Kusner's conjecture: Exact values and linear bounds

Hong-Jun Ge* Zixiang Xu† Yang Zhou*

Abstract

In 1983, Kusner conjectured that the largest equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$ has cardinality $n+1$ when $1<p<\infty$ and $2n$ when $p=1$. The conjecture is false for $1<p<2$, while in the range $p \geqslant 2$ it was previously known only in the isolated cases $p=2$ and $p=4$. The best previously known general upper bound was $O_p(n^{\frac{2p+2}{2p-1}})$ due to the celebrated work of Alon and Pudlák [GAFA, 2003]. Our main contributions are as follows.

(1) We prove Kusner's conjecture for every dimension $n \geqslant 1$ when $2 \leqslant p \leqslant 4$. More generally, for every integer $k \geqslant 0$ and every $p \in [4k+2, 4k+4]$, every equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$ has cardinality at most $(2k+1)n+1$. On the complementary intervals $p \in (4k, 4k+2)$ with $p \geqslant 1$, we obtain the almost linear bound $O_p(n \log n)$.

(2) We also consider the analogous problem on the torus $(\mathbb{T}^n, d_p)$, recently initiated by Alon, where the cyclic distance makes the problem substantially more delicate than in $\mathbb{R}^n$. We prove the almost linear bound $O_p(n \log n)$ for $1 \leqslant p \leqslant 2$ and $O_p(n^{\frac{3}{2}-\frac{1}{p}})$ for every fixed $p>2$, improving Alon's bounds $O_p(n^{2+\frac{2}{\lceil p \rceil}})$ for all $1 \leqslant p < \infty$.

1 Introduction

A subset A of a metric space is called *equilateral* if all distances between distinct points of A are equal. Equilateral sets are among the most basic and extensively studied configurations in metric and discrete geometry. Besides their classical role in Euclidean and normed spaces, they are closely connected to coding theory, spherical codes, and other problems in algebraic and extremal combinatorics [3, 7, 11, 14, 23]. In this paper, we study the maximum cardinality of an equilateral set in two families of spaces: the normed spaces $(\mathbb{R}^n, \|\cdot\|_p)$ and their toroidal metric analogues $(\mathbb{T}^n, d_p)$.

1.1 Kusner's conjecture

The study of equilateral sets in normed spaces is a classical topic in discrete geometry and Banach space theory. For a normed space X, let $e(X)$ denote the maximum cardinality of an equilateral set in X. For $1 \leqslant p < \infty$, we write $\ell_p^n := (\mathbb{R}^n, \|\cdot\|_p)$, where $\mathbb{R}^n$ is equipped with the ℓ_p-norm $\|\boldsymbol{x}\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$. The Euclidean case is completely understood: every equilateral set in $(\mathbb{R}^n, \|\cdot\|_2)$ has at most $n+1$ points, with equality precisely for the vertex sets of regular simplices. Hence $e(\ell_2^n) = n+1$. For general ℓ_p-norms, however, the situation is much more delicate. A theorem of Petty [16] shows that every equilateral set in an n-dimensional normed space has at most 2^n points, with equality only for ℓ_∞^n. On the other hand, for every $1<p<\infty$, the standard basis vectors together with a suitable multiple of the all-ones vector form an equilateral set of size $n+1$, while in $(\mathbb{R}^n, \|\cdot\|_1)$ the set $\{\pm \boldsymbol{e}_i : 1 \leqslant i \leqslant n\}$ is an equilateral set of size $2n$. Thus $n+1$ is the natural lower bound for finite $p>1$, and $2n$ is the natural lower bound for $p=1$.

*School of Mathematical Sciences, University of Science and Technology of China, Hefei, China. Emails: gehj22@mail.ustc.edu.cn and zy19700816@mail.ustc.edu.cn.

†School of Mathematical Sciences, Zhejiang University, Hangzhou, China. Email: zixiangxu@zju.edu.cn.

In his list of open problems, Guy recorded the following conjecture of Kusner [10], which has become one of the central problems in the theory of equilateral sets in normed spaces.

Conjecture 1.1 (Kusner's conjecture). *$e(\ell_1^n) = 2n$ and $e(\ell_p^n) = n+1$ for all $1 < p < \infty$.*

These questions have driven much of the subsequent work on equilateral sets in $(\mathbb{R}^n, \|\cdot\|_p)$. Prior to the present work, the exact picture was known only in a few cases. The Euclidean case $p = 2$ is classical. Swanepoel [21] proved the quartic case $e(\ell_4^n) = n+1$ for all $n \geqslant 1$ via polynomial methods, see also [22] for another proof. For $e(\ell_1^n)$, the conjectured value $2n$ is known only in small dimensions: it holds for $n \leqslant 4$, with the cases $n = 3$ and $n = 4$ established by Bandelt, Chepoi and Laurent [4] and Koolen, Laurent and Schrijver [13], respectively.

