A_m and their corresponding indices are shown. For the second part we consider as our set A_m the lattice points inside a regular hexagon of side $m/2 - 1$ with sides parallel to the lattice. Clearly A_m contains at most $m-1$ collinear points. Also $|A_m| = (3m^2 - 6m + 4)/4$ and $S_\triangle(A_m) = (7m^4 - 28m^3 + 36m^2 - 16m)/64$ (see [1]), therefore the result follows from Theorem 1. $\square$

After some elementary estimations, the index on last theorem satisfies that

$$i_\triangle(A_m) \geq 2 - \frac{\log(12/7)}{2\log m} + \Theta(\log m)^{-2} > 2 - \frac{0.269}{\log m} + \Theta(\log m)^{-2}.$$

This suggests that $S_\triangle(n,m) \geq \Omega\left(n^{2-0.269/\log m}\right)$, however the constant term hidden in the Ω may depend on A, and thus on m. We see this with more detail on the next section.

3. When m grows together with n

We investigate the function $S_P(n,m)$ when the pattern P is fixed and $m = m(n) \to \infty$ when $n \to \infty$. For instance, the maximum number of squares in a n-point set without $\log n$ points on a line, is at least $\Omega(n^{2-c(\log\log n)^{-1}})$. For the proof of our result, we use the following two theorems which give the best bounds for the function $S_P(n)$ without restrictions.

Theorem A (Elekes and Erdős [6]). *For any pattern P there are constants $a, b, c > 0$ such that*

$$S_P(n) \geq cn^{2-a(\log n)^{-b}},$$

moreover, if the coordinates of P are algebraic or if $|P| = 3$, then $S_P(n) \geq cn^2$.

If $u, v, w, z \in \mathbb{C}$ then the *cross-ratio* of the 4-tuple (u, v, w, z) is defined as

$$\frac{(w-u)(z-v)}{(z-u)(w-v)}.$$

Theorem B (Laczkovich and Ruzsa [12]). *$S_P(n) = \Theta(n^2)$ if and only if the cross-ratio of every 4-tuple in P is algebraic.*

For the sake of clarity, let us call a pattern P *cross-algebraic* if the cross-ratio of every 4-tuple is algebraic, and *cross-transcendental* otherwise.

Theorem 3. *Let P be an arbitrary pattern and suppose $m = m(n) \to \infty$, then for every $\varepsilon > 0$ there is a threshold function $N_0 = N_0(\varepsilon, P)$ such that*

$$S_P(n, m) \geq n^{2-\varepsilon} \text{ for every } n \geq N_0.$$

Proof. Suppose $m = m(n) \to \infty$. We actually prove the following stronger result.

(i) If $\log m \leq \sqrt{\log n}$ and P is cross-algebraic, then there is a constant $c_1 > 0$ depending only on P such that

$$S_P(n, m) \geq \Omega(n^{2-c_1/\log m}).$$

(ii) If $\log m \leq \sqrt{\log n}$ and P is cross-transcendental, then there are constants $c_1, c_2, c_3 > 0$ depending only on P such that

$$S_P(n, m) \geq \Omega(n^{2-c_2/(\log m)^{c_3} - c_1/\log m}).$$

(iii) If $\log m > \sqrt{\log n}$, then $S_P(n, m) \geq S_P(n, e^{\sqrt{\log n}})$ and thus either (i) or (ii) holds with $\log m = \sqrt{\log n}$.

Let us recall that, as part of the proof of Theorem 1 (see [3]), we constructed a set A_j^* with $|A_j^*| = |A|^j$ such that $I \cdot S_P(A_j^*) + |A|^j \geq (I \cdot S_P(A) + |A|)^j$, where $I = \mathrm{Iso}^+(P)$. It follows that for $|A|^j \leq n < |A|^{j+1}$,

$$S_P(n, m) \geq S_P(A_j^*) \geq \frac{1}{I}((I \cdot S_P(A) + |A|)^j - |A|^j) \quad (1)$$

$$\geq \frac{1}{I}\left(|A|^{j \cdot i_P(A)} - n\right) \geq \frac{1}{I}\left(\left(\frac{n}{|A|}\right)^{i_P(A)} - n\right) \quad (2)$$

