Almost Sharp Bounds on the Number of Discrete Chains in the Plane

Nóra Frankl
Department of Mathematics, London School of Economics and Political Science, UK
Laboratory of Combinatorial and Geometric Structures at MIPT, Moscow, Russia
n.frankl@lse.ac.uk

Andrey Kupavskii
Moscow Institute of Physics and Technology, Moscow, Russia
Institute for Advanced Study, Princeton, NJ, US
G-SCOP, CNRS, Grenoble, France
kupavskii@yandex.ru

Abstract

The following generalisation of the Erdős unit distance problem was recently suggested by Palsson, Senger and Sheffer. For a sequence $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)$ of k distances, a $(k+1)$-tuple $(p_1, \dots, p_{k+1})$ of distinct points in $\mathbb{R}^d$ is called a $(k, \boldsymbol{\delta})$-*chain* if $\|p_j - p_{j+1}\| = \delta_j$ for every $1 \le j \le k$. What is the maximum number $C_k^d(n)$ of $(k, \boldsymbol{\delta})$-chains in a set of n points in $\mathbb{R}^d$, where the maximum is taken over all $\boldsymbol{\delta}$? Improving the results of Palsson, Senger and Sheffer, we essentially determine this maximum for all k in the planar case. It is only for $k \equiv 1 \pmod 3$ that the answer depends on the maximum number of unit distances in a set of n points. We also obtain almost sharp results for even k in dimension 3.

2012 ACM Subject Classification Mathematics of computing → Combinatoric problems

Keywords and phrases unit distance problem, unit distance graphs, discrete chains

Digital Object Identifier 10.4230/LIPIcs.SoCG.2020.48

Related Version A related version of this paper is available at `https://arxiv.org/abs/1912.00224`.

Funding The authors acknowledge the financial support from the Ministry of Education and Science of the Russian Federation in the framework of MegaGrant no. 075-15-2019-1926.
Nóra Frankl: Research was partially supported by the National Research, Development, and Innovation Office, NKFIH Grant K119670.
Andrey Kupavskii: Research supported by the Russian Foundation for Basic Research (grant no. 18-01-00355) and the Council for the Support of Leading Scientific Schools of the President of the Russian Federation (grant no. HW-6760.2018.1). Research was directly supported by the IAS Fund for Math, the Director's Fund and indirectly supported by the National Science Foundation Grant No. CCF-1900460. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Acknowledgements We thank Konrad Swanepoel and the referees for helpful comments on the manuscript.

1 Introduction

Determining the maximum possible number of pairs $u_d(n)$ at distance 1 apart in a set of n points in $\mathbb{R}^d$ for $d = 2, 3$ is one of the central questions in combinatorial geometry. The planar version, determining $u_2(n)$ is also known as the Erdős unit distances problem. The question dates back to 1946, and despite much effort, the best known upper and lower bounds are still very far apart. For some constants $C, c > 0$, we have

$$n^{1+c/\log\log n} \le u_2(n) \le Cn^{4/3},$$

where the lower bound is due to Erdős [3] and the upper bound is due to Spencer, Szemerédi and Trotter [9].

36th International Symposium on Computational Geometry (SoCG 2020).
Editors: Sergio Cabello and Danny Z. Chen; Article No. 48; pp. 48:1–48:15
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

As in the planar case, the best known upper and lower bounds in the 3-dimensional case are also far apart. For some $c, C > 0$, we have

$$cn^{4/3} \log\log n \leq u_3(n) \leq Cn^{295/137+\varepsilon}, \tag{1}$$

where the lower bound is due to Erdős [4], and the upper bound is due to Zahl [10]. The latter is a recent improvement upon the upper bound $O(n^{3/2})$ by Kaplan, Matoušek, Safernová, and Sharir [5], and Zahl [11]. In contrast, for $d \geq 4$ we have $u_d(n) = \Theta(n^2)$.

Palsson, Senger and Sheffer [8] suggested the following generalisation of the unit distance problem. Let $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_k)$ be a sequence of k positive reals. A $(k+1)$-tuple $(p_1, \ldots, p_{k+1})$ of distinct points in $\mathbb{R}^d$ is called a *$(k, \boldsymbol{\delta})$-chain* if $\|p_i - p_{i+1}\| = \delta_i$ for all $i = 1, \ldots, k$. For every fixed k determine $C_k^d(n)$, the maximum number of $(k, \boldsymbol{\delta})$-chains that can be spanned by a set of n points in $\mathbb{R}^d$, where the maximum is taken over all $\boldsymbol{\delta}$. In the planar case, the following upper bounds were found in [8] in terms of the maximum number of unit distances.

▶ **Proposition 1** (Palsson, Senger, and Sheffer [8]).

$$C_k^2(n) = \begin{cases} O\left(n \cdot u_2(n)^{k/3}\right) & \text{if } k \equiv 0 \pmod 3, \\ O\left(u_2(n)^{(k+2)/3}\right) & \text{if } k \equiv 1 \pmod 3, \\ O\left(n^2 \cdot u_2(n)^{(k-2)/3}\right) & \text{if } k \equiv 2 \pmod 3. \end{cases}$$

If $u_2(n) = O(n^{1+\varepsilon})$ for any $\varepsilon > 0$, which is conjectured to hold, then the upper bounds in the proposition above almost match the lower bounds given in Theorem 2. However, as we have already mentioned, determining the order of magnitude of $u_2(n)$ is very far from being done, and in general it proved to be a very hard problem. Thus, it is interesting to obtain "unconditional" bounds, that depend on the value of $u_2(n)$ as little as possible. In [8], the following "unconditional" upper bounds were proved in the planar case.

▶ **Theorem 2** (Palsson, Senger, and Sheffer [8]). *$C_2^2(n) = \Theta(n^2)$, and for every $k \geq 3$ we have*

$$C_k^2(n) = \Omega\left(n^{\lfloor (k+1)/3 \rfloor + 1}\right)$$

and

$$C_k^2(n) = O\left(n^{2k/5+1+\gamma(k)}\right),$$

where $\gamma_k \leq \frac{1}{12}$, and $\gamma_k \to \frac{4}{75}$ as $k \to \infty$.

In our main result, in two-third of the cases we almost determine the value of $C_k^2(n)$, no matter what the value of $u_2(n)$ is, by matching the lower bounds given in Theorem 2. Further, we show that in the remaining cases determining $C_k^2(n)$ essentially reduces to determining the maximum number of unit distances.

▶ **Theorem 3.** *For every integer $k \geq 1$ we have*

$$C_k^2(n) = \tilde{\Theta}\left(n^{\lfloor (k+1)/3 \rfloor + 1}\right) \ \textit{if } k \equiv 0, 2 \pmod 3,$$

and for any $\varepsilon > 0$ we have

$$C_k^2(n) = \Omega\left(n^{(k-1)/3} u_2(n)\right) \ \textit{and } C_k^2(n) = O\left(n^{(k-1)/3+\varepsilon} u_2(n)\right) \ \textit{if } k \equiv 1 \pmod 3.$$

Here and in what follows $f(n) = \tilde{O}(g(n))$ means that there exist positive constants c, C such that $f(n)/g(n) \le C \log^c n$ for every n. We write $f(n) = \tilde{\Omega}(g(n))$ if $g(n) = \tilde{O}(f(n))$, and $f(n) = \tilde{\Theta}(g(n))$ if $f(n) = \tilde{O}(g(n))$ and $g(n) = \tilde{O}(f(n))$.

Let us turn our attention to the 3-dimensional case. The following was proved in [8].

▶ **Theorem 4** (Palsson, Senger, and Sheffer [8]). *For any integer $k \ge 2$, we have*

$$C_k^3(n) = \Omega\left(n^{\lfloor k/2 \rfloor + 1}\right),$$

and

$$C_k^3(n) = \begin{cases} O\left(n^{2k/3+1}\right) & \text{if } k \equiv 0 \pmod 3, \\ O\left(n^{2k/3+23/33+\varepsilon}\right) & \text{if } k \equiv 1 \pmod 3, \\ O\left(n^{2k/3+2/3}\right) & \text{if } k \equiv 2 \pmod 3. \end{cases}$$

We improve their upper bound and essentially settle the problem for even k.