A major turning point was Swanepoel's discovery that Kusner's conjecture is false throughout the whole range $1 < p < 2$ in sufficiently high dimension [21]. More precisely, for every fixed $1 < p < 2$, Swanepoel [21] constructed equilateral sets in $(\mathbb{R}^n, \|\cdot\|_p)$ of cardinality at least $(1+\varepsilon_p)n$ for some $\varepsilon_p > 0$. Thus the difficult part of Kusner's conjecture is not the sub-Euclidean regime $1 < p < 2$, which is now known to be false, but rather the super-Euclidean regime $p > 2$.

As for upper bounds, the first nontrivial polynomial estimate goes back to Smyth [19], who proved $e(\ell_p^n) \leqslant C_p n^{\frac{p+1}{p-1}}$ for every fixed $1 < p < \infty$, see Smyth's survey [20]. This was substantially sharpened by Alon and Pudlák [2], who established $e(\ell_p^n) \leqslant c_p n^{\frac{2p+2}{2p-1}}$ for every fixed $p \geqslant 1$ together with the stronger bound $e(\ell_p^n) \leqslant c_p n \log n$ for every odd integer $p \geqslant 1$. For even integers p, Swanepoel [21] obtained linear bounds:

$$e(\ell_p^n) \leqslant \begin{cases} \left(\frac{p}{2} - 1\right) \cdot n + 1, & p \equiv 0 \pmod 4, \\ \frac{p}{2} \cdot n + 1, & p \equiv 2 \pmod 4. \end{cases}$$

Later, Konyagin [12] gave a different approach which yields $e(\ell_p^n) \leqslant C_p n \log n$ for all $1 \leqslant p < 2$. This recovers the best currently known order of magnitude for $e(\ell_1^n)$ and extends the logarithmic bound from odd integers to the entire sub-Euclidean range. More recently, Chen, Gui, Tang, and Xiong [8] obtained an improvement in the regime where p is significantly larger than n.

Before the present work, the known results around Kusner's conjecture exhibited a striking asymmetry. On the one hand, Swanepoel [21] showed that the conjecture fails for every fixed $1 < p < 2$ in sufficiently high dimension, while the exact simplex bound is known for the isolated cases $p = 2$ and $p = 4$. On the other hand, in the super Euclidean regime $p > 2$, the only previously known exact case was $p = 4$, and linear upper bounds were known essentially only for even integer exponents. Thus, before the present work, even the most basic question, whether the equilateral number $e(\ell_p^n)$ should remain linear in n, was open for every fixed $p > 2$ except those even integers p.

Our main result changes this picture substantially. It shows that for a whole half of the parameters $p \geqslant 2$, the cardinality of an equilateral set is always bounded linearly in the dimension.

Theorem 1.2. *Let k and n be integers with $k \geqslant 0$ and $n \geqslant 1$, and let $p \in [4k+2, 4k+4]$. If A is an equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$, then*

$$|A| \leqslant (2k+1)n + 1.$$

Theorem 1.2 has two immediate consequences. First, it gives a linear upper bound for all

$$p \in [2,4] \cup [6,8] \cup [10,12] \cup \cdots,$$

a set of exponents of density one half in $[2, \infty)$. Thus Kusner's conjecture is confirmed up to a constant factor on a large subset of the super Euclidean range. Second, in the first interval $2 \leqslant p \leqslant 4$, the bound is exactly $n+1$, and hence Kusner's conjecture holds there for every dimension n. To the best of our knowledge, this is the first verification of the conjectured simplex bound on a continuum of exponents.

On the complementary intervals $(4k, 4k+2)$, the situation is more delicate. The finite-dimensional correction used above no longer has the required sign pattern. Nevertheless, a different argument yields an almost-linear bound, losing only a logarithmic factor.

Theorem 1.3. *Let k and n be integers with $k \geqslant 0$ and $n \geqslant 1$, and let $p \in (4k, 4k+2)$ with $p \geqslant 1$. If A is an equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$, then*

$$|A| \leqslant C_p n \log(2n)$$

for some constant $C_p > 0$ depending only on p.

