Andrey Kupavskii, Nabil H. Mustafa and Konrad Swanepoel

Bounding the size of an almost-equidistant set in Euclidean space

Article (Accepted version) (Refereed)

Bounding the size of an almost-equidistant set in Euclidean space

Andrey Kupavskii* Nabil H. Mustafa†‡ Konrad J. Swanepoel§

Abstract

A set of points in d-dimensional Euclidean space is *almost equidistant* if among any three points of the set, some two are at distance 1. We show that an almost-equidistant set in $\mathbb{R}^d$ has cardinality $O(d^{4/3})$.

1 Introduction

A set of lines through the origin of Euclidean d-space $\mathbb{R}^d$ is *almost orthogonal* if among any three of the lines, some two are orthogonal. Erdős asked (see [12]) what is the largest cardinality of an almost-orthogonal set of lines in $\mathbb{R}^d$? By taking the union of two sets of d pairwise orthogonal lines, we see that $2d$ is possible. Rosenfeld [12] showed that $2d$ is the maximum by considering the eigenvalues of the Gramian of the unit vectors spanning the lines. His result was subsequently given simpler proofs by Pudlák [11] and Deaett [6].

In this note we consider the analogous notion obtained by replacing orthogonal pairs of lines by pairs of points at unit distance. A subset V of Euclidean d-space $\mathbb{R}^d$ is *almost equidistant* if among any three points in V, some two are at Euclidean distance 1. We investigate the largest size, which we denote by $f(d)$, of an almost-equidistant set in $\mathbb{R}^d$. Although asking for the size of this function is a very natural question, it seems to be harder than the question of Erdős, which can be refomulated as asking for the largest size of an almost-equidistant set on a sphere of radius $1/\sqrt{2}$ in $\mathbb{R}^d$. Before stating our main result, we give an overview of what is known about $f(d)$.

Bezdek, Naszódi and Visy [5] showed that $f(2) \leq 7$, and István Talata (personal communication, 2007) showed that the only almost-equidistant set in $\mathbb{R}^2$ with 7 points is the Moser spindle. Györey [7] showed that $f(3) \leq 10$ and that there is a unique almost-equidistant set of 10 points in $\mathbb{R}^3$, a configuration originally considered by Nechushtan [9]. The Moser spindle can be generalized to higher dimensions, giving an almost-equidistant set of $2d+3$ points in $\mathbb{R}^d$ [4]. (We mention that Bezdek and Langi [4] considered the variant of Erdős's problem where the radius of the sphere is arbitrary instead of $1/\sqrt{2}$.) A

*Moscow Institute of Physics and Technology, Ecole Polytechnique Fédérale de Lausanne. Email: kupavskii@yandex.ru

†Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, ESIEE Paris, France. Email: mustafan@esiee.fr

‡The work of Nabil H. Mustafa in this paper has been supported by the grant ANR SAGA (JCJC-14-CE25-0016-01).

§Department of Mathematics, London School of Economics and Political Science, London. Email: k.swanepoel@lse.ac.uk

construction of Larman and Rogers [8] shows that $f(5) \geq 16$. Since there does not exist a set of $d+2$ points in $\mathbb{R}^d$ that are pairwise at distance 1, it follows that $f(d) \leq R(d+2, 3)-1$, where the Ramsey number $R(a, b)$ is the smallest n such that whenever each edge of the complete graph on n vertices is coloured blue or red, there is either a blue clique of size a or a red clique of size b. Ajtai, Komlós, and Szemerédi [1] showed $R(k, 3) = O(k^2/\log k)$, which implies the asymptotic upper bound $f(d) \leq O(d^2/\log d)$. Balko, Pór, Scheuer, Swanepoel, and Valtr [3] generalized the Nechushtan configuration to higher dimensions, giving $f(d) \geq 2d + 4$ for all $d \geq 3$. They also obtained the asymptotic upper bound $f(d) = O(d^{3/2})$ by an argument based on Deaett's paper [6]. Using computer search and ad hoc geometric arguments, they obtained the following bounds for small d: $f(4) \leq 13$, $f(5) \leq 20$, $18 \leq f(6) \leq 26$, $20 \leq f(7) \leq 34$, and $f(9) \geq f(8) \geq 24$. Polyanskii [10] subsequently improved the asymptotic upper bound to $f(d) = O(d^{13/9})$.