▶ **Theorem 5.** *For any integer $k \ge 2$ we have*

$$C_k^3(n) = \tilde{O}\left(n^{k/2+1}\right).$$

In particular, for even k we have

$$C_k^3(n) = \tilde{\Theta}\left(n^{k/2+1}\right).$$

We also improve the lower bound from Theorem 4 for odd k. Let $us_3(n)$ be the maximum number of pairs at unit distance apart between a set of n points in $\mathbb{R}^3$ and a set of n points on a sphere in $\mathbb{R}^3$.

▶ **Proposition 6.** *Let $k \ge 3$ odd. Then we have*

$$C_k^3(n) = \Omega\left(\max\left\{\frac{u_3(n)^k}{n^{k-1}}, us_3(n)n^{(k-1)/2}\right\}\right).$$

Note that $us_3(n)$ equals the maximum number of incidences between a set of n points and a set of n circles (not necessarily of the same radii) in the plane. Thus we have

$$cn^{4/3} \le us_3(n) = \tilde{O}\left(n^{15/11}\right)$$

(see [1, 2, 6, 7]). Therefore, in general we cannot tell which of the two bounds in Proposition 6 is better. However, for large k the second term is larger than the first due to (1).

Finally, we note that for $d \ge 4$ we have $C_k^d(n) = \Theta(n^{k+1})$. Indeed, we clearly have $C_k^d(n) = O(n^{k+1})$. To see that $C_k^d(n) = \Omega(n^{k+1})$, take two orthogonal circles of radius $1/\sqrt{2}$ centred at the origin and choose $n/2$ points on each of them.

2 Preliminaries

We denote by $u_d(m,n)$ the maximum number of incidences between a set of m points and n spheres[1] of fixed radius in $\mathbb{R}^d$. In other words, $u_d(m,n)$ is the maximum number of red-blue pairs spanning a given distance in a set of m red and n blue points in $\mathbb{R}^d$. By the result of Spencer, Szemerédi and Trotter [9], we have

$$u_2(m,n) = O\left(m^{\frac{2}{3}}n^{\frac{2}{3}} + m + n\right). \tag{2}$$

We say that a point p is *n^α-rich* with respect to a set $P \subseteq \mathbb{R}^d$ and to a distance δ, if the sphere of radius δ around p contains at least n^α points of P. If $P \subseteq \mathbb{R}^2$ and $|P| = n^x$, then (2) implies that the number of points that are n^α-rich with respect to P and to a given distance δ is

$$O\left(n^{2x-3\alpha} + n^{x-\alpha}\right). \tag{3}$$

The bound

$$u_3(m,n) = O\left(m^{\frac{3}{4}}n^{\frac{3}{4}} + m + n\right) \tag{4}$$

is due to Zahl [10] and Kaplan, Matoušek, Safernová, and Sharir [5]. It implies that for $P \subseteq \mathbb{R}^3$ with $|P| = n^x$ the number of points that are n^α-rich with respect to P and to a given distance δ is

$$O\left(n^{3x-4\alpha} + n^{x-\alpha}\right). \tag{5}$$

3 Bounds in $\mathbb{R}^2$

For $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)$ and $P_1 \dots, P_{k+1} \subseteq \mathbb{R}^2$ we denote by $\mathcal{C}_k^{\boldsymbol{\delta}}(P_1, \dots, P_k)$ the family of $(k+1)$-tuples $(p_1, \dots, p_{k+1})$ with $p_i \in P_i$ for all $i \in [k+1]$, $\|p_i - p_{i+1}\| = \delta_i$ for all $i \in [k]$ and with $p_i \neq p_j$ for $i \neq j$. Let $C_k^{\boldsymbol{\delta}}(P_1, \dots, P_{k+1}) = |\mathcal{C}_k^{\boldsymbol{\delta}}(P_1, \dots, P_{k+1})|$ and

$$C_k(n_1, \dots, n_{k+1}) = \max C_k^{\boldsymbol{\delta}}(P_1, \dots, P_{k+1}),$$

where the maximum is taken over all choices of $\boldsymbol{\delta}$ and sets $P_1, \dots, P_{k+1}$ subject to $|P_i| \leq n_i$ for all $i \in [k+1]$.

It is easy to see that $C_k^2(n) \leq C_k(n, \dots, n) \leq C_k^2\left((k+1)n\right)$. Since we are only interested in the order of magnitude of $C_k^2(n)$ for fixed k, we are going to bound $C_k(n, \dots, n)$ instead of $C_k^2(n)$.

In Section 3.1, we are going to prove the lower bounds from Theorem 3. In Section 3.2, we are going to prove an upper bound on $C_k(n, \dots, n)$, which is almost tight for $k \equiv 0, 2 \pmod 3$. The case $k \equiv 1 \pmod 3$ is significantly more complicated. We will the case $k = 4$ case separately in Section 3.3, and then the general case in Section 3.4.

3.1 Lower bounds

For completeness, we present constructions for all congruence classes modulo 3. For $k \equiv 0, 2$ they were described in [8].

[1] circles, if $d = 2$

First, note that $C_0(n) = n$ and $C_1(n,n) = u_2(n,n) = \Theta(u_2(n))$. For $k = 2$, let $P_2 = \{x\}$ for some point x, and let P_1, P_3 be disjoint sets of n points on the unit circle around x. It is not hard to see that $C_2^{\boldsymbol{\delta}}(P_1, P_2, P_3) = n^2$ with $\boldsymbol{\delta} = (1,1)$, implying the lower bound $C_2(n,n,n) = \Omega(n^2)$. To obtain lower bounds in Theorem 3, it is thus sufficient to show that

$$C_{k+3}(n,\dots,n) \geq nC_k(n,\dots,n).$$

To see this take, a construction with $k+1$ parts $P_1,\dots,P_{k+1}$ of size n that contains $C_k(n,\dots,n)$ $(k,\boldsymbol{\delta})$-chains for some $\boldsymbol{\delta} = (\delta_1,\dots,\delta_k)$. Next, fix an arbitrary point x on the plane and choose distances $\delta_{k+1}, \delta_{k+2}$ to be sufficiently large so that x can be connected to each of the points in P_{k+1} by a 2-chain with distances δ_{k+2} and δ_{k+1}. Set $P_{k+3} = \{x\}$ and let P_{k+2} be the set of intermediate points of the 2-chains described above. Finally, let $\delta_{k+3} = 1$, and P_{k+4} be a set of n points (disjoint from P_{k+2}) on the unit circle around x. It is easy to see that the number of $(k+3,\boldsymbol{\delta})$-chains with $\boldsymbol{\delta} = (\delta_1,\dots,\delta_{k+3})$ in $P_1 \times \dots \times P_{k+4}$ is at least $nC_k(n)$.

Note that it is not hard to modify this construction to show that for any given $\boldsymbol{\delta}$ there is a set of n points with $\Omega(n^{k/3+1})$ many $(k,\boldsymbol{\delta})$-chains if $k \equiv 0 \pmod 3$ and with $\Omega(n^{(k+4)/3})$ many $(k,\boldsymbol{\delta})$-chains if $k \equiv 2 \pmod 3$. However, for $k \equiv 1 \pmod 3$, our construction to find sets of n points with $\Omega(n^{(k-1)/3}u_2(n))$ many $(k,\boldsymbol{\delta})$-chains only works if δ_1 is much smaller than δ_2 and δ_3.

3.2 Upper bound for $k \equiv 0, 2 \pmod 3$

We fix $\boldsymbol{\delta} = (\delta_1,\dots,\delta_k)$ throughout the remainder of Section 3 and leave $\boldsymbol{\delta}$ out of the notation. All logs are base 2.

▶ **Theorem 7.** *For any fixed integer $k \geq 0$ and $x, y \in [0,1]$, we have*

$$C_k(n^x, n, \dots, n, n^y) = \tilde{O}\left(n^{\frac{f(k)+x+y}{3}}\right),$$

where $f(k) = k+2$ if $k \equiv 2 \pmod 3$ and $f(k) = k+1$ otherwise.

Theorem 7 implies the upper bounds in Theorem 3 for $k \equiv 0, 2 \pmod 3$ by taking $x = y = 1$. It is easier, however, to prove this more general statement than the upper bounds in Theorem 3 directly. Having varied sizes of the first and the last groups of points allows for a seamless use of induction.