1.2 Largest equilateral sets on the torus

We next turn to the analogous problem on the torus, recently studied by Alon [1]. One motivation for this toroidal version comes from the work of Bilyk, Nagel, and Ruohoniemi on tensor product energies on $\mathbb{T}^n$, which arise naturally in discrepancy theory and quasi-Monte Carlo integration [6]. A key role in their work is played by Latin hypercube configurations with a single distance vector. In finite dimension, the single-distance-vector property is equivalent to requiring the point set to be equilateral for the toroidal ℓ_p-metric d_p, $1 \leqslant p \leqslant \infty$. Thus, for a fixed p, the toroidal equilateral problem can be viewed as a natural relaxation of this more structured setting. Let $\mathbb{T} := \mathbb{R}/\mathbb{Z}$, which we identify with the interval $[0,1)$ endowed with the cyclic distance

$$d_{\mathbb{T}}(u,v) := \min\{|u-v|,\ 1-|u-v|\}.$$

For $\boldsymbol{x} = (x_1, \ldots, x_n), \boldsymbol{y} = (y_1, \ldots, y_n) \in \mathbb{T}^n$ and $1 \leqslant p < \infty$, define

$$d_p(\boldsymbol{x}, \boldsymbol{y}) := \left(\sum_{i=1}^n d_{\mathbb{T}}(x_i, y_i)^p\right)^{1/p} = \left(\sum_{i=1}^n \min\{|x_i - y_i|,\ 1-|x_i-y_i|\}^p\right)^{1/p},$$

and for $p = \infty$, define $d_\infty(\boldsymbol{x}, \boldsymbol{y}) := \max_{1 \leqslant i \leqslant n} d_{\mathbb{T}}(x_i, y_i)$. Thus $(\mathbb{T}^n, d_p)$ is the n-dimensional torus equipped with the ℓ_p-product metric. Similarly, we write $e_p(\mathbb{T}^n) := e\big((\mathbb{T}^n, d_p)\big)$.

While equilateral sets in normed spaces have been studied for decades, the corresponding problem on the torus is much more recent. Alon [1] initiated the systematic study of the numbers $e_p(\mathbb{T}^n)$ and proved a general polynomial upper bound together with natural linear lower bounds.

Theorem 1.4 (Alon [1]). *For every fixed $p \geqslant 1$, there exists a constant $c(p) > 0$ such that for every integer $n \geqslant 1$,*

$$2n \leqslant e_p(\mathbb{T}^n) \leqslant c(p) n^{2+2/\lfloor p \rfloor}.$$

Moreover, if $2n+1$ is a prime, then

$$e_p(\mathbb{T}^n) \geqslant 2n+1.$$

In the same work, Alon also determined the exact value in the extremal case $p = \infty$, proving that $e_\infty(\mathbb{T}^n) = 3^n$. For finite p, however, the correct order of magnitude remained widely open. The lower bounds above strongly suggest that $e_p(\mathbb{T}^n)$ should be linear in n, and Alon explicitly conjectured that this is indeed the case, in particular that $e_1(\mathbb{T}^n) = \Theta(n)$ for all sufficiently large n.

Our next result gives the upper bounds that substantially improve Alon's general polynomial estimate for all finite $p \geqslant 1$.

Theorem 1.5. *Let $p \geqslant 1$. Then there exists a constant $C_p > 0$ such that for every integer $n \geqslant 1$,*

$$e_p(\mathbb{T}^n) \leqslant \begin{cases} C_p n \log(2n), & 1 \leqslant p \leqslant 2, \\ C_p n^{\frac{3}{2} - \frac{1}{p}}, & p > 2. \end{cases}$$

In particular, for $1 \leqslant p \leqslant 2$, this brings the toroidal analogue of Kusner-type problems to within a logarithmic factor of the conjectured linear behavior. For $p > 2$, it gives a substantially smaller exponent than Alon's bound $O_p(n^{2+2/\lfloor p \rfloor})$.

We also give two constructions which strengthen the known lower bounds. The first extends Alon's construction from prime values of $2n+1$ to odd prime powers.

Theorem 1.6. *Let q be an odd prime power. If $n = \frac{q-1}{2}$, then there exists a set $X \subseteq \mathbb{T}^n$ of cardinality $2n+1$ such that X is equilateral in $(\mathbb{T}^n, d_p)$ for every $1 \leqslant p \leqslant \infty$.*

The second construction shows that, for suitable choices $p = p_q$ depending on a prime q, the lower bound can be improved along an infinite sequence of dimensions.