Now, we first prove (i). Suppose that P is cross-algebraic. By Theorem B there is a constant c, depending only on P, and a $(\lceil m\rceil-1)$-set A such that $|A| + S_P(A) \geq cm^2$. Clearly A does not have m points on a line. By Inequality (2), we have that

$$\begin{aligned} S_P(n,m) &\geq \frac{1}{I}\left(\left(\frac{n}{|A|}\right)^{i_P(A)} - n\right) \\ &> \frac{1}{I}\left(\left(\frac{n}{m}\right)^{\frac{\log(cm^2)}{\log m}} - n\right) = \frac{1}{Ic}\left(n^{2+\frac{\log c}{\log m}-\frac{2\log m}{\log n}} - cn\right). \end{aligned}$$

By assumption, $2\log m/\log n \leq 2/\log m$. Because $c<1$ it follows that $\log c < 0$. Let $c_1 = 2 - \log c > 0$, then

$$S_P(n,m) \geq \frac{1}{Ic}\left(n^{2+\frac{\log c}{\log m}-\frac{2}{\log m}} - n\right) \geq \frac{1}{Ic}\left(n^{2-c_1/\log m} - n\right).$$

That is, $S_P(n,m) \geq \Omega(n^{2-c_1/\log m})$, where the constant in the Ω term does not depend on n or m. Similarly, to prove (ii), assume P is cross-transcendental, then by Theorem A there are constants $c, c_2, c_3 > 0$, depending only on P, such that $S_P(n) \geq cn^{2-c_2/(\log n)^{c_3}}$. Then there is a $(\lceil m\rceil-1)$-set A such that $|A|+S_P(A) \geq cm^{2-c_2/(\log m)^{c_3}}$. Again A does not have m points on a line and setting $c_1 = 2-\log c$ we get

$$\begin{aligned} S_P(n,m) &\geq \frac{1}{I}\left(\left(\frac{n}{|A|}\right)^{i_P(A)} - n\right) \geq \frac{1}{I}\left(\left(\frac{n}{m}\right)^{2+\frac{\log c}{\log m}-\frac{c_2}{(\log m)^{c_3}}} - n\right) \\ &\geq \frac{1}{Ic}\left(n^{2-c_2/(\log m)^{c_3}-c_1/\log m} - cn\right) \end{aligned}$$

for n and m large enough depending only on P. That is, $S_P(n,m) \geq \Omega(n^{2-c_2/(\log m)^{c_3}-c_1/\log m})$, where the constant in the Ω term does not depend on n or m. □

If m grows like a fixed power of n and P is cross-algebraic (equivalently by Theorem B, $S_P(n) = \Theta(n^2)$), then we can improve our bound.

Theorem 4. *If P is cross-algebraic and $m = m(n) \geq n^\alpha$ for some fixed $0<\alpha<1$, then there is $c_1 = c_1(P,\alpha) > 0$ such that*

$$S_P(n,m) \geq c_1 n^2 \text{ for every } n \geq |P|.$$

Proof. Choose an integer $j \geq 2$ such that $\alpha > 1/j$. Consider an optimal set A for the function $S_P(\lfloor n^{1/j} \rfloor$. Then $|A| = \lfloor n^{1/j} \rfloor$ and by Theorem B, there is a constant $c = c(P)$ such that $S_P(A) = S_P(|A|) \geq c|A|^2$. Since $|A| \leq n^{1/j} < n^\alpha \leq m$, then A has no m collinear points. By Inequality (1),

$$\begin{aligned} S_P(n, m) \geq S_P(|A|^j, m) &\geq \frac{1}{I}\left((I \cdot S_P(A) + |A|)^j - |A|^j\right) \\ &\geq I^{j-1} S_P(A)^j \geq c^j I^{j-1} |A|^{2j}. \end{aligned}$$

Now, if $n \geq |P|$ then $|A| \geq n^{1/j} - 1 \geq (1 - |P|^{-1/j})n^{1/j}$. By letting $c_1 = c^j I^{j-1}(1 - |P|^{-1/j})^{2j}$ we get $S_P(n, m) \geq c_1 n^2$. □

4. Parallelogram-free sets

We consider the restriction of the function $S_P(n)$ to sets of points A in general position (no 3 points on a line) and without parallelograms. We say that such a set A is *parallelogram-free*. This immediately prohibits the use of Minkovski Sums to obtain good constructions. More precisely, for a parallelogram-free pattern P, define

$$S_P^{\nparallel}(n) = \max\left\{S_P(A) : |A| = n \text{ and } A \text{ is parallelogram-free}\right\}.$$

We obtain the following upper bound on $S_P^{\nparallel}(n)$.