Proof of Theorem 7. The proof is by induction on k. Let us first verify the statement for $k \leq 2$. (Note that, for $k = 0$, we should have $x = y$.) We have

$$C_0(n^x) \leq n^x = O\left(n^{\frac{1+x+y}{3}}\right),$$

$$C_1(n^x, n^y) \leq u_2(n^x, n^y) = O\left(n^{\frac{2}{3}(x+y)} + n^x + n^y\right) = O\left(n^{\frac{2+x+y}{3}}\right), \tag{6}$$

$$C_2(n^x, n, n^y) \leq n^x n^y = O\left(n^{\frac{4+x+y}{3}}\right), \tag{7}$$

where (6) follows from (2) and (7) follows from the fact that each pair (p_1, p_3) can be extended to a 2-chain (p_1, p_2, p_3) in at most 2 different ways.

Next, let $k \geq 3$. Take $P_1, \dots, P_{k+1} \subseteq \mathbb{R}^2$ with $|P_1| = n^x$, $|P_{k+1}| = n^y$, and $|P_i| = n$ for $2 \leq i \leq k$. Denote by $P_2^\alpha \subseteq P_2$ the set of those points in P_2 that are at least n^α-rich but at most $2n^\alpha$-rich with respect to P_1 and δ_1. Similarly, we denote by $P_k^\beta \subseteq P_k$ the set of those points in P_k that are at least n^β-rich but at most $2n^\beta$-rich with respect to P_{k+1} and δ_k.

It is not hard to see that

$$\mathcal{C}_k(P_1, P_2 \dots, P_k, P_{k+1}) \subseteq \bigcup_{\alpha,\beta} \mathcal{C}_k(P_1, P_2^\alpha, P_3, \dots, P_{k-1}, P_k^\beta, P_{k+1}),$$

where the union is taken over all $\alpha, \beta \in \{\frac{i}{\log n} : i = 0, \dots, \lceil \log n \rceil\}$. Since the cardinality of the latter set is at most $\log n + 2$, it is sufficient to prove that for every α and β we have

$$C_k(P_1, P_2^\alpha, P_3, \dots, P_{k-1}, P_k^\beta, P_{k+1}) = \tilde{O}\left(n^{\frac{f(k)+x+y}{3}}\right). \tag{8}$$

To prove this, we consider three cases.

Case 1: $\alpha \geq \frac{x}{2}$**.** By (3) we have $|P_2^\alpha| = O(n^{x-\alpha})$. Therefore the number of pairs $(p_1, p_2) \in P_1 \times P_2^\alpha$ with $\|p_1 - p_2\| = \delta_1$ is at most $O(n^x)$. Since every pair $(p_1, p_2) \in P_1 \times P_2^\alpha$ and every $(k-3)$-chain $(p_4, \dots, p_{k+1}) \in P_4 \times \dots \times P_k^\beta \times P_{k+1}$ can be extended to a k-chain $(p_1, \dots, p_{k+1}) \in P_1 \times \dots \times P_{k+1}$ in at most two different ways, we obtain

$$C_k(P_1, P_2^\alpha, \dots, P_k^\beta, P_{k+1}) \leq 4O(n^x) C_{k-3}(P_4, \dots, P_k^\beta, P_{k+1}).$$

By induction we have

$$C_{k-3}(P_4, \dots, P_k^\beta, P_{k+1}) = \tilde{O}\left(n^{\frac{f(k-3)+1+y}{3}}\right).$$

These two displayed formulas and the fact that $f(k-3) = f(k) - 3$ imply (8).

Case 2: $\beta \geq \frac{y}{2}$**.** By symmetry, this case can be treated in the same way as Case 1.

Case 3: $\alpha \leq \frac{x}{2}$ **and** $\beta \leq \frac{y}{2}$**.** By (3) we have $|P_2^\alpha| = O\left(n^{2x-3\alpha}\right)$ and $|P_k^\beta| = O\left(n^{2y-3\beta}\right)$. The number of $(k-2)$-chains in $P_2^\alpha \times P_3 \times \dots \times P_{k-1} \times P_k^\beta$ is $C_{k-2}(P_2^\alpha, P_3, \dots, P_{k-1}, P_k^\beta)$, and every $(k-2)$-chain $(p_2, \dots, p_k) \in P_2^\alpha \times P_3 \times \dots \times P_{k-1} \times P_k^\beta$ can be extended at most $4n^{\alpha+\beta}$ ways to a k-chain in $P_1 \times P_2^\alpha \times \dots \times P_k^\beta \times P_{k+1}$. Thus

$$C_k(P_1, P_2^\alpha, \dots, P_k^\beta, P_{k+1}) \leq 4n^{\alpha+\beta} C_{k-2}(P_2^\alpha, \dots, P_k^\beta).$$

By induction we have

$$C_{k-2}(P_2^\alpha, \dots, P_k^\beta) = \tilde{O}\left(n^{\frac{f(k-2)+2x-3\alpha+2y-3\beta}{3}}\right).$$

For $k \equiv 0, 2 \pmod 3$ we have $f(k) \geq f(k-2) + 2$, and thus

$$\begin{aligned} C_k(P_1, P_2^\alpha, \dots, P_k^\beta, P_{k+1}) &= \tilde{O}\left(n^{\alpha+\beta} n^{\frac{f(k-2)+2x-3\alpha+2y-3\beta}{3}}\right) \\ &= \tilde{O}\left(n^{\frac{f(k)-2+2x+2y}{3}}\right) = \tilde{O}\left(n^{\frac{f(k)+x+y}{3}}\right). \end{aligned}$$

If $k \equiv 1 \pmod 3$ then $f(k) < f(k-2) + 2$, and thus the argument above does not work. However, we then have $f(k) = f(k-1) + 1$, and we can use the bound

$$C_k(P_1, P_2^\alpha, \dots, P_k^\beta, P_{k+1}) \leq 2n^\alpha C_{k-1}(P_2^\alpha, P_3, \dots, P_{k+1}),$$

obtained in an analogous way. This gives

$$C_k(P_1, P_2^\alpha, P_3, \dots, P_{k+1}) = \tilde{O}\left(n^\alpha n^{\frac{f(k-1)+2x-3\alpha+y}{3}}\right) = \tilde{O}\left(n^{\frac{f(k)-1+2x+y}{3}}\right) = \tilde{O}\left(n^{\frac{f(k)+x+y}{3}}\right).$$

◀

▶ Remark 8. The proof above is not sufficient to obtain an almost sharp bound in the $k \equiv 1 \pmod 3$ case for two reasons. First, for these k any analogue of Theorem 7 would involve taking maximums of two expressions, where one contains $u_2(n^x, n)$ and the other contains $u_2(n^y, n)$. However, due to our lack of good understanding of how $u_2(n^x, n)$ changes as x is increasing, this is difficult to work with.

Second, on a more technical side, while Case 1 and Case 2 in the above proof would go through with any reasonable inductive statement, Case 3 would fail. The main reason for this is that C_k as a function of k makes jumps at every third value of k, and remains essentially the same, or changes by $u(n, n)/n$ for the other values of k. Thus one would need to remove three vertices from the path to make the induction work. However, the path has only two ends, and removing vertices other than the endpoints turns out to be intractable.

3.3 Upper bound for $k = 4$

In this section we prove the upper bound in Theorem 3 for $k = 4$. Let $P_1, \ldots, P_5$ be five sets of n points. We will show that $C_4(P_1, \ldots, P_5) = \tilde{O}(u_2(n)n)$, which is slightly stronger than what is stated in Theorem 3.

Instead of (3) we need the following more general bound on the number of rich points.

▶ **Observation 9** (Richness bound). *Let n^y be the maximum possible number of points that are n^α-rich with respect to a set of n^x points and some distance δ. Then we have*

$$n^{y+\alpha} \le u_2(n^x, n^y), \tag{9}$$

or, equivalently

$$n^\alpha \le \frac{u_2(n^x, n^y)}{n^y}.$$

The proof of (9) follows immediately from the definition of n^α richness and $u_2(n^x, n^y)$.