Theorem 1.7. *There exist infinitely many primes q with the following property: there is a real number $p_q > \frac{q-1}{4}$ such that for every integer $m \geqslant 1$, if $n = \frac{q^m-1}{4}$, then $e_{p_q}(\mathbb{T}^n) \geqslant 4n+1$.*

2 Kusner's conjecture for $p \in [4k+2, 4k+4]$

We start with two pieces of notation that will be used throughout this section. The first is a convenient way to write distance matrices. For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $q > 0$, write $\ell_q(\boldsymbol{x}, \boldsymbol{y}) := (\sum_{r=1}^n |x_r - y_r|^q)^{1/q}$. Thus $\ell_p(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|_p$, which is the distance induced by the ℓ_p-norm on $\mathbb{R}^n$. The second is the standard notion of conditional negative definiteness. This notion is useful here because, after double centering, a conditionally negative definite kernel becomes a positive semidefinite matrix, which is exactly the linear algebraic mechanism behind our rank argument.

Definition 2.1. Let X be a set. A symmetric kernel $K : X \times X \to \mathbb{R}$ is called *conditionally negative definite* on X if for any $t \geqslant 2$ points $x_1, \ldots, x_t \in X$ and coefficients $a_1, \ldots, a_t \in \mathbb{R}$ with $\sum_{i=1}^t a_i = 0$, we have $\sum_{i,j=1}^t a_i a_j K(x_i, x_j) \leqslant 0$. Moreover, K is called *strictly conditionally negative definite* on X if $\sum_{i,j=1}^t a_i a_j K(x_i, x_j) < 0$ whenever $\sum_{i=1}^t a_i = 0$ and $(a_1, \ldots, a_t) \neq (0, \ldots, 0)$.

2.1 Useful tools

We collect the auxiliary results needed for the proof. The main input is a one-dimensional kernel decomposition. Its role is to correct the kernel $|x-y|^p$, on a prescribed finite set of real numbers, by a quadratic term and finitely many oscillatory terms so that the remaining kernel is conditionally negative definite. The crucial point is the sign control on the oscillatory coefficients: at most k of them are positive. This is what ultimately leads to the factor $(2k+1)n$ in the rank bound.

Theorem 2.2. *Let $p \in [4k+2, 4k+4]$ for some integer $k \geqslant 0$, and set $d := \lfloor p/2 \rfloor$. Let $A = \{x_1, \ldots, x_N\} \subseteq \mathbb{R}$ be a finite set. Then there exist constants $\theta_0 = \theta_0(p, A) > 0$ and $C_p(A) > 0$ such that, for every $0 < \theta < \theta_0$, there exist a symmetric function $E_\theta : A \times A \to \mathbb{R}$ and real coefficients $b_0 = b_0(\theta), b_1 = b_1(\theta), \ldots, b_{d-1} = b_{d-1}(\theta)$ with $|E_\theta(x,y)| \leqslant C_p(A)\theta^2$ for $x, y \in A$, such that the kernel*

$$K_\theta(x,y) := b_0|x-y|^2 + \sum_{h=1}^{d-1} b_h\big(1 - \cos(h\theta(x-y))\big) - |x-y|^p + E_\theta(x,y)$$

is conditionally negative definite on A. Moreover, among the coefficients $b_1, \ldots, b_{d-1}$, at most k are positive.

The remaining tools are elementary linear algebraic facts. The first one explains why conditionally negative definite kernels are useful for us: after double centering, they become positive semidefinite matrices.

Lemma 2.3. *Let $X = \{x_1, x_2, \ldots, x_m\}$ be a finite set, and let $K : X \times X \to \mathbb{R}$ be a symmetric kernel that is conditionally negative definite on X. Then the matrix*

$$-C_m\big(K(x_i, x_j)\big)_{i,j=1}^m C_m$$

is positive semidefinite, where $C_m = I_m - m^{-1}J_m$.

The second lemma is the rank-forcing step. It says that a positive semidefinite matrix, after adding the centered identity C_m and a sufficiently small symmetric error, must still have rank at least $m-1$.

Lemma 2.4. *Let $C_m := I_m - m^{-1}J_m$. Assume that*

$$G = NN^{\mathsf{T}} + C_m + E,$$

where E is a symmetric matrix satisfying $\lambda_{\min}(E) > -1$. Then

$$\operatorname{rank}(G) \geqslant m-1.$$

The last lemma is the low-rank input for the oscillatory terms. After centering, each basic cosine distance matrix is controlled by the centered cosine Gram matrix, which has rank at most 2.

Lemma 2.5. *Let $u_1, \dots, u_m \in \mathbb{R}$, let $\theta > 0$ and $r \geqslant 1$, and define*

$$D := \big(1 - \cos(r\theta(u_i - u_j))\big)_{i,j=1}^{m}.$$

Then the matrix $-C_m D C_m$ is positive semidefinite and has rank at most 2, where $C_m := I_m - m^{-1}J_m$.