In this note we obtain a further improvement to the upper bound.

Theorem 1. *An almost-equidistant set in $\mathbb{R}^d$ has cardinality $O(d^{4/3})$.*

Its proof is based on the approach of [3] to show the upper bound $O(d^{3/2})$, which is in turn based on Deaett's proof [6] of Rosenfeld's result. Before proving Theorem 1 in Section 3, we establish notation and collect some lemmas in the next section.

2 Preliminaries

We denote the Euclidean norm of $x \in \mathbb{R}^d$ by $\|x\|$ and the inner product of $x, y \in \mathbb{R}^d$ by $\langle x, y\rangle$. The cardinality of a finite set A is denoted by $|A|$. We call a finite non-empty subset C of $\mathbb{R}^d$ (the vertex set of) a *unit simplex* if the distance between any two points in C equals 1. It has already been mentioned in the Introduction that if C is a unit simplex then $|C| \leq d + 1$. Given any finite $V \subset \mathbb{R}^d$, we define the *unit-distance graph* $G = (V, E)$ *on* V to be the graph with $vw \in E$ iff $\|v - w\| = 1$. Thus, $C \subset V$ is a unit simplex iff it is a clique in G. We denote the set of neighbours of $v \in V$ in G by $N(v)$.

The following well-known lemma gives a lower bound for the rank of a square matrix in terms of its entries [2, 6, 11].

Lemma 1. *For any non-zero $n \times n$ symmetric matrix $A = [a_{i,j}]$,*

$$\operatorname{rank}(A) \geq \frac{(\sum_i a_{i,i})^2}{\sum_{i,j} a_{i,j}^2}.$$

For the sake of completeness, we include the proofs of the following three lemmas on the vertices and centroids of unit simplices.

Lemma 2. *Let C be a unit simplex with centroid $c = \frac{1}{|C|}\sum_{v\in C} v$. Then*

$$\|v - c\|^2 = \frac{1}{2}\left(1 - \frac{1}{|C|}\right) \quad \text{for all } v \in C,$$

and

$$\langle v - c, v' - c\rangle = -\frac{1}{2\,|C|} \quad \text{for all distinct } v, v' \in C.$$

Proof. We may translate C so that c is the origin o. Write $C = \{p_1, \ldots, p_k\}$. By symmetry, $\alpha := \|p_i\|^2$ is independent of i, and $\beta := \langle v_i, v_j \rangle$ $(i \neq j)$ is independent of i and j. Then

$$0 = \left\| \sum_{i=1}^{k} p_i \right\|^2 = k\alpha + k(k-1)\beta$$

and

$$1 = \|p_i - p_j\|^2 = 2\alpha - 2\beta.$$

Solving these two linear equations in α and β, we obtain $\alpha = \frac{1}{2} - \frac{1}{2k}$ and $\beta = -\frac{1}{2k}$. $\square$

Lemma 3. *Let C be a unit simplex with centroid $c = \frac{1}{|C|}\sum_{v \in C} v$, and let $F \subset C$ be a unit simplex with centroid $f = \frac{1}{|F|}\sum_{v \in F} v$. Then*

$$\|c - f\|^2 = \frac{1}{2}\left(\frac{1}{|F|} - \frac{1}{|C|}\right).$$

Proof. Let $k := |C|$, $\ell := |F|$. Then

$$\begin{aligned} \|f - c\|^2 = \left\| \frac{1}{\ell} \sum_{v \in F} (v - c) \right\|^2 &= \frac{1}{\ell^2}\left(\ell \cdot \frac{1}{2}\Big(1 - \frac{1}{k}\Big) - \frac{\ell(\ell-1)}{2k}\right) \quad \text{by Lemma 2} \\ &= \frac{1}{2}\left(\frac{1}{\ell} - \frac{1}{k}\right). \end{aligned}$$