Let $\Lambda := \left\{\frac{i}{\log n} : i = 0, \ldots, \lceil \log n \rceil\right\}^4$. For any $\boldsymbol{\alpha} = (\alpha_2, \alpha_3, \alpha_4, \alpha_5) \in \Lambda$ let $Q_1^{\boldsymbol{\alpha}} = P_1$ and for $i = 2, \ldots, 5$ define recursively $Q_i^{\boldsymbol{\alpha}}$ to be the set of those points in P_i that are at least n^{α_i}-rich but at most $2n^{\alpha_i}$-rich with respect to Q_{i-1} and δ_i.

It is not difficult to see that

$$\mathcal{C}_4(P_1, \ldots, P_5) = \bigcup_{\boldsymbol{\alpha} \in \Lambda} \mathcal{C}_4\left(Q_1^{\boldsymbol{\alpha}}, \ldots, Q_5^{\boldsymbol{\alpha}}\right).$$

We have $|\Lambda| = \tilde{O}(1)$ and thus, in order to prove the theorem, it is sufficient to show that for every $\boldsymbol{\alpha} \in \Lambda$ we have

$$C_4\left(Q_1^{\boldsymbol{\alpha}}, \ldots, Q_5^{\boldsymbol{\alpha}}\right) = O\left(n \cdot u_2(n, n)\right).$$

From now on, fix $\boldsymbol{\alpha} = (\alpha_2, \ldots, \alpha_5)$, and denote $Q_i = Q_i^{\boldsymbol{\alpha}}$. Choose $x_i \in [0, 1]$ so that $|Q_i| = n^{x_i}$. Then we have

$$C_4(Q_1, \ldots, Q_5) = O\left(n^{x_5 + \alpha_5 + \alpha_4 + \alpha_3 + \alpha_2}\right). \tag{10}$$

Indeed, each chain $(p_1, \ldots, p_5)$ with $p_i \in Q_i$ can be obtained in the following five steps.

- **Step 1:** Pick $p_5 \in Q_5$.
- **Step i ($2 \le i \le 5$):** Pick a point $p_{6-i} \in Q_{6-i}$ at distance δ_{6-i} from p_{7-i}.

In the first step we have n^{x_5} choices, and for $i \geq 2$ in the i-th step we have at most $2n^{\alpha_{6-i}}$ choices. Further, by Observation 9, for each $i \geq 2$ we have

$$n^{\alpha_i} \leq \frac{u_2(n^{x_{i-1}}, n^{x_i})}{n^{x_i}}. \tag{11}$$

Combining (10) and (11), we obtain

$$C_4(Q_1, \ldots, Q_5) = O\left(u_2(n^{x_4}, n^{x_5})\frac{u_2(n^{x_3}, n^{x_4})}{n^{x_4}}\frac{u_2(n^{x_2}, n^{x_3})}{n^{x_3}}\frac{u_2(n^{x_1}, n^{x_2})}{n^{x_2}}\right). \tag{12}$$

By (2) we have

$$u_2(n^{x_{i-1}}, n^{x_i}) = O\left(\max\left\{n^{\frac{2}{3}(x_i+x_{i-1})}, n^{x_i}, n^{x_{i-1}}\right\}\right).$$

Note that the maximum is attained on the second (third) term iff $x_{i-1} \leq \frac{x_i}{2}$ ($x_i \leq \frac{x_{i-1}}{2}$). To bound $C_4(Q_1, \ldots, Q_5)$ we consider several cases depending on which of these three terms the maximum above is attained on for different i.

Case 1: For all $2 \leq i \leq 5$ we have $u_2(n^{x_{i-1}}, n^{x_i}) = O\left(n^{\frac{2}{3}(x_i+x_{i-1})}\right)$. Then

$$\frac{u_2(n^{x_4}, n^{x_5})u_2(n^{x_3}, n^{x_4})u_2(n^{x_2}, n^{x_3})}{n^{x_2+x_3+x_4}} = O\left(n^{\frac{2}{3}x_5+\frac{1}{3}x_4+\frac{1}{3}x_3-\frac{1}{3}x_2}\right)$$

and

$$\frac{u_2(n^{x_3}, n^{x_4})u_2(n^{x_2}, n^{x_3})u_2(n^{x_1}, n^{x_2})}{n^{x_2+x_3+x_4}} = O\left(n^{-\frac{1}{3}x_4+\frac{1}{3}x_3+\frac{1}{3}x_2+\frac{2}{3}x_1}\right).$$

Substituting each of these two displayed formulas into (12) and taking their product, we obtain

$$C_4(Q_1, \ldots, Q_5)^2 = O\left(u_2(n^{x_1}, n^{x_2})u_2(n^{x_4}, n^{x_5}) \cdot n^{\frac{2}{3}x_1+\frac{2}{3}x_3+\frac{2}{3}x_5}\right) = O\left(u_2(n, n)^2 \cdot n^2\right),$$

which concludes the proof in this case.

Case 2: There is an $2 \leq i \leq 5$ such that

$$\min\{x_{i-1}, x_i\} \leq \frac{1}{2}\max\{x_{i-1}, x_i\} \quad \text{and thus} \quad u_2(n^{x_{i-1}}, n^{x_i}) = O\left(\max\{n^{x_{i-1}}, n^{x_i}\}\right). \tag{13}$$

We distinguish three cases based on for which i holds.

Case 2.1: (13) holds for $i = 2$ or 5. In particular, this implies that $u_2(n^{x_1}, n^{x_2}) = O(n)$ or $u_2(n^{x_4}, n^{x_5}) = O(n)$. The following lemma finishes the proof in this case.

▶ **Lemma 10.** *Let $R_1, \ldots, R_5 \subseteq \mathbb{R}^2$ such that $|R_i| \leq n$ for every $i \in [5]$. If $u_2(R_1, R_2) = O(n)$ or $u_2(R_4, R_5) = O(n)$ holds, then $C_4(R_1, \ldots, R_5) = O(n \cdot u_2(n, n))$.*

Proof. We have

$$C_4(R_1, \ldots, R_5) \leq 2u_2(R_1, R_2)u_2(R_4, R_5) = O\left(n \cdot u_2(n, n)\right).$$

Indeed, every 4-tuple (r_1, r_2, r_4, r_5) with $r_i \in R_i$ can be extended in at most two different ways to a 4-chain $(r_1, \ldots, r_5) \in R_1 \times \cdots \times R_5$. At the same time, the number of 4-tuples with $\|r_1 - r_2\| = \delta_1$, $\|r_4 - r_5\| = \delta_4$ is at most $u_2(R_1, R_2)u_2(R_4, R_5)$. ◀

Case 2.2: (13) holds for $i = 4$. Note that if $x_4 \leq \frac{x_3}{2} \leq \frac{1}{2}$, then $u_2(n^{x_5}, n^{x_4}) = O(n)$, and we can apply Lemma 10 to conclude the proof in this case. Thus we may assume that $x_3 \leq \frac{x_4}{2}$, and hence $u_2(n^{x_4}, n^{x_3}) = O(n^{x_4})$. This means that $n^{\alpha_4} = O(1)$ by Observation 9. Thus to finish the proof of this case, it is sufficient to prove the following claim.

▶ **Claim 11.** *Let $R_1, \ldots, R_5 \subseteq \mathbb{R}^2$ such that $|R_i| \leq n$ for all $i \in [5]$ and every point of R_4 is $O(1)$ rich with respect to R_3 and δ_3. Then $C_4(R_1, \ldots, R_5) = O\left(n \cdot u_2(n, n)\right)$.*

Proof. Every 4-chain $(r_1, \ldots, r_5)$ can be obtained in the following steps.

- Pick a pair $(r_4, r_5) \in R_4 \times R_5$ with $\|r_4 - r_5\| = \delta_4$.
- Choose $r_3 \in R_3$ at distance δ_3 from r_4.
- Pick a point $r_1 \in R_1$.
- Extend (r_1, r_3, r_4, r_5) to a 4-chain.