2.2 Sketch of the proof

We give a short overview of the proof before entering the details. The argument is a rank comparison. Starting from an equilateral set $A = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_m\} \subseteq \mathbb{R}^n$ with $p \in [4k+2, 4k+4]$, we normalize so that $\ell_p(\boldsymbol{x}_i, \boldsymbol{x}_j) = 1$ for all $i \neq j$. Our goal is to prove $m \leqslant (2k+1)n + 1$. The strategy is to build a matrix identity whose right hand side has rank at least $m-1$, while its left hand side has rank at most $(2k+1)n$.

The key input is the one-dimensional kernel decomposition in Theorem 2.2. Applied to the finite set of all coordinate values appearing in A, it expresses the kernel $|x-y|^p$, up to a small error, as a sum of a quadratic term, finitely many oscillatory terms of the form $1 - \cos(h\theta(x-y))$, and a conditionally negative definite remainder. The important point is not merely that such a decomposition exists, but that among the oscillatory coefficients $b_1, \dots, b_{d-1}$, at most k are positive. After summing over all coordinates, the full distance matrix of A is decomposed into a Euclidean quadratic part, a controlled number of positive-coefficient oscillatory parts, a conditionally negative definite part, and a small error.

The next step is to double-center the resulting matrix identity by

$$C_m = I_m - m^{-1}J_m.$$

This is where the equilateral condition enters in a very clean way. Since the matrix $\Lambda = \big(\ell_p(\boldsymbol{x}_i, \boldsymbol{x}_j)^p\big)_{i,j=1}^{m}$ is equal to $J_m - I_m$, double centering gives $C_m \Lambda C_m = -C_m$. Thus the equilateral condition contributes exactly the centered identity C_m, with the correct sign after rearrangement.

The conditionally negative definite part is favorable after this centering operation. Indeed, Lemma 2.3 says that if K is conditionally negative definite, then its centered matrix $-C_m(K(x_i, x_j))C_m$ is positive semidefinite. The oscillatory terms with negative coefficients are also favorable: by Lemma 2.5, each centered cosine kernel is positive semidefinite. Therefore all these favorable terms can be moved to the same side and absorbed into a positive semidefinite matrix. In schematic form, the centered identity becomes

$$2b_0 \widetilde{G}_2 + \sum_{t=1}^{n} G_\theta^+(t) = NN^{\mathsf{T}} + C_m + E.$$

Here $\widetilde{G}_2$ is the centered Gram matrix coming from the quadratic term $|x-y|^2$, the matrix $G_\theta^+(t)$ collects the positive-coefficient oscillatory contributions from the t-th coordinate, NN^{T} absorbs the positive semidefinite terms, and E is the small centered error.

The smallness of E is used only at this point. By taking $\theta > 0$ sufficiently small, the matrix E satisfies $\lambda_{\min}(E) > -1$. Lemma 2.4 then applies to the right hand side and gives the lower bound

$$\operatorname{rank}\big(NN^{\mathsf{T}} + C_m + E\big) \geqslant m - 1.$$

It remains to bound the rank of the left hand side from above. The matrix $\widetilde{G}_2$ is a centered Gram matrix of m vectors in $\mathbb{R}^n$, so $\operatorname{rank}(\widetilde{G}_2) \leqslant n$. For each fixed coordinate t, the matrix $G_\theta^+(t)$ is a sum of at most k centered cosine kernels, and each has rank at most 2 by Lemma 2.5. Hence $\operatorname{rank}(G_\theta^+(t)) \leqslant 2k$. Consequently,

$$\operatorname{rank}\Big(2b_0\widetilde{G}_2 + \sum_{t=1}^{n} G_\theta^+(t)\Big) \leqslant n + 2kn = (2k+1)n.$$

Comparing the lower and upper rank bounds gives $m - 1 \leqslant (2k+1)n$.

2.3 Proof of Theorem 1.2

Set $d := \lfloor p/2 \rfloor$. Let $A = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m\} \subseteq \mathbb{R}^n$ be an equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$. After scaling, we may assume that $\ell_p(\boldsymbol{x}_i, \boldsymbol{x}_j) = 1$ for all $i \neq j$. After translating the whole set, we may further assume that $\boldsymbol{x}_1 = \boldsymbol{0}$. Write $\boldsymbol{x}_i = (x_i^{(1)}, \ldots, x_i^{(n)})$ for $1 \leqslant i \leqslant m$. Since $\boldsymbol{x}_1 = \boldsymbol{0}$ and $\|\boldsymbol{x}_i - \boldsymbol{x}_1\|_p = 1$, every coordinate of every point lies in $[-1, 1]$. Thus $\boldsymbol{x}_i \in [-1,1]^n$ for all $1 \leqslant i \leqslant m$.