Theorem 5. *Let P be a parallelogram-free pattern with $|P| \geq 3$. Then for all n,*

$$S_P^{\nparallel}(n) \leq n^{3/2} + n.$$

Proof. Suppose A is an n-set in the plane in general position and with no parallelograms. Let p_1, p_2, p_3 be three points in P. Consider the following bipartite graph B. The vertex bipartition is (A, A); the edges are the pairs $(a_1, a_2) \in A \times A$, $a_1 \neq a_2$ such that there is a point $a_3 \in A$ with $\triangle a_1a_2a_3 \sim \triangle p_1p_2p_3$. Every similar copy of P in A has at least one edge (a_1, a_2) associated to it. Thus the number of edges E in our graph satisfies that $E \geq S_P(A)$. By a theorem of Kővari et al. [11] (also referred in the literature as Zarankiewicz problem [14]), it is known that a bipartite graph with n vertices on each class and without subgraphs isomorphic to $K_{2,2}$ contains at most $(n-1)n^{1/2} + n$ edges. To finish our proof we now show that B has no subgraphs isomorphic to $K_{2,2}$. Suppose by contradiction that $(a_1, a_3), (a_1, a_4), (a_2, a_3), (a_2, a_4)$ are edges in B. Let $\lambda = (p_3 -$

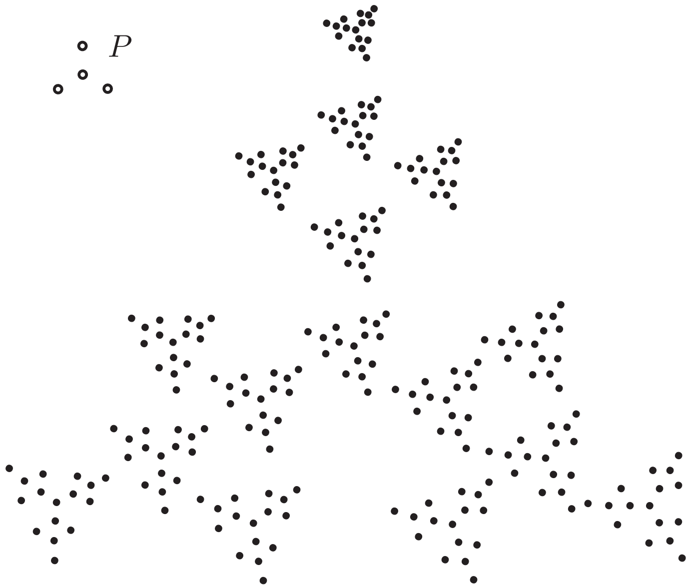

FIGURE 2. A parallelogram-free point set with n points and $cn \log n$ similar copies of P.

$p_1)/(p_2-p_1)$. By definition, there are points $a_{13}, a_{14}, a_{23}, a_{24} \in A$ such that $\triangle p_1p_2p_3 \sim \triangle a_1a_3a_{13} \sim \triangle a_1a_4a_{14} \sim \triangle a_2a_3a_{23} \sim \triangle a_2a_4a_{24}$. Thus $a_{13} = a_1+\lambda(a_3-a_1)$, $a_{14} = a_1+\lambda(a_4-a_1)$, $a_{23} = a_2+\lambda(a_3-a_2)$, and $a_{24} = a_2 + \lambda(a_4 - a_2)$. Then $a_{13} - a_{14} = a_{23} - a_{24} = \lambda(a_3 - a_4)$, which means that $a_{13}a_{14}a_{24}a_{23}$ is a parallelogram. This contradicts the parallelogram-free assumption on A. □

We make no attempt to optimize the coefficient of the $n^{3/2}$ term, since we do not believe that $n^{3/2}$ is the right order of magnitude.

Theorem 6. *Let P be a parallelogram-free pattern with $|P| \geq 3$. Then there is a constant $c = c(P)$ such that for $n \geq |P|$,*

$$S_P^{\nparallel}(n) \geq cn \log n.$$

For every pattern P, we recursively construct a parallelogram-free point set with many occurrences of P. For any $u, v \in \mathbb{C}$, we define

$$\begin{aligned} Q(P, A, u, v) &= \bigcup_{p\in P} (up + (vp - p + 1)\, A) \\ &= \bigcup_{p\in P} \bigcup_{a\in A} (up + (vp - p + 1)\, a)\,. \end{aligned} \tag{3}$$

Almost all selections of u and v yield a set $Q = Q(P, A, u, v)$ that is parallelogram-free and such that all the terms in the double union are pairwise different. The proof of this technical fact is given by next lemma.

