Acta Math. Univ. Comenianae
Vol. LXXXVIII, 3 (2019), pp. 689–693

NEARLY k-DISTANCE SETS

N. FRANKL AND A. KUPAVSKII

ABSTRACT. We say that $S \subset \mathbb{R}^d$ is an *ε-nearly k-distance set* if there exist $1 \le t_1 \le \cdots \le t_k$ such that the distance between any two distinct points of S falls into $[t_1, t_1 + \varepsilon] \cup \cdots \cup [t_k, t_k + \varepsilon]$. In this abstract, we propose to study the quantity $M_k(d) := \lim_{\varepsilon\to 0} \max\{|S| : S$ is an ε-nearly k-distance set in $\mathbb{R}^d\}$. Let $m_k(d)$ be the maximal cardinality of a k-distance set in $\mathbb{R}^d$. We show that $M_k(d) = m_k(d)$ if either $d \ge d(k)$ or $k \le 3$.

We also address a closely related Turán-type problem, studied by Erdős, Makai, Pach, and Spencer in the 80's: given n points in $\mathbb{R}^d$, how many pairs out of them form a distance that belongs to $[t_1, t_1 + 1] \cup \cdots \cup [t_k, t_k + 1]$, where $t_1, \ldots, t_k$ are fixed and any two points in the set are at distance at least 1 apart? We obtain an exact answer for the same k, d as above.

1. INTRODUCTION

We call any point set that determines at most k distances a *k-distance set*. Let us denote by $m_k(d)$ the cardinality of the largest k-distance set in $\mathbb{R}^d$. Determining the value of $m_k(d)$ is a well studied hard question, which is in general wide open. The best known bounds are

$$\binom{d+1}{k} \le m_k(d) \le \binom{d+k}{k}. \tag{1}$$

Here, the lower bound follows from a simple construction in $\{0,1\}^{d+1}$, and the upper bound is due to Bannai, Bannai and Stanton [**1**].

In this abstract, we consider an "approximate" version of this problem. A set of points $S \subseteq \mathbb{R}^d$ is called an *ε-nearly k-distance set* if there exist $1 \le t_1 \le \cdots \le t_k$ such that

$$\|p - q\| \in [t_1, t_1 + \varepsilon] \cup \cdots \cup [t_k, t_k + \varepsilon]$$

for all $p \ne q \in S$.

We put $M_k(d) := \lim_{\varepsilon\to 0} \max\{|S| : S$ is an ε-nearly k-distance set in $\mathbb{R}^d\}$. The quantity $M_k(d)$ was generally overlooked in the literature. The only reference we

Received June 7, 2019.
2010 *Mathematics Subject Classification.* Primary 52C10.
N. Frankl was partially supported by the National Research, Development, and Innovation Office, NKFIH Grant K119670.
A. Kupavskii was 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. -6760.2018.1).

found is a brief mention by Erdős, Makai and Pach in a recent preprint [**7**, page 19], where they also write: "For k fixed, d sufficiently large probably $M_k(d) = m_k(d)$." We confirm this.

Theorem 1.1. *Fix a positive integer k. $M_k(d) = m_k(d)$ holds if either $d \geq d(k)$ or $k \leq 3$.*

We also study the following related problem. We call S *separated* if the distance between any two of its points is at least 1. Let $M_k(d,n)$ denote the maximum M, such that there exist numbers $1 \leq t_1 \leq \cdots \leq t_k$ and a separated set S of n points in $\mathbb{R}^d$ with at least M pairs of points at distance that falls into $[t_1, t_1+1] \cup \cdots \cup [t_k, t_k+1]$.