$\square$

Lemma 4. *Let A and B be disjoint unit simplices with centroids $a = \frac{1}{|A|}\sum_{v \in A} v$ and $b = \frac{1}{|B|}\sum_{v \in B} v$, respectively, such that $A \cup B$ is also a unit simplex. Then*

$$\|a - b\|^2 = \frac{1}{2}\left(\frac{1}{|A|} + \frac{1}{|B|}\right).$$

Proof. Let $C := A \cup B$ have centroid c. Then by Lemma 3, $\|a - c\|^2 = \frac{1}{2}\left(\frac{1}{|A|} - \frac{1}{|C|}\right)$ and $\|b - c\|^2 = \frac{1}{2}\left(\frac{1}{|B|} - \frac{1}{|C|}\right)$. It follows that

$$\begin{aligned} \|a - b\|^2 &= (\|a - c\| + \|b - c\|)^2 \\ &= \frac{1}{2}\left(\frac{1}{|A|} + \frac{1}{|B|} - \frac{2}{|C|}\right) + \sqrt{\left(\frac{1}{|A|} - \frac{1}{|C|}\right)\left(\frac{1}{|B|} - \frac{1}{|C|}\right)} \\ &= \frac{1}{2}\left(\frac{1}{|A|} + \frac{1}{|B|}\right) \end{aligned}$$

$\square$

3 Proof of Theorem 1

Let G be the unit-distance graph of a given almost-equidistant set V. Then the complement of G is K_3-free, and the non-neighbours of any vertex form a unit simplex. Let C be a clique of maximum cardinality in G. Write $k = |C|$. Each $v \in V \setminus C$ is a non-neighbour of some point in C, and it follows that $|V| \leq |C| + |C|\,k = k^2 + k$. Thus, without loss of generality, $k > d^{2/3}$.

We split V up into two parts, each to be bounded separately. Let

$$N = \left\{ v \in V : |N(v) \cap C| \geq k - k^{4/3} d^{-2/3} \right\}.$$

Note that $k^{4/3}d^{-2/3} = O(d^{-1/3})k$. We first bound the complement of N. Consider the set

$$X = \{(u, v) \in C \times V \setminus N : uv \notin E(G)\}.$$

For each $v \in V \setminus N$, there are more than $k^{4/3}d^{-2/3}$ points $u \in C$ such that $u \notin N(v)$, hence $|X| > k^{4/3}d^{-2/3}\,|V \setminus N|$. On the other hand, for each $u \in C$, the set of non-neighbours of u forms a clique, so has cardinality at most k, and $|X| \leq |C|\,k = k^2$. It follows that

$$|V \setminus N| < k^{2/3}d^{2/3}. \tag{1}$$

Next, we estimate $|N|$. Without loss of generality, $\frac{1}{k}\sum_{v \in C} v = o$ and $N = \{v_1, \ldots, v_n\}$. We want to apply Lemma 1 to the $n \times n$ matrix $A = [\langle v_i, v_j \rangle]$, which has rank at most d.

Claim 1. *For each $i = 1 \ldots, n$, $\|v_i\|^2 = \frac{1}{2} + O(k^{-1/3}d^{-1/3})$, and for each $v_iv_j \in E(G)$, $\langle v_i, v_j \rangle = O(k^{-1/3}d^{-1/3})$.*

Proof of Claim 1. Let $C_i := N(v_i) \cap C$, $k_i := |C_i|$, and $c_i := \frac{1}{k_i}\sum_{v \in C_i} v$. Then $k_i \geq k - k^{4/3}d^{-2/3}$. By Lemma 4 applied to $A = \{v_i\}$ and $B = C_i$, $\|v_i - c_i\|^2 = \frac{1}{2}\left(1 + \frac{1}{k_i}\right)$, hence $\|v_i - c_i\| = \frac{1}{\sqrt{2}} + O(k^{-1})$. By Lemma 3 applied to C and $F = C_i$,

$$\|c_i\| = \sqrt{\frac{1}{2}\left(\frac{1}{k_i} - \frac{1}{k}\right)} \leq \sqrt{\frac{1}{2}\left(\frac{1}{k - k^{4/3}d^{-2/3}} - \frac{1}{k}\right)} = O(k^{-1/3}d^{-1/3}).$$