Lemma 1. *Let A and P be parallelogram-free sets. If $\mathcal{S}$ is the set of points $(u, v) \in \mathbb{C}^2$ for which $Q = Q(P, A, u, v)$ satisfies that $|Q| < |A|\,|P|$, Q has three collinear points, or Q has a parallelogram; then $\mathcal{S}$ has zero Lebesgue measure.*

We defer the proof of this lemma and instead proceed to bound the number of similar copies of P in Q.

Lemma 2. *If A and P are finite parallelogram-free sets , and $Q = Q(P, A, u, v)$ defined in (3) satisfies that $|Q| = |A|\,|P|$, then*

$$S_P(Q) \geq |P|\, S_P(A) + |A|\,.$$

Proof. Because $|Q| = |A|\,|P|$ it follows that each term in the first union contributes exactly $S_P(A)$ similar copies of P, all of them pairwise different. In addition note that

$$Q = \bigcup_{a \in A} \left(a + (u + va - a)\, P\right).$$

So each term in the new union is a similar copy of P, all of them different and also different from the ones we had counted before. Therefore $S_P(Q) \geq |P|\, S_P(A) + |A|$. □

We now prove the theorem.

Proof of Theorem 6. Let $A_1 = P$ and for $m \geq 1$ let A_{m+1} be the parallelogram-free set Q obtained from Lemma 1 with $A = A_m$. Because $|A_{m+1}| = |A_1|\,|A_m|$, it follows that $|A_m| = |P|^m$ for all m. Further, by Lemma 2, for every $0 \leq k \leq m-2$, $S_P(A_{m-k}) \geq |P|\, S_P(A_{m-k-1}) + |A_{m-k-1}| = |P|\, S_P(A_{m-k-1}) + |P|^{m-k-1}$. Thus

$$\begin{aligned} S_P(A_m) &\geq |P|\, S_P(A_{m-1}) + |P|^{m-1} \geq |P|^2\, S_P(A_{m-2}) + 2\,|P|^{m-1} \\ &\geq \cdots \geq |P|^{m-1}\, S_P(A_1) + (m-1)\,|P|^{m-1} = m\,|P|^{m-1}\,. \end{aligned}$$

Suppose $|P|^m \leq n < |P|^{m+1}$ with $m \geq 2$. Let $c = 1/(2|P|^2 \log |P|)$, then

$$S_P^{\nparallel}(n) \geq S_P(A_m) \geq m\,|P|^{m-1} > \left(\frac{\log n}{\log |P|} - 1\right) \frac{n}{|P|^2} \geq cn \log n.$$

If $|P| \leq n < |P|^2$, then $S_P^{\nparallel}(n) \geq S_P(P) = 1 > cn \log n$. □

Finally, we present the proof of Lemma 1.

Proof of Lemma 1. We show that $\mathcal{S}$ is made of a finite number of algebraic sets, all of them of real dimension at most three. This immediately implies that the Lebesgue measure of such a set is zero. For every $p \in P$ and $a \in A$, let $q(a,p) = (up + (vp - p + 1)\,a)$. Suppose that $q(a_1,p_1) = q(a_2,p_2)$ with $(a_1,p_1) \neq (a_2,p_2)$. Then

$$(p_1 - p_2)\,u + (p_1a_1 - p_2a_2)\,v + a_1(1-p_1) - a_2\,(1-p_2) = 0.$$

This is the equation of a complex-line in $\mathbb{C}^2$ (with real-dimension two) unless the coefficients of u and v, as well as the independent term are equal to zero. That is, $p_1 - p_2 = 0$, $(p_1a_1 - p_2a_2) = 0$, and $a_1(1-p_1) - a_2\,(1-p_2) = 0$. These equations imply that $(a_1,p_1) = (a_2,p_2)$ which contradicts our assumption. Thus the set of pairs (u,v) for which $|Q| < |A|\,|P|$ is the union of $\binom{|A||P|}{2}$ sets of real-dimension two. Assume that $q(a_1,p_1), q(a_2,p_2)$, and $q(a_3,p_3)$ are three collinear points. Thus there is a real $\lambda \neq 0,1$ such that $q(a_2,p_2) - q(a_1,p_1) = \lambda\,(q(a_3,p_3) - q(a_1,p_1))$. Then

$$\begin{aligned}(p_2 - p_1 - \lambda(p_3 - p_1))\,u + (p_2a_2 - p_1a_1 - \lambda(p_3a_3 - p_1a_1))\,v + \\ a_2\,(1-p_2) - a_1(1-p_1) - \lambda\,(a_3\,(1-p_3) - a_1(1-p_1)) = 0.\end{aligned}$$