This quantity was studied by Erdős, Makai, Pach and Spencer [**6, 7, 8, 10**]. In [**8**], they showed that $M_1(d,n) = T(d,n)$ holds for sufficiently large n, where $T(d,n)$ is the number of edges in a balanced complete d-partite graph on n vertices. In [**7**], they proved that

$$M_2(d,n) = \frac{n^2}{2}\left(1 - \frac{1}{m_2(d)} + o(1)\right),$$

moreover, for $d \neq 4, 5$ they determined the exact value of $M_2(d,n)$ and showed that the same remains true if in the definition of $M_2(d,n)$ we change the intervals to be of the form $[t_i, t_i + cn^{1/d}]$ for some constant $c > 0$.

We strengthen and extend their results.

Theorem 1.2. *Fix k. If either $k \leq 3$ or $d \geq d(k)$, then for sufficiently large n we have*

$$M_k(d,n) = T(m_k(d-1), n).$$

Moreover, the same holds for intervals of the form $[t_i, t_i + cn^{1/d}]$ with some $c = c(k,d) > 0$.

Concerning the "$\geq$" part of the displayed inequality above, one can see that a stronger bound $M_k(d,n) \geq T(M_k(d-1), n) \geq T(m_k(d-1), n)$ is true for any $k \geq 1$, $d \geq 2$. This is shown by the following construction, which is similar to those that appeared in the works of Erdős, Makai, Pach and Spencer. Embed a $\frac{1}{2}$-nearly k-distance set $S \subseteq \mathbb{R}^{d-1}$ of size $M_k(d-1)$ and with distances $2n^2 \leq t_1 \leq \cdots \leq t_k$ in a hyperplane γ in $\mathbb{R}^d$. Replace each point $p \in S$ by an arithmetic progression A_p of length $\lfloor n/M_k(d-1) \rfloor$ or $\lceil n/M_k(d-1) \rceil$ and of difference 1, in the direction orthogonal to γ. One can easily check that in $\bigcup_{p \in S} A_p$ there are at least $T(M_k(d-1), n)$ pairs forming a distance in $[t_1, t_1+1] \cup \cdots \cup [t_k, t_k+1]$.

2. Comparing $m_k(d)$ with $M_k(d)$, and constructions

In this section we explore differences between k-distance and ε-nearly k-distance sets. Note that $M_k(d) \geq m_k(d)$ for all $k, d \geq 1$.

Proposition. $M_k(d) \geq \max\left\{\prod_{i=1}^s m_{k_i}(d_i) \,:\, \sum_{i=1}^s k_i = k, \sum_{i=1}^s d_i = d\right\}$ *holds for every $k, d \geq 1$.*

Proof. Take non-negative integers k_i, d_i, such that $\sum_{i=1}^{s} k_i = k$ and $\sum_{i=1}^{s} d_i = d$. Then there is an ε-nearly k-distance set in $\mathbb{R}^d$ of cardinality $\prod_{i=1}^{s} m_{k_i}(d_i)$ given by the following construction. For each i let S_i be a k_i-distance set in $\mathbb{R}^{d_i}$ of cardinality $m_{k_i}(d_i)$ and such that the distances in S_i are much larger (in terms of ε) than the distances in S_{i-1}. Then $S_1 \times \cdots \times S_s$ is an ε-nearly equal k-distance set in $\mathbb{R}^d$ of cardinality $\prod_{i=1}^{s} m_{k_i}(d_i)$. □

We may use this bound to show that $m_k(d) = M_k(d)$ is not always the case. Indeed, for example, in $\mathbb{R}^2$ the cardinality of the largest 6-distance set is 13, while the product of two arithmetic progressions of length 4 gives a ε-nearly 6-distance sets of cardinality 16. We however believe that $M_k(d) = M'_k(d)$ is true for all but very few pairs (k, d).

The difference between $m_k(d)$ and $M_k(d)$ for fixed d and growing k is very significant. In $\mathbb{R}^d$ the cardinality of a k-distance set is $O\left(k^{\frac{d}{2}+1}\right)$, by combining the result of Solymosi and Vu [**11**] with the result of Guth and Katz [**9**]. However, the product of d arithmetic progressions of size $\lfloor k/d \rfloor + 1$ gives a ε-nearly k-distance set of cardinality $(\lfloor k/d \rfloor + 1)^d \geq (k/d)^d$. While determining the order of magnitude of $m_k(d)$ for fixed d is a difficult open problem, we could find the order of magnitude of $M_k(d)$ for large k by a simple inductive proof.