We now apply the one-dimensional kernel decomposition to all coordinate values that occur in A. For each coordinate $t \in [n]$, define

$$A_t := \{x_i^{(t)} : 1 \leqslant i \leqslant m\} \subseteq [-1, 1],$$

and set $B := \bigcup_{t=1}^n A_t$. It is clear that B is finite. Choose $\theta > 0$ sufficiently small so that $\theta < \theta_0(p, B)$ and $mnC_p(B)\theta^2 < 1$, where $\theta_0(p, B)$ and $C_p(B)$ are given by Theorem 2.2. Applying Theorem 2.2 to B, we obtain a symmetric function $E_\theta : B \times B \to \mathbb{R}$ and real coefficients $b_0, b_1, \ldots, b_{d-1}$ such that $|E_\theta(x, y)| \leqslant C_p(B)\theta^2$ for $x, y \in B$, the kernel

$$K(x, y) := b_0|x - y|^2 + \sum_{r=1}^{d-1} b_r\big(1 - \cos(r\theta(x - y))\big) - |x - y|^p + E_\theta(x, y)$$

is conditionally negative definite on B, and among $b_1, \ldots, b_{d-1}$ at most k are positive.

For every pair of points $\boldsymbol{x}, \boldsymbol{y} \in A$, the definition of K gives

$$\sum_{t=1}^{n}\big(K(x^{(t)}, y^{(t)}) - E_\theta(x^{(t)}, y^{(t)})\big) + \ell_p(\boldsymbol{x}, \boldsymbol{y})^p = b_0\|\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \sum_{t=1}^{n}\sum_{r=1}^{d-1} b_r\big(1 - \cos(r\theta(x^{(t)} - y^{(t)}))\big). \quad (1)$$

We now turn this scalar identity into a matrix identity. Define $M_K := \sum_{t=1}^n \big(K(x_i^{(t)}, x_j^{(t)})\big)_{i,j=1}^m$ and $M_E := \sum_{t=1}^n \big(E_\theta(x_i^{(t)}, x_j^{(t)})\big)_{i,j=1}^m$, and let $\Lambda := \big(\ell_p(\boldsymbol{x}_i, \boldsymbol{x}_j)^p\big)_{i,j=1}^m$. Since A is equilateral and $\ell_p(\boldsymbol{x}_i, \boldsymbol{x}_j) = 1$ for $i \neq j$, we have $\Lambda = J_m - I_m$.

We separate the oscillatory terms according to the signs of their coefficients. Put

$$R_+ := \{r \in \{1, 2, \ldots, d-1\} : b_r > 0\},$$

and

$$R_- := \{r \in \{1, 2, \ldots, d-1\} : b_r < 0\}$$

respectively. By our selection and Theorem 2.2, $|R_+| \leqslant k$. For each $t \in [n]$, define

$$D_\theta^+(t) := \sum_{r \in R_+} b_r \big(1 - \cos(r\theta(x_i^{(t)} - x_j^{(t)}))\big)_{i,j=1}^m$$

and

$$D_\theta^-(t) := \sum_{r \in R_-} (-b_r) \big(1 - \cos(r\theta(x_i^{(t)} - x_j^{(t)}))\big)_{i,j=1}^m.$$

Thus both $D_\theta^+(t)$ and $D_\theta^-(t)$ are nonnegative linear combinations of the basic cosine-distance matrices. Set $G_\theta^+(t) := -C_m D_\theta^+(t) C_m$ and $G_\theta^-(t) := -C_m D_\theta^-(t) C_m$. By Lemma 2.5, each of $G_\theta^+(t)$ and $G_\theta^-(t)$ is positive semidefinite. Moreover,

$$\operatorname{rank}(G_\theta^+(t)) \leqslant \sum_{r \in R_+} \operatorname{rank}\Big(-C_m \big(1 - \cos(r\theta(x_i^{(t)} - x_j^{(t)}))\big)_{i,j=1}^m C_m\Big) \leqslant 2|R_+| \leqslant 2k.$$

Recall that $C_m = I_m - m^{-1} J_m$. Next let $D_2 := \big(\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2^2\big)_{i,j=1}^m$ and $\widetilde{G}_2 := -\frac{1}{2} C_m D_2 C_m$. Notice that $\widetilde{G}_2$ is the Gram matrix of the centered vectors $\boldsymbol{x}_i - \overline{\boldsymbol{x}}$, where $\overline{\boldsymbol{x}} := \frac{1}{m}\sum_{j=1}^m \boldsymbol{x}_j$. Since these vectors lie in $\mathbb{R}^n$, we have

$$\operatorname{rank}(\widetilde{G}_2) \leqslant n.$$

By (1), we have

$$M_K - M_E + \Lambda = b_0 D_2 + \sum_{t=1}^n D_\theta^+(t) - \sum_{t=1}^n D_\theta^-(t).$$