For every λ, the last equation represents a complex-line in $\mathbb{C}^2$. Considering λ as a real variable, this equation represents an algebraic set of real-dimension three. This happens unless the coefficients of u and v, as well as the independent term are equal to zero. That is, $p_2 - p_1 - \lambda(p_3 - p_1) = 0$, $p_2a_2 - p_1a_1 - \lambda(p_3a_3 - p_1a_1) = 0$, and $a_2 - a_1 - \lambda\,(a_3 - a_1) = 0$. If p_1, p_2, p_3 are three different points then, since no three points in P are collinear, $p_2 - p_1 - \lambda(p_3 - p_1) \neq 0$. If any two of p_1, p_2, p_3 are equal and the remaining is different, then we still have $p_2 - p_1 - \lambda(p_3 - p_1) \neq 0$. Thus $p_1 = p_2 = p_3$, and by symmetry $a_1 = a_2 = a_3$; which contradicts the fact that we started with three distinct points $q(a_j,p_j)$. Thus the set of pairs (u,v) for which $|Q|$ has collinear points is the union of $\binom{|A||P|}{3}$ sets of real-dimension three. Finally, assume that $q(a_1,p_1)q(a_2,p_2)q(a_4,p_4)q(a_3,p_3)$ is a parallelogram. That is $q(a_2,p_2) - q(a_1,p_1) = q(a_4,p_4) - q(a_3,p_3)$, and thus

$$\begin{aligned}(p_2 - p_1 - (p_4 - p_3))\,u + (p_2a_2 - p_1a_1 - (p_4a_4 - p_3a_3))\,v + \\ a_2\,(1-p_2) - a_1(1-p_1) - (a_4\,(1-p_4) - a_3(1-p_3)) = 0.\end{aligned}$$

Again this equation represents a complex-line in $\mathbb{C}^2$, unless the coefficients of u and v, as well as the independent term are equal to zero. That is, $p_2 - p_1 - (p_4 - p_3) = 0$, $p_2a_2 - p_1a_1 - (p_4a_4 - p_3a_3) = 0$, and $a_2 - a_1 - (a_4 - a_3) = 0$. If p_1, p_2, p_3, p_4 are four different points then, since P has no parallelograms, $p_2 - p_1 - (p_4 - p_3) \neq 0$. Since no three points of P are collinear there are two extra possibilities: $(p_1, p_2) = (p_3, p_4)$ or $(p_1, p_3) = (p_2, p_4)$. By symmetry we also have $(a_1, a_2) = (a_3, a_4)$ or $(a_1, a_3) = (a_2, a_4)$. If $(p_1, p_2) = (p_3, p_4)$ and $(a_1, a_2) = (a_3, a_4)$, then $q(a_1, p_1) = q(a_3, p_3)$; which contradicts our assumption. Similarly, if $(p_1, p_3) = (p_2, p_4)$ and $(a_1, a_3) = (a_2, a_4)$, then $q(a_1, p_1) = q(a_2, p_2)$. Assume $(p_1, p_2) = (p_3, p_4)$ and $(a_1, a_3) = (a_2, a_4)$. Then the equation $p_2a_2 - p_1a_1 - (p_4a_4 - p_3a_3) = 0$ becomes $(p_2 - p_1)(a_1 - a_3) = 0$. But if $p_1 = p_2$ then $q(p_1, a_1) = q(p_2, a_2)$, and if $a_1 = a_3$ then $q(p_1, a_1) = q(p_3, a_3)$; a contradiction in both cases. The remaining case when $(p_1, p_3) = (p_2, p_4)$ and $(a_1, a_2) = (a_3, a_4)$ follows by symmetry. Thus the set of pairs (u, v) for which $|Q|$ has parallelograms is the union of $\binom{|A||P|}{4}$ sets of real-dimension two. $\square$