Theorem 2.1. *$M_k(d) = \Theta\left(k^d\right)$ holds for any fixed $d \geq 2$.*

3. Proof outlines

Proof of Theorem 1.1 for fixed k, large d. First note that if we required that for a fixed constant K $\frac{t_i}{t_{i-1}} \leq K$ holds for any $1 < i \leq k$, then a standard compactness argument would imply $M_k(d) = m_k(d)$. Therefore we may assume that $\max_{1<i\leq k} \frac{t_i}{t_{i-1}} > K$ for some sufficiently large K.

For a sufficiently small ε consider an ε-nearly k-distance set $S \subseteq \mathbb{R}^d$ with distances $1 \leq t_1 \leq \cdots \leq t_k$, and assume that $1 < i \leq k$ is the largest index for which $\frac{t_i}{t_{i-1}} > K$. Colour a pair $\{p, q\}$ $(p, q \in S)$ with blue if $\|p - q\| \geq t_i$ and with red otherwise. Let B be the largest blue clique. Then by the triangle inequality S can be partitioned into $|B|$ vertex-disjoint red cliques $R_1, \ldots, R_{|B|}$ satisfying the following properties.

1. Each R_i shares exactly one vertex with B.
2. If $p \in R_i$, $q \in R_j$, $i \neq j$, then $\{p, q\}$ is blue.

We obtain $|B| \leq m_{k-i+1}(d)$ by using a compactness argument. We wish to bound each $|R_b|$ by using induction on k. Let $c > 0$ be a sufficiently small constant. We separate two cases.

Case 1: There is an $k \geq l > i$ such that $\frac{t_l}{t_{l-1}} \leq 1 + c$. Then by a compactness argument we obtain $|B| \leq m_{k-i}(d)$. Moreover, by induction on k we have $|R_b| \leq m_{i-1}$ for each $b \in \{1, \ldots, |B|\}$. Thus using (1) we conclude that

$$|S| = |R_1| + \cdots + |R_{|B|}| \leq m_{i-1}(d)m_{k-i}(d) \leq m_k(d)$$

holds for large d.

Case 2: $\min_{k\geq l>i} \frac{t_l}{t_{l-1}} > c$. Then, if K is sufficiently large,[1] by the triangle inequality we obtain the following. For each red clique R and each $b \in B \setminus R$ there is a $q \geq i$ such that $\|b - r\| \in [t_q, t_q + \varepsilon]$ for every $r \in R$. Using this, we obtain that if B is "essentially contained" in a subspace of dimension j, then each R_b is "essentially contained" in a subspace of dimension $d - j$. Hence by inducting on k we conclude that

$$|S| = |R_1| + \cdots + |R_{|B|}| \leq m_i(j)m_{k-i}(d-j) \leq m_k(d).$$

□

Proof of Theorem 1.2. Observe that $M_k(d,n) \leq T(M_k(d),n)$ is obvious for any $d, k \geq 1$ from Turan's theorem and the definition of $M_k(d)$. The difficulty in proving Theorem 1.2 lies in relating $M_k(d,n)$ with k-distance sets in one dimension smaller.

In our proof we combine the methods similar to those of of Erdős, Makai, Pach and Spencer with new ideas that use extremal graph theory and probabilistic arguments.

We use the following result of Erdős.