In the first step, we have at most $u_2(n, n)$ choices, in the third at most n choices, and in the other two steps at most $O(1)$. ◀

Case 2.3: (13) holds for $i = 3$ *only.* Arguing as in Case 2.2, we may assume that $u_2(n^{x_3}, n^{x_2}) = O(n^{x_2})$. Then we have

$$C_4(Q_1, \ldots, Q_5) = O\left(u_2(n^{x_4}, n^{x_5}) \frac{u_2(n^{x_3}, n^{x_4})}{n^{x_4}} \frac{u_2(n^{x_2}, n^{x_3})}{n^{x_3}} \frac{u_2(n^{x_1}, n^{x_2})}{n^{x_2}}\right)$$

$$= O\left(u_2(n^{x_1}, n^{x_2}) \cdot n^{\frac{2}{3}(x_4+x_5)+\frac{2}{3}(x_3+x_4)-x_4-x_3}\right) = O\left(u_2(n, n) \cdot n\right),$$

which finishes the proof.

3.4 Upper bound for $k \equiv 1 \pmod 3$

We will prove the upper bound in Theorem 3 for $k \equiv 1$ by induction. The $k = 1$ case follows from the definition of $u_2(n, n)$, thus we may assume that $k \geq 4$. For the rest of the section fix $\varepsilon' > 0$, and sets $P_1, \ldots, P_{k+1} \subseteq \mathbb{R}^2$ of size n, further let $\varepsilon = \frac{\varepsilon'}{4k}$. We are going to show that $C_k(P_1, \ldots, P_{k+1}) = O(n^{(k-1)/3+\varepsilon'} u_2(n))$.

The first step of the proof is to find a certain covering of $P_1 \times \cdots \times P_{k+1}$, which resembles the one used for the $k = 4$ case, although is more elaborate.[2] (The goal of this covering is to make the corresponding graph between each of the two consecutive parts "regular in both directions" in a certain sense.)

Let

$$\Lambda = \left\{i\varepsilon : i = 0, \ldots, \left\lfloor \frac{1}{\varepsilon} \right\rfloor\right\}^{k+1}.$$

We cover the product $\mathbf{P} = P_1 \times \cdots \times P_{k+1}$ by fine-grained classes $P_1^{\boldsymbol{\gamma}} \times \ldots \times P_{k+1}^{\boldsymbol{\gamma}}$ encoded by the sequence $\boldsymbol{\gamma} = (\boldsymbol{\gamma^1}, \boldsymbol{\gamma^2}, \ldots)$ of length at most $(k+1)\varepsilon^{-1} + 1$ with $\boldsymbol{\gamma^j} \in \Lambda$ for each $j = 1, 2, \ldots$. One property that we shall have is

$$P_1 \times \cdots \times P_{k+1} = \bigcup_{\boldsymbol{\gamma}} P_1^{\boldsymbol{\gamma}} \times \ldots \times P_{k+1}^{\boldsymbol{\gamma}}.$$

To find the covering, first we define a function D that receives a parity digit $j \in \{0, 1\}$, a product set $\mathbf{R} := R_1 \times \ldots \times R_{k+1}$ and an $\boldsymbol{\alpha} \in \Lambda$, and outputs a product set $D(j, \boldsymbol{R}, \boldsymbol{\alpha}) = \mathbf{R}(\boldsymbol{\alpha}) = R_1(\boldsymbol{\alpha}) \times \ldots \times R_{k+1}(\boldsymbol{\alpha})$.

[2] This covering brings in the ε-error term in the exponent, that we could avoid in the $k = 4$ case.

Definition of D.

- If $j = 1$ then let $R_1(\boldsymbol{\alpha}) := R_1$ and for $i = 2, \dots, k+1$ define $R_i(\boldsymbol{\alpha})$ iteratively to be the set of points in R_i that are at least n^{α_i}, but at most $n^{\alpha_i+\varepsilon}$-rich with respect to $R_{i-1}(\boldsymbol{\alpha})$ and δ_{i-1}.
- If $j = 0$ then apply the same procedure, but in reverse order. That is, let $R_{k+1}(\boldsymbol{\alpha}) = R_{k+1}$ and for $i = k, k-1, \dots, 1$ define $R_i(\boldsymbol{\alpha})$ iteratively to be the set of points in R_i that are at least n^{α_i} but at most $n^{\alpha_i+\varepsilon}$-rich with respect to $R_{i+1}(\boldsymbol{\alpha})$ and δ_i.

Note that

$$\mathbf{R} = \bigcup_{\boldsymbol{\alpha}\in\Lambda} \mathbf{R}(\boldsymbol{\alpha}). \tag{14}$$

For a sequence $\boldsymbol{\gamma} = (\boldsymbol{\gamma^1}, \boldsymbol{\gamma^2}, \dots)$ with $\boldsymbol{\gamma^j} \in \Lambda$, we define $\mathbf{P}^{\boldsymbol{\gamma}}$ recursively as follows. Let $\mathbf{P}^{\emptyset} := \mathbf{P}$, and for each $j \geq 1$ let

$$\mathbf{P}^{(\boldsymbol{\gamma^1},\dots,\boldsymbol{\gamma^j})} = D(j \pmod 2, \mathbf{P}^{(\boldsymbol{\gamma^1},\dots,\boldsymbol{\gamma^{j-1}})}, \boldsymbol{\gamma^j}).$$

We say that a sequence $\boldsymbol{\gamma}$ is *stable at j* if

$$\left|\mathbf{P}^{(\boldsymbol{\gamma^1},\dots,\boldsymbol{\gamma^j})}\right| \geq \left|\mathbf{P}^{(\boldsymbol{\gamma^1},\dots,\boldsymbol{\gamma^{j-1}})}\right| \cdot n^{-\varepsilon}.$$

Otherwise $\boldsymbol{\gamma}$ is *unstable at j*.

▶ **Definition 12.** *Let Υ be the set of those sequences $\boldsymbol{\gamma}$ that are stable at their last coordinate, but are not stable for any previous coordinate, and for which $\boldsymbol{P}^{\boldsymbol{\gamma}}$ is non-empty.*

The set Υ has several useful properties, some of which are summarised in the following lemma.

▶ **Lemma 13.**

1. *Any $\boldsymbol{\gamma} \in \Upsilon$ has length at most $(k+1)\varepsilon^{-1} + 1$.*
2. $|\Upsilon| = O_\varepsilon(1)$.
3. $\boldsymbol{P} = \bigcup_{\boldsymbol{\gamma}\in\Upsilon} \boldsymbol{P}^{\boldsymbol{\gamma}}$.

Proof.

1. If $\boldsymbol{\gamma}$ is unstable at j then

$$|\mathbf{P}^{(\gamma^1,\dots,\gamma^j)}| \leq |\mathbf{P}^{(\gamma^1,\dots,\gamma^{j-1})}| \cdot n^{-\varepsilon}.$$

Since $|\mathbf{P}| = n^{k+1}$ and $|\mathbf{P}^{\boldsymbol{\gamma}}| \geq 1$, we conclude that $\boldsymbol{\gamma}$ is unstable at at most $(k+1)\varepsilon^{-1}$ indices j.

2. It follows from part 1 by counting all possible sequences of length at most $(k+1)\varepsilon^{-1} + 1$ of elements from the set Λ. (Note that $|\Lambda| = O_\varepsilon(1)$.)
3. For a nonnegative integer j let $\Lambda^{\leq j}$ be the set of all sequences of length at most j of elements from Λ. Let

$$\Upsilon_j := \left(\Upsilon \cap \Lambda^{\leq j}\right) \cup \Psi_j, \text{ where } \Psi_j := \left\{\boldsymbol{\gamma} \in \Lambda^j : \boldsymbol{\gamma} \text{ is not stable for any } \ell \leq j\right\}.$$

By part 1 of the lemma, $\Upsilon_j = \Upsilon$ for $j > (k+1)\varepsilon^{-1}$. We prove by induction on j that $\mathbf{P} = \bigcup_{\boldsymbol{\gamma}\in\Upsilon_j} \mathbf{P}^{\boldsymbol{\gamma}}$.