Remark 1. *If the sets A and P have no two parallel segments then it can be proved, along the lines of last lemma, that almost all the sets $Q(A, P, u, v)$ are free of pairs of parallel segments as well.*

5. Concluding Remarks and Conjectures

The construction of Theorem 2 in [3] can be carried out the same way for a pattern P with no m points on a line. The result would be a set A with $|A| = k^2 - k + 1$ and $S_P(A) = 2k - 1$. Then Theorem 1 would give

$$S_P(n, m) \geq \Omega(n^{\log(k^2+k)/\log(k^2-k+1)}).$$

There should be a better general construction, particularly when m is larger than the maximum number of collinear points in P.

Problem 1. *Let P be an arbitrary pattern with $|P| = k$. For every $m \geq 4$ construct a set A without m collinear points such that $i_P(A) > \log(k^2 + k)/\log(k^2 - k + 1)$.*

According to Theorem 3, if we let the number of allowed collinear points to increase with n, then we can achieve $n^{2-\varepsilon}$ similar copies of a pattern P. We actually believe this is true even when m is constant.

Conjecture 1. *Let $m \geq 3$ be a positive integer and P a finite pattern with no m collinear points. For every real $\varepsilon > 0$, there is $N(\varepsilon) > 0$ such that for all $n \geq N(\varepsilon)$,*

$$S_P(n, m) \geq n^{2-\varepsilon}.$$

A proof of this conjecture cannot follow from Theorem 1, so a proof would require a different way of constructing sets with no m collinear points and with many similar copies of the pattern P. In contrast, we believe that the construction in Theorem 6 for the function $S_P^{\nparallel}(n)$ is close to optimal. Here we believe that a stronger upper bound is needed.

Conjecture 2. *Let P be a parallelogram-free pattern. For every real $\varepsilon > 0$, there is $N(\varepsilon) > 0$ such that for all $n \geq N(\varepsilon)$,*

$$S_P^{\nparallel}(n) \leq n^{1+\varepsilon}.$$

References

[1] B. M. Ábrego, Problemas combinatorios sobre conjuntos finitos de puntos, *Aportaciones Matemáticas: Comunicaciones* **19** (1997), SMM, México.

[2] B. M. Ábrego and S. Fernández-Merchant, On the maximum number of equilateral triangles I, *Discrete and Computational Geometry* **23** (2000), 129–135.

[3] B. M. Ábrego and S. Fernández-Merchant, Point-sets in general position with many similar copies of a pattern, to appear in *Geombinatorics.*

[4] B. M. Ábrego, S. Fernández-Merchant, and Gy. Elekes. Structural results for planar sets with many similar subsets, *Combinatorica* **24**(4) (2004), 541–554.

[5] P. Brass, W. Moser, J. Pach, *Research Problems in Discrete Geometry*, Springer-Verlag, New York, 2005.

[6] G. Elekes and P. Erdős, Similar configurations and pseudogrids, in *Intuitive Geometry*, Colloquia Mathematica Societatis János Bolyai. North Holland, Amsterdam, (1994), 85–104.

[7] P. Erdős, On sets of distances of n points, *Amer. Math. Monthly* **53** (1946) 248-250.

[8] P. Erdős and G. Purdy. Some extremal problems in geometry. *Journal of Combinatorial Theory* **10** (1971), 246–252.

[9] P. Erdős and G. Purdy. Some extremal problems in geometry III. *Proc. 6th Southeastern Conference in Combinatorics, Graph Theory and Comp.* (1975), 291–308.

[10] P. Erdős and G. Purdy. Some extremal problems in geometry IV. *Proc. 7th Southeastern Conference in Combinatorics, Graph Theory and Comp.* (1976), 307–322.

[11] T. Kővari, V. Sós, and P. Turán, On a problem of K. Zarankiewicz, *Colloquium Mathematicum* **3** (1954), 50–57.

[12] M. Laczkovich and I.Z.Ruzsa, The Number of Homothetic Subsets, *The mathematics of Paul Erdős II*, Springer Verlag (1997), 294–302.

[13] K. J. Swanepoel and P. Valtr, The unit distance problem on spheres, in *Towards a Theory of Geometric Graphs*, J. Pach, ed., *Contemporary Mathematics* **342**, AMS (2004), 273–279.

[14] K. Zarankiewicz, Problem 101, *Colloquium Mathematicum* **2** (1951), 301.