By the triangle inequality,

$$\|v_i\| = \|v_i - c_i\| + O(\|c_i\|) = \frac{1}{\sqrt{2}} + O(k^{-1/3}d^{-1/3}),$$

and $\|v_i\|^2 = \frac{1}{2} + O(k^{-1/3}d^{-1/3})$. Also, $2\langle v_i, v_j \rangle = \|v_i\|^2 + \|v_j\|^2 - 1 = O(k^{-1/3}d^{-1/3})$. $\square$

Claim 2. *For each $i = 1, \ldots, n$,*

$$\sum_{\substack{j=1 \\ v_iv_j \notin E(G)}}^{n} \langle v_i, v_j \rangle^2 = O(k^{2/3}d^{-1/3}).$$

Proof of Claim 2. The non-neighbours $N \setminus N(v_i)$ of v_i form a unit simplex with cardinality $t := |N \setminus N(v_i)| \leq k$ and with centroid c, say. If $t = d + 1$, remove one point v_j from the unit simplex, which decreases the sum by $\langle v_i, v_j \rangle^2 = O(1)$. Thus, without loss of generality, $t \leq d$, and there exists a point $p \in \mathbb{R}^d$ such that $p - c$ is orthogonal to the affine hull of $N \setminus N(v_i)$, $\|p - c\| = 1/\sqrt{2t}$, the set $\{v_j - p : v_j \in N \setminus N(v_i)\}$ is orthogonal, and $\|v_j - p\| = 1/\sqrt{2}$ for each non-neighbour v_j of v_i. Then, by the finite Bessel inequality,

$$\sum_{v_j \in N \setminus N(v_i)} \langle v_i, v_j - p \rangle^2 \leq \frac{1}{2}\|v_i\|^2,$$

hence by applying Cauchy–Schwarz a few times,

$$\begin{aligned}
\sum_{v_j \in N \setminus N(v_i)} \langle v_i, v_j \rangle^2 &= \sum_{v_j \in N \setminus N(v_i)} \big(\langle v_i, v_j - p \rangle + \langle v_i, p - c \rangle + \langle v_i, c \rangle \big)^2 \\
&\le 3 \sum_{v_j \in N \setminus N(v_i)} \big(\langle v_i, v_j - p \rangle^2 + \langle v_i, p - c \rangle^2 + \langle v_i, c \rangle^2 \big) \\
&\le 3 \left(\frac{1}{2} \|v_i\|^2 + t \|v_i\|^2 \|p - c\|^2 + t \|v_i\|^2 \|c\|^2 \right) \\
&\le 3(\|v_i\|^2 + t \|v_i\|^2 \|c\|^2). \qquad (2)
\end{aligned}$$

By Claim 1, $\|v_i\|^2 = \frac{1}{2} + O(k^{-1/3} d^{-1/3})$ and

$$\begin{aligned}
\|c\|^2 = \Big\| \frac{1}{t} \sum_{v_j \in N \setminus N(v_i)} v_j \Big\|^2 &= \frac{1}{t^2} \Bigg(\sum_{v_j \in N \setminus N(v_i)} \|v_j\|^2 + \sum_{\substack{v_j, v_{j'} \in N \setminus N(v_i) \\ v_j \neq v_{j'}}} \langle v_j, v_{j'} \rangle \Bigg) \\
&\le \frac{1}{t^2} \left(t \left(\frac{1}{2} + O(k^{-1/3} d^{-1/3}) \right) + t(t-1) O(k^{-1/3} d^{-1/3}) \right) \\
&= \frac{1}{2t} + O(k^{-1/3} d^{-1/3}).
\end{aligned}$$