Υ_0 consists of an empty sequence, thus the statement is clear for $j = 0$. Next, assume that the statement holds for j. We have

$$\mathbf{P} = \bigcup_{\boldsymbol{\gamma}\in\Upsilon_j} \mathbf{P}^{\boldsymbol{\gamma}} = \bigcup_{\boldsymbol{\gamma}\in\Lambda^{\leq j}} \mathbf{P}^{\boldsymbol{\gamma}} \cup \bigcup_{\boldsymbol{\gamma}\in\Psi_j} \mathbf{P}^{\boldsymbol{\gamma}}.$$

By (14) we have that $\mathbf{P}^{\boldsymbol{\gamma}} = \bigcup_{\boldsymbol{\gamma}'} \mathbf{P}^{\boldsymbol{\gamma}'}$ holds for any $\boldsymbol{\gamma} \in \Psi_j$, where the union is taken over the sequences from Λ^{j+1} that coincide with $\boldsymbol{\gamma}$ on the first j entries. This, together with $\boldsymbol{\gamma}' \in \left(\Upsilon \cap \Lambda^{j+1}\right) \cup \Psi_{j+1}$ when $\mathbf{P}^{\boldsymbol{\gamma}'}$ is nonempty finishes the proof. ◀

Parts 2 and 3 of Lemma 13 imply that in order to complete the proof of the $k \equiv 1 \pmod 3$ case, it is sufficient to show that for any $\boldsymbol{\gamma} \in \Upsilon$ we have

$$C_k(P_1^{\boldsymbol{\gamma}}, \ldots, P_{k+1}^{\boldsymbol{\gamma}}) = O\left(u_2(n) \cdot n^{\frac{k-1}{3}+4k\varepsilon}\right). \tag{15}$$

From now on fix $\boldsymbol{\gamma} \in \Upsilon$. For each $i = 1, \ldots, k+1$ let $R_i := P_i^{\boldsymbol{\gamma}}$ and $Q_i := P_i^{\boldsymbol{\gamma}'}$, where $\boldsymbol{\gamma}'$ is obtained from $\boldsymbol{\gamma}$ by removing the last element of the sequence. Without loss of generality, assume that the length ℓ of $\boldsymbol{\gamma}$ is even. For each $i = 1, \ldots, k+1$, choose x_i, y_i such that

$$|Q_i| = n^{x_i}, \qquad |R_i| = n^{y_i}.$$

Let $\alpha_i := \gamma_i^{\ell-1}$ and $\beta_i := \gamma_i^{\ell}$. By the definition of $\mathbf{P}^{\boldsymbol{\gamma}}$ we have that each point in Q_i is at least n^{α_i}-rich but at most $n^{\alpha_i+\varepsilon}$-rich with respect to Q_{i-1} and δ_{i-1}, and each point in R_i is at least n^{β_i}-rich but at most $n^{\beta_i+\varepsilon}$-rich with respect to R_{i+1} and δ_i.

By Observation 9, we have

$$n^{\alpha_i} \le \frac{u_2(n^{x_{i-1}}, n^{x_i})}{n^{x_i}} \quad \text{ and } \quad n^{\beta_i} \le \frac{u_2(n^{y_i}, n^{y_{i+1}})}{n^{y_i}} \le \frac{u_2(n^{x_i}, n^{x_{i+1}})}{n^{x_i-\varepsilon}}. \tag{16}$$

The last inequality follows from two facts: first $u_2(n^{y_i}, n^{y_{i+1}}) \le u_2(n^{x_i}, n^{x_{i+1}})$ and, second, since $\boldsymbol{\gamma}$ is stable at its last coordinate[3], we have $n^{y_i} = |R_i| \ge |Q_i| \cdot n^{-\varepsilon} = n^{x_i-\varepsilon}$.

In the same fashion as in the beginning of Section 3.3, we can show that

$$C_k(R_1, \ldots, R_{k+1}) \le n^{y_1} n^{\beta_1+\cdots+\beta_k+k\varepsilon}, \text{ and}$$

$$C_k(R_1, \ldots, R_{k+1}) \le C_k(Q_1, \ldots, Q_{k+1}) \le n^{x_{k+1}} n^{\alpha_{k+1}+\alpha_k+\cdots+\alpha_2+k\varepsilon}.$$

Combining the first of these displayed inequalities with (16), we have

$$C_k(R_1, \ldots, R_{k+1}) \le u_2(n^{x_1}, n^{x_2}) \prod_{2 \le i \le k} \frac{u_2\left(n^{x_i}, n^{x_{i+1}}\right)}{n^{x_i}} n^{2k\varepsilon}.$$

Recall that

$$u_2(n^{x_i}, n^{x_{i+1}}) = O\left(\max\{n^{\frac{2}{3}(x_i+x_{i+1})}, n^{x_i}, n^{x_{i+1}}\}\right). \tag{17}$$

To bound $C_k(R_1, \ldots, R_{k+1})$, we consider several cases based on which of these three terms can be used to bound $u_2(n^{x_i}, n^{x_{i+1}})$ for different values of i.

Case 1: Either $u_2(n^{x_1}, n^{x_2}) = O(n)$ or $u_2(n^{x_k}, n^{x_{k+1}}) = O(n)$ holds. As in the proof of Lemma 10, we have

$$\begin{aligned} & C_k(R_1, \ldots, R_{k+1}) \\ & \qquad \le \min\left\{2u_2(n^{y_1}, n^{y_2}) C_{k-3}(R_4, \ldots, R_{k+1}), 2u_2(n^{y_k}, n^{y_{k+1}}) C_{k-3}(R_1, \ldots, R_{k-2})\right\}. \end{aligned}$$

By induction we obtain $C_{k-3}(R_4, \ldots, R_{k+1}), C_{k-3}(R_1, \ldots, R_{k-2}) = O\left(n^{\frac{k-4}{3}+\varepsilon} \cdot u_2(n)\right)$. Together with the assumption of Case 1, and the fact that $u_2(n^{y_1}, n^{y_2}) \le u_2(n^{x_1}, n^{x_2})$ and $u_2(n^{y_k}, n^{y_{k+1}}) \le u_2(n^{x_k}, n^{x_{k+1}})$, this implies (15) and finishes the proof.

[3] This is essentially the only place where we use the stability of $\boldsymbol{\gamma}$.

Case 2: For some $i = 1, \ldots, (k-1)/3$, one of the following holds:
- $u_2(n^{x_{3i+1}}, n^{x_{3i+2}}) = O(\max\{n^{x_{3i+1}}, n^{x_{3i+2}}\})$;
- $u_2(n^{x_{3i-1}}, n^{x_{3i}}) = O(n^{x_{3i-1}})$;
- $u_2(n^{x_{3i}}, n^{x_{3i+1}}) = O(n^{x_{3i+1}})$.

We will show how to conclude in the first case. The other cases are very similar and we omit the details of their proofs. If $u_2(n^{x_{3i+1}}, n^{x_{3i+2}}) = O(n^{x_{3i+2}})$ then $n^{\alpha_{3i+2}} = O(1)$ by (16). Every chain $(r_1, \ldots, r_{k+1}) \in \mathcal{C}_k(Q_1, \ldots, Q_{k+1})$ can be obtained as follows.

1. Pick a $(3i-2)$-chain $(r_1, \ldots, r_{3i-1})$ with $r_j \in Q_j$ for every j.
2. Pick a $(k-3i-1)$-chain $(r_{3i+2}, r_{3i+3}, \ldots, r_{k+1})$ with $r_j \in Q_j$ for every j.
3. Extend $(r_{3i+2}, r_{3i+3}, \ldots, r_{k+1})$ to a $(k-3i-2)$ chain $(r_{3i+1}, r_{3i+2}, \ldots, r_{k+1})$.
4. Connect $(r_1, \ldots, r_{3i-1})$ and $(r_{3i+1}, r_{3i+2}, \ldots, r_{k+1})$ to obtain a k-chain.

In the first step, we have $O\left(n^{\frac{3i-3}{3}+\varepsilon} \cdot u_2(n)\right)$ choices by induction on k. In the second step, we have $\tilde{O}\left(n^{\frac{k-3i+2}{3}}\right)$ choices by the $k \equiv 0 \pmod 3$ case of Theorem 3. In the third step, we have at most $n^{\alpha_{3i+2}+\varepsilon} = O(n^{\varepsilon})$ choices. Finally, in the fourth step we have at most 2 choices. Thus the number of k-chains is at most

$$O\left(n^{\frac{3i-3}{3}+\varepsilon} \cdot u_2(n)\right) \cdot \tilde{O}\left(n^{\frac{k-3i+2}{3}}\right) \cdot O\left(n^{\varepsilon}\right) \cdot 2 = O\left(n^{\frac{k-1}{3}+3\varepsilon} \cdot u_2(n)\right),$$

finishing the proof of the first case.