Since $C_m J_m = J_m C_m = 0$ and $C_m^2 = C_m$, we have

$$C_m \Lambda C_m = C_m (J_m - I_m) C_m = -C_m.$$

Multiplying the preceding matrix identity by $-C_m$ on the left and by C_m on the right, we obtain

$$-C_m M_K C_m + C_m M_E C_m + C_m = 2b_0 \widetilde{G}_2 + \sum_{t=1}^n G_\theta^+(t) - \sum_{t=1}^n G_\theta^-(t),$$

which implies that

$$2b_0 \widetilde{G}_2 + \sum_{t=1}^n G_\theta^+(t) = \Big(-C_m M_K C_m + \sum_{t=1}^n G_\theta^-(t)\Big) + C_m + C_m M_E C_m. \tag{2}$$

We then estimate the rank of the right-hand side from below. First, observe that the coordinatewise sum of K is still conditionally negative definite on A. Indeed, let $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ satisfy $\sum_{i=1}^m \lambda_i = 0$. Then

$$\sum_{i,j=1}^m \lambda_i \lambda_j \sum_{t=1}^n K(x_i^{(t)}, x_j^{(t)}) = \sum_{t=1}^n \sum_{i,j=1}^m \lambda_i \lambda_j K(x_i^{(t)}, x_j^{(t)}).$$

For each fixed t, all points $x_i^{(t)}$ belong to $B = \bigcup_{t=1}^n A_t$. Since K is conditionally negative definite on B, the inner sum is non-positive. If some of the values $x_i^{(t)}$ coincide, we simply combine the corresponding coefficients: for $z \in B$, set $\mu_z := \sum_{i : x_i^{(t)} = z} \lambda_i$. Then $\sum_z \mu_z = 0$, and

$$\sum_{i,j=1}^m \lambda_i \lambda_j K(x_i^{(t)}, x_j^{(t)}) = \sum_{z, z' \in B} \mu_z \mu_{z'} K(z, z') \leqslant 0.$$

Thus every coordinate contribution is conditionally negative definite, and their sum is conditionally negative definite as well. Hence the kernel

$$(\boldsymbol{x}_i, \boldsymbol{x}_j) \mapsto \sum_{t=1}^{n} K(x_i^{(t)}, x_j^{(t)})$$

is conditionally negative definite on A. By Lemma 2.3, $-C_m M_K C_m$ is positive semidefinite. Since each $G_\theta^-(t)$ is also positive semidefinite, the matrix

$$-C_m M_K C_m + \sum_{t=1}^{n} G_\theta^-(t)$$

is positive semidefinite. Since every real positive semidefinite matrix admits a Gram factorization, we may write

$$-C_m M_K C_m + \sum_{t=1}^{n} G_\theta^-(t) = NN^\mathsf{T}$$

for some real matrix N.

It remains to check that the error term is small enough. Recall that for a real symmetric matrix M, $\|M\|_{\mathrm{OP}}$ denotes the largest absolute value of its eigenvalues. Each entry of M_E has absolute value at most $nC_p(B)\theta^2$. Hence, by the Gershgorin circle theorem [9],

$$\|M_E\|_{\mathrm{OP}} \leqslant \max_{1\leqslant i\leqslant m} \sum_{j=1}^{m} |(M_E)_{ij}| \leqslant mnC_p(B)\theta^2 < 1.$$

Since C_m is an orthogonal projection,

$$\|C_m M_E C_m\|_{\mathrm{OP}} \leqslant \|M_E\|_{\mathrm{OP}} < 1.$$

The matrix $C_m M_E C_m$ is symmetric, and hence

$$\lambda_{\min}(C_m M_E C_m) > -1.$$

Thus the right hand side of (2) has the form

$$NN^\mathsf{T} + C_m + E$$

with $E = C_m M_E C_m$ and $\lambda_{\min}(E) > -1$. Then Lemma 2.4 gives

$$\operatorname{rank}\Big(\Big(-C_m M_K C_m + \sum_{t=1}^{n} G_\theta^-(t)\Big) + C_m + C_m M_E C_m\Big) \geqslant m-1.$$

Therefore, by (2),

$$m-1 \leqslant \operatorname{rank}\Big(2b_0\widetilde{G}_2 + \sum_{t=1}^{n} G_\theta^+(t)\Big).$$

On the other hand, the rank of the left hand side is controlled by the Euclidean part and the positive oscillatory parts:

$$\operatorname{rank}\Big(2b_0\widetilde{G}_2 + \sum_{t=1}^{n} G_\theta^+(t)\Big) \leqslant \operatorname{rank}(\widetilde{G}_2) + \sum_{t=1}^{n} \operatorname{rank}(G_\theta^+(t)) \leqslant n + 2kn = (2k+1)n.$$

Combining the two inequalities gives $m-1 \leqslant (2k+1)n$, and hence

$$m \leqslant (2k+1)n + 1.$$

This completes the proof.