Therefore,

$$t \|c\|^2 = \frac{1}{2} + O(t k^{-1/3} d^{-1/3}) = O(k^{2/3} d^{-1/3}).$$

Substitute this back into (2) to finish the proof of Claim 2. □

We now finish the proof of the theorem. By Claim 2,

$$\begin{aligned}
\sum_{j=1}^{n} \langle v_i, v_j \rangle^2 &= \|v_i\|^4 + \sum_{v_j \in N(v_i)} \langle v_i, v_j \rangle^2 + O(k^{2/3} d^{-1/3}) \\
&= n O(k^{-2/3} d^{-2/3}) + O(k^{2/3} d^{-1/3}) \quad \text{by Claim 1.}
\end{aligned}$$

Also by Claim 1, $\sum_{i=1}^{n} \|v_i\|^2 = \Omega(n)$. Therefore, by Lemma 1,

$$d \ge \operatorname{rank}(A) \ge \frac{\left(\sum_{i=1}^{n} \|v_i\|^2 \right)^2}{\sum_{i,j=1}^{n} \langle v_i, v_j \rangle^2} = \frac{\Omega(n^2)}{n\big(n O(k^{-2/3} d^{-2/3}) + O(k^{2/3} d^{-1/3})\big)},$$

hence $n = O(n k^{-2/3} d^{1/3}) + O(k^{2/3} d^{2/3})$. Since $O(k^{-2/3} d^{1/3}) = o(1)$, it follows that $|N| = n = O(k^{2/3} d^{2/3})$. Recalling (1), we obtain that

$$|V| = |N| + |V \setminus N| = O(k^{2/3} d^{2/3}) = O(d^{4/3}).$$

□

Acknowledgements

We thank Nóra Frankl and Alexandr Polyanskii for helpful conversations.

References

[1] Miklós Ajtai, János Komlós, and Endre Szemerédi, *A note on Ramsey numbers*, J. Combin. Theory Ser. A **29** (1980), no. 3, 354–360.

[2] Noga Alon, *Problems and results in extremal combinatorics. I*, Discrete Math. **273** (2003), no. 1-3, 31–53.

[3] Martin Balko, Attila Pór, Manfred Scheucher, Konrad Swanepoel, and Pavel Valtr, *Almost-equidistant sets*, 2017, `http://arxiv.org/abs/1706.06375`

[4] Károly Bezdek and Zsolt Lángi, *Almost equidistant points on S^{d-1}*, Period. Math. Hungar. **39** (1999), no. 1-3, 139–144.

[5] Károly Bezdek, Márton Naszódi, and Balázs Visy, *On the mth Petty numbers of normed spaces*, Discrete geometry, Monogr. Textbooks Pure Appl. Math., vol. 253, Dekker, New York, 2003, pp. 291–304.

[6] Louis Deaett, *The minimum semidefinite rank of a triangle-free graph*, Linear Algebra Appl. **434** (2011), no. 8, 1945–1955.

[7] Bernadett Györey, *Diszkrét metrikus terek beágyazásai*, Master's thesis, Eötvös Loránd University, Budapest, 2004, in Hungarian.

[8] David G. Larman and Claude A. Rogers, *The realization of distances within sets in Euclidean space*, Mathematika **19** (1972), 1–24.

[9] Oren Nechushtan, *On the space chromatic number*, Discrete Math. **256** (2002), no. 1–2, 499–507.

[10] Alexandr Polyanskii, *On almost-equidistant sets*, 2017, `http://arxiv.org/abs/1707.00295`

[11] Pavel Pudlák, *Cycles of nonzero elements in low rank matrices*, Combinatorica **22** (2002), no. 2, 321–334.

[12] Moshe Rosenfeld, *Almost orthogonal lines in E^d*, Applied geometry and discrete mathematics. The Victor Klee Festschrift (Peter Gritzmann and Bernd Sturmfels, eds.), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 4, Amer. Math. Soc., Providence, RI, 1991, pp. 489–492.