If $u_2(n^{x_{3i+1}}, n^{x_{3i+2}}) = O(n^{x_{3i+1}})$ then $n^{\beta_{3i+1}} = O(n^{\varepsilon})$ by (16).[4] We proceed similarly in this case, but we count the k-chains now in $R_1 \times \ldots \times R_{k+1}$ instead in $Q_1 \times \ldots \times Q_{k+1}$ (and get an extra factor of n^{ε} in the bound). In all cases, we obtain (15).

Case 3: Neither the assumptions of Case 1 nor that of Case 2 hold. We define four sets S', S'_+, S'_{++}, and S'_- of indices in $\{2, \ldots, k\}$ as follows. Let

$$S' := \left\{i : u_2(n^{x_i}, n^{x_{i-1}}) = O(n^{\frac{2}{3}(x_i + x_{i-1})}) \text{ and } u_2(n^{x_{i+1}}, n^{x_i}) = O(n^{\frac{2}{3}(x_{i+1}+x_i)})\right\},$$

$$S'_+ := \Big\{i : u_2(n^{x_i}, n^{x_{i-1}}) = O(n^{\frac{2}{3}(x_i + x_{i-1})}) \text{ and } u_2(n^{x_{i+1}}, n^{x_i}) = O(n^{x_i}), \text{ or}$$
$$u_2(n^{x_i}, n^{x_{i-1}}) = O(n^{x_i}) \text{ and } u_2(n^{x_{i+1}}, n^{x_i}) = O(n^{\frac{2}{3}(x_{i+1}+x_i)})\Big\},$$

$$S'_{++} := \left\{i : u_2(n^{x_i}, n^{x_{i-1}}) = O(n^{x_i}) \text{ and } u_2(n^{x_{i+1}}, n^{x_i}) = O(n^{x_i})\right\}, \text{ and}$$

$$S'_- := \Big\{i : u_2(n^{x_i}, n^{x_{i-1}}) = O(n^{\frac{2}{3}(x_i + x_{i-1})}) \text{ and } u_2(n^{x_{i+1}}, n^{x_i}) = O(n^{x_{i+1}}), \text{ or}$$
$$u_2(n^{x_i}, n^{x_{i-1}}) = O(n^{x_{i-1}}) \text{ and } u_2(n^{x_{i+1}}, n^{x_i}) = O(n^{\frac{2}{3}(x_{i+1}+x_i)})\Big\}.$$

Since the conditions of Case 2 are not satisfied, we have

$$\{2, \ldots, k\} \subseteq S' \cup S'_+ \cup S'_{++} \cup S'_-.$$

Indeed, for each $i \in \{2, \ldots, k\}$, there are 9 possible pairs of maxima in (17) with $i, i+1$. The four sets above encompass 6 possibilities. In total, there are 4 possible pairs of maxima with

[4] This is the key application of (16), and the reason why we needed a decomposition with regularity in both directions between the consecutive parts.

only the two last terms from (17) used. For $i \equiv 1,2 \pmod 3$, any of those 4 are excluded due to the first condition in Case 2 (in fact, then $i \in S' \cup S'_-$). If $i \equiv 0 \pmod 3$, then the second and the third condition in Case 2 rule out all possibilities but the one defining S'_{++}.

From these, it is also easy to see that if $i \in S'_{++}$, then $i-1, i+1 \in S'_-$, while if $i \in S'_+$ then one of $i-1, i+1$ is in S'_-. (Recall that $i \in S'_+ \cup S'_{++}$ only if $i \equiv 0 \pmod 3$.) These together imply

$$|S'_+| + 2|S'_{++}| \le |S'_-|. \tag{18}$$

We partition $\{2,\dots,k\}$ using these sets as follows: let $S_- = S'_-, S = S' \setminus S'_-, S_+ = S'_+ \setminus (S'_- \cup S')$ and $S_{++} = \{2,\dots,k\} \setminus S'_- \cup S' \cup S'_+$. Note that the analogue of (18) holds for the new sets. That is, we have

$$|S_+| + 2|S_{++}| \le |S_-|.$$

Recall that

$$C_k(R_1,\dots,R_{k+1}) \le u_2(n^{x_1}, n^{x_2}) \prod_{2\le i\le k} \frac{u_2\left(n^{x_i}, n^{x_{i+1}}\right)}{n^{x_i}} n^{2k\varepsilon}. \tag{19}$$

Since the assumptions of Case 1 and 2 do not hold, we have $2,k \in S$. Indeed, $2,k \ne 0 \pmod 3$ and thus $2,k \notin S_+, S_{++}$. Further, if say $k \in S_- = S'_-$ then by the definition of S'_- we either have $u_2(n^{x_{k+1}}, n^{x_k}) = O(n)$, or $u_2(n^{x_k}, n^{x_{k-1}}) = O(n^{x_{k-1}})$. The first case cannot hold since the assumption of Case 1 does not hold. Further, the second case cannot hold either, since it would imply $x_k \le \frac{x_{k-1}}{2} \le \frac{1}{2}$, meaning $u_2(n^{x_{k+1}}, n^{x_k}) = O(n)$. Using $2,k \in S$ and expanding (19), we obtain

$$C_k(R_1,\dots,R_{k+1}) \le n^{2k\varepsilon} u_2(n^{x_1}, n^{x_2}) n^{-\frac{1}{3}x_2} n^{\frac{2}{3}x_{k+1}} \prod_{\substack{i\in S,\\ i\ne 2}} n^{\frac{1}{3}x_i} \prod_{i\in S_+} n^{\frac{2}{3}x_i} \prod_{i\in S_{++}} n^{x_i} \prod_{i\in S_-} n^{-\frac{1}{3}x_i}, \tag{20}$$

and

$$C_k(R_1,\dots,R_{k+1}) \le n^{2k\varepsilon} u_2(n^{x_k}, n^{x_{k+1}}) n^{-\frac{1}{3}x_k} n^{\frac{2}{3}x_1} \prod_{\substack{i\in S,\\ i\ne k}} n^{\frac{1}{3}x_i} \prod_{i\in S_+} n^{\frac{2}{3}x_i} \prod_{i\in S_{++}} n^{x_i} \prod_{i\in S_-} n^{-\frac{1}{3}x_i}. \tag{21}$$

Taking the product of (20) and (21) we obtain

$$C_k(R_1,\dots,R_{k+1})^2 \le$$

$$n^{4k\varepsilon} \cdot u_2(n^{x_1}, n^{x_2}) u_2(n^{x_k}, n^{x_{k+1}}) n^{\frac{2}{3}(x_1 + x_{k+1})} \left(\prod_{\substack{i\in S,\\ i\ne 2,k}} n^{\frac{1}{3}x_i} \prod_{i\in S_+} n^{\frac{2}{3}x_i} \prod_{i\in S_{++}} n^{x_i} \prod_{i\in S_-} n^{-\frac{1}{3}x_i} \right)^2$$

$$\le n^{4k\varepsilon} \cdot u_2(n,n)^2 \cdot n^{2\left(\frac{2}{3} + \frac{1}{3}|S\setminus\{2,k\}| + \frac{2}{3}|S_+| + |S_{++}|\right)} = u_2(n,n)^2 \cdot n^{\frac{2(k-1)}{3} + 4k\varepsilon}.$$

The last equality follows from $|S_+| + 2|S_{++}| \le |S_-|$, which is equivalent to $\frac{2}{3}|S_+| + |S_{++}| \le \frac{1}{3}(|S_+| + |S_{++}| + |S_-|)$, and from the fact that S, S_+,S_{++}, and S_- partition $\{2,\dots,k\}$. This finishes the proof.