Theorem 3.1 ([**5**]). *Every n-vertex graph with at least $T(n, \ell - 1) + 1$ edges contains an edge that is contained in $\delta n^{\ell-2}$ cliques of size ℓ, where δ is a constant that depends on ℓ.*

The proof goes indirectly as follows. Assume that there is a separated set $S \subseteq \mathbb{R}^d$ and distances $t_1 \leq \cdots \leq t_k$ such that there are more than $T(n, m_k(d-1))$ pairs in S that span a distance in $[t_1, t_1 + 1] \cup \cdots \cup [t_k, t_k + 1]$. Then conclude that there are ε-nearly k-distance sets in $\mathbb{R}^{d-1}$ of cardinality $m_k(d-1) + 1$. This, together with Theorem 1.1, leads to a contradiction.

The main conceptual difference with the proofs of Erdős, Makai, Pach and Spencer is that, while they reduced the problem to the existence of k-distance sets of cardinality $m_k(d-1) + 1$ directly, we passed by the "intermediate" step and showed that $M_k(d-1) = m_k(d-1)$. This helped to isolate the difficulties of both parts of the argument (that have a very different nature) and make the approach more transparent and thus more powerful.

The main technical difficulty is that, rather than working with ε-nearly k-distance sets, we had to work with *almost flat ε-nearly k-distance sets S*: for all but at most two points $p \in S$, all vectors $p - q$, $q \in S$, form a very small angle with a fixed hyperplane. We thus needed to show that, for the k and d in question, the maximal cardinality of almost flat nearly k-distance sets in $\mathbb{R}^d$ is the same as that of k-distance sets. Working with almost flat ε-nearly k-distance sets involves a lot of additional geometric considerations. □

Acknowledgment. We thank Konrad Swanepoel for introducing us to the topic and for many fruitful discussions.

[1]The choice of K varies for different values of i.

References

1. Bannai E., Bannai E. and Stanton D., *An upper bound for the cardinality of an s-distance subset in real Euclidean space, II*, Combinatorica **3** (1983), 147–152.

2. Erdős P., *On a lemma of Littlewood and Offord*, Bull. Amer. Math. Soc. **51** (1945), 898–902.

3. Erdős P., *On sets of distances of n points*, Amer. Math. Monthly **53** (1946), 248–250.

4. Erdős P., *On extremal problems of graphs and generalized graphs*, Israel J. Math. **2** (1964), 183–190.

5. Erdős P., *On the number of complete subgraphs and circuits contained in graphs*, Časopis Pěst. Mat. **94** (1969), 290–296.

6. Erdős P., Makai Jr. E. and Pach J., *Nearly equal distances in the plane*, Combin. Probab. Comput. **2** (1993), 401–408.

7. Erdős P., Makai Jr. E. and Pach J., *Two nearly equal distances in* $\mathbb{R}^d$, `arXiv:1901.01055`.

8. Erdős P., Makai Jr. E., Pach J. and Spencer J., *Gaps in difference sets and the graph of nearly equal distances*, in: Applied Geometry and Discrete Mathematics: The Victor Klee Festschrift, DIMACS Series in Discr. Math. and Theoretical Comput. Sci. 4, Amer. Math. Soc., 1991, 265–273.

9. Guth L. and Katz N. H., *On the Erdős distinct distances problem in the plane*, Ann. of Math. (2) **181** (2015), 155–190.

10. Makai Jr. E., Pach J. and Spencer J., *New results on the distribution of distances determined by separated point sets*, in: Paul Erdős and his Mathematics II, Bolyai Society Mathematical Studies 11 (G. Halász, L. Lovász, M. Simonovits, V. T. Sós, eds.), János Bolyai Math. Soc., 2002, 499–511.

11. Solymosi J. and Van H., *Near optimal bounds for the Erdős distinct distances problem in high dimensions*, Combinatorica **28** (2008), 113–125.

N. Frankl, The London School of Economics and Political Science, London, UK,
e-mail: `n.frankl@lse.ac.uk`

A. Kupavskii, Moscow Institute of Physics and Technology, Moscow, Russia;
University of Oxford, Oxford, United Kingdom,
e-mail: `kupavskii@yandex.ru`