4 Bounds in $\mathbb{R}^3$

Similarly as in the planar case, for $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_k)$ and $P_1 \ldots, P_{k+1} \subseteq \mathbb{R}^3$ we denote by $\mathcal{C}_k^{3,\boldsymbol{\delta}}(P_1, \ldots, P_k)$ the family of $(k+1)$-tuples $(p_1, \ldots, p_{k+1})$ with $p_i \in P_i$ for all $i \in [k+1]$ and with $\|p_i - p_{i+1}\| = \delta_i$ for all $i \in [k]$. Let $C_k^{3,\boldsymbol{\delta}}(P_1, \ldots, P_{k+1}) = |\mathcal{C}_k^{3,\boldsymbol{\delta}}(P_1, \ldots, P_{k+1})|$ and

$$C_k^3(n_1, \ldots, n_{k+1}) = \max C_k^{3,\boldsymbol{\delta}}(P_1, \ldots, P_{k+1}),$$

where the maximum is taken over all choices of $\boldsymbol{\delta}$ and sets $P_1, \ldots, P_{k+1}$ subject to $|P_i| \le n_i$ for all $i \in [k+1]$.

It is easy to see that $C_k^3(n) \le C_k^3(n, \ldots, n) \le C_k^3((k+1)n)$. Since we are only interested in the order of magnitude of $C_k^3(n)$ for fixed k, sometimes we are going to work with $C_k^3(n, \ldots, n)$ instead of $C_k^3(n)$.

4.1 Lower bounds

For completeness, we recall the constructions from [8] for even $k \ge 2$. For every even $2 \le i \le k$, let $P_i = \{p_i\}$ be a single point such that the unit spheres centred at p_i and p_{i+2} intersect in a circle. Further, let P_1 and P_{k+1} be a set of n points contained in the unit sphere centred at p_2 and p_k respectively. Finally, for every odd $3 \le i \le k-1$, let P_i be a set of n points contained in the intersection of the unit spheres centred at p_{i-1} and p_{i+1}. Then it is not hard to see that $P_1 \times \cdots \times P_{k+1}$ contains $n^{\frac{k}{2}+1}$ many $(k, \boldsymbol{\delta})$-chains for $\boldsymbol{\delta} = (1, \ldots, 1)$.

Next, we prove the lower bounds for odd $k \ge 3$ given in Proposition 6.

Proof of Proposition 6. First we show that $C_k^3(n) = \Omega\left(\frac{u_3(n)^k}{n^{k-1}}\right)$. Take a set $P' \subset \mathbb{R}^3$ of size n that contains $u_3(n)$ point pairs at unit distance apart. It is a standard exercise in graph theory to show that there is $P \subset P'$ such that $\frac{n}{2} \le |P| \le n$ and for every $p \in P$ there are at least $\frac{u_3(n)}{4n}$ points $p' \in P$ at distance 1 from p. Then P contains $\Omega\left(\frac{u_3(n)^k}{n^{k-1}}\right)$ many $(k, \boldsymbol{\delta})$-chains with $\boldsymbol{\delta} = (1, \ldots, 1)$.

To prove $C_k^3(n) = \Omega\left(us_3(n)n^{k-2}\right)$, we modify and extend the construction used for $k-1$ as follows. Let $P_1, \ldots, P_{k-1}$ be as in the construction for $(k-1)$-chains (from the even case). Further, let P_k be a set of n points on the unit sphere around p_{k-1}, and P_{k+1} be a set of n points such that $u_3(P_k, P_{k+1}) = us_3(n)$. It is not hard to see that $P_1 \times \cdots \times P_{k+1}$ contains $\Omega\left(us_3(n)n^{k-2}\right)$ many $(k, \boldsymbol{\delta})$-chains with $\boldsymbol{\delta} = (1, \ldots, 1)$. ◀

4.2 Upper bound

We again fix $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_k)$ throughout the section and, omit it from the notation. The following result with $x = 1$ implies the upper bound in Theorem 5.

▶ **Theorem 14.** *For any fixed integer $k \ge 0$ and $x \in [0, 1]$, we have*

$$C_k^3(n^x, n, \ldots, n) = \tilde{O}\left(n^{\frac{k+1+x}{2}}\right).$$

Proof. The proof is by induction on k. For $k = 0$ the bound is trivial, and for $k = 1$ it follows from (4).

For $k \ge 2$ let $P_1, \ldots, P_{k+1} \subseteq \mathbb{R}^3$ be sets of points satisfying $|P_1| = n^x$, and $|P_i| = n$ for $2 \le n \le k+1$. Denote by $P_2^\alpha \subseteq P_2$ the set of those points in P_2 that are at least n^α-rich but at most $2n^\alpha$-rich with respect to P_1 and δ_1.

It is not hard to see that

$$\mathcal{C}_k^3(P_1, P_2 \dots, P_{k+1}) \subseteq \bigcup_{\alpha\in\Lambda} \mathcal{C}_k^3(P_1, P_2^\alpha, P_3, \dots, P_{k+1}),$$

where $\Lambda := \{\frac{i}{\log n} : i = 0, 1, \dots, \lfloor \log n \rfloor\}$. Since $|\Lambda| = \tilde{O}(1)$, it is sufficient to prove that, for every $\alpha \in \Lambda$, we have

$$C_k^3(P_1, P_2^\alpha, P_3, \dots, P_{k+1}) = \tilde{O}\left(n^{\frac{k+1+x}{2}}\right).$$

Assume that $|P_2^\alpha| = n^y$. The number of $(k-1)$-chains in $P_2^\alpha \times P_3 \times \dots \times P_{k+1}$ is at most $C_{k-1}^3(n^y, n, \dots, n)$, and each of them may be extended in $2n^\alpha$ ways. By induction, we get

$$C_k^3(P_1, P_2^\alpha, P_3, \dots, P_{k+1}) = \tilde{O}\left(n^\alpha \cdot n^{\frac{k+y}{2}}\right),$$

and we are done as long as

$$2\alpha + k + y \leq k + 1 + x. \tag{22}$$

To show this, we need to consider several cases depending on the value of α. Note that $\alpha \leq x$.

- If $\alpha \geq \frac{2x}{3}$, then by (5) we have $y \leq x-\alpha$, and the LHS of (22) is at most $\alpha+k+x \leq 1+k+x$.
- If $\frac{x}{2} \leq \alpha \leq \frac{2x}{3}$ then by (5) we have $y \leq 3x-4\alpha$. The LHS of (22) is at most $k+3x-2\alpha \leq k+2x \leq k+1+x$.
- If $\alpha \leq \frac{x}{2}$ then we use a trivial bound $y \leq 1$. The LHS of (22) is at most $2\alpha + k + 1 \leq x + k + 1$. ◀

References

1. P. K. Agarwal, E. Nevo, J. Pach, R. Pinchasi, M. Sharir, and S. Smorodinsky. Lenses in arrangements of pseudo-circles and their applications. *J. ACM*, 51(2):139–186, 2004.
2. B. Aronov and M. Sharir. Cutting circles into pseudo-segments and improved bounds for incidences. *Discrete Comput. Geom.*, 28(4):475–490, 2002.
3. P. Erdős. On sets of distances of n points. *Amer. Math. Monthly*, 53(5):248–250, 1946.
4. P. Erdős. On sets of distances of n points in euclidean space. *Magyar Tud. Akad. Mat. Kutato Int. Közl.*, 5:165–169, 1960.
5. H. Kaplan, J. Matoušek, Z. Safernová, and M. Sharir. Unit distances in three dimensions. *Comb. Probab. Comput.*, 21(4):597–610, 2012.
6. A. Marcus and G. Tardos. Intersection reverse sequences and geometric applications. *J. Combin. Theory Ser. A*, 113(4):675–691, 2006.
7. J. Pach and M. Sharir. Geometric incidences. In *Towards a theory of geometric graphs*, volume 342 of *Contemp. Math.*, pages 185–223. Amer. Math. Soc., Providence, RI, 2004.
8. E. A. Palsson, S. Senger, and A. Sheffer. On the number of discrete chains, 2019. `arXiv:1902.08259`.
9. J. Spencer, E. Szemerédi, and W. T Trotter. Unit distances in the Euclidean plane. In *Graph theory and combinatorics*, pages 294–304. Academic Press, 1984.
10. J. Zahl. An improved bound on the number of point-surface incidences in three dimensions. *Contrib. Discrete Math.*, 8(1):100–121, 2013.
11. J. Zahl. Breaking the 3/2 barrier for unit distances in three dimensions. *Int. Math. Res. Notices*, 2019(20):6235–6284, 2019.