2.4 Proofs of the tools

We now prove the auxiliary results stated above. The main task is to prove the one-dimensional decomposition, Theorem 2.2. The proof has two stages. First we show that, on any finite set of real numbers, the kernel $-|x-y|^p$ can be corrected by adding a suitable even polynomial in $x-y$, so that the resulting kernel is strictly conditionally negative definite. Then we approximate this correcting polynomial by a finite cosine sum with coefficients of prescribed signs.

2.4.1 Proof of Theorem 2.2

We begin with the polynomial correction. In the range $p \in (4k+2, 4k+4)$, one has $d = \lfloor p/2 \rfloor = 2k+1$. The oddness of d is the key point: it is what allows us to add a suitable even polynomial in $x-y$ to $-|x-y|^p$, so that the resulting kernel becomes strictly conditionally negative definite.

Lemma 2.6. *Let $p \in (4k+2, 4k+4)$ for some integer $k \geqslant 0$, and set $d := \lfloor p/2 \rfloor = 2k+1$. Let $A = \{x_1, \ldots, x_m\} \subseteq \mathbb{R}$ be a finite set. Then there exist real constants $C_2(A,p), C_4(A,p), \ldots, C_{2d}(A,p)$ such that the kernel*

$$K(x,y) := \sum_{r=1}^{d} C_{2r}(A,p)(x-y)^{2r} - |x-y|^p$$

is strictly conditionally negative definite on A.

Before giving the proof, we keep the notation $A = \{x_1, \ldots, x_m\}$ from the statement and introduce the following filtration by vanishing moments. For an even function $f : \mathbb{R} \to \mathbb{R}$, define

$$B_f(\boldsymbol{a}, \boldsymbol{b}) := \sum_{i,j=1}^{m} a_i b_j f(x_i - x_j).$$

In particular, we write

$$Q_f(\boldsymbol{a}) := B_f(\boldsymbol{a}, \boldsymbol{a}).$$

For each integer $r \geqslant 0$, define the r-th moment of $\boldsymbol{a} = (a_1, \ldots, a_m) \in \mathbb{R}^m$ by $M_r(\boldsymbol{a}) := \sum_{i=1}^m a_i x_i^r$, and set

$$V_r := \{\boldsymbol{a} \in \mathbb{R}^m : M_0(\boldsymbol{a}) = M_1(\boldsymbol{a}) = \cdots = M_r(\boldsymbol{a}) = 0\}.$$

Thus V_r consists of coefficient vectors whose moments up to order r vanish. In particular, $V_0 = \{\boldsymbol{a} \in \mathbb{R}^m : \sum_{i=1}^m a_i = 0\}$. These spaces form the decreasing filtration

$$V_0 \supseteq V_1 \supseteq \cdots \supseteq V_d.$$

The role of the correction polynomial is to make the resulting quadratic form negative definite on V_0. We achieve this by decreasing the moment order one step at a time, passing from V_d to V_{d-1} and eventually to V_0.

The first ingredient is the following elementary identity. It shows that, on V_{r-1}, the quadratic form induced by the kernel $(x-y)^{2r}$ depends only on the r-th moment.

Lemma 2.7. *Let $r \geqslant 1$. If $\boldsymbol{a}, \boldsymbol{b} \in V_{r-1}$, then*

$$B_{t^{2r}}(\boldsymbol{a}, \boldsymbol{b}) = (-1)^r \binom{2r}{r} M_r(\boldsymbol{a}) M_r(\boldsymbol{b}).$$

In particular:

(1) *if $\boldsymbol{a} \in V_{r-1}$ and $\boldsymbol{b} \in V_r$, then $B_{t^{2r}}(\boldsymbol{a}, \boldsymbol{b}) = 0$;*

(2) *if* $0 \neq \boldsymbol{a} \in V_{r-1} \backslash V_r$, *then* $Q_{(-1)^{r+1}t^{2r}}(\boldsymbol{a}) = -\binom{2r}{r} M_r(\boldsymbol{a})^2 < 0$.

Proof of Lemma 2.7. Expanding $(x_i - x_j)^{2r}$ gives

$$(x_i - x_j)^{2r} = \sum_{\ell=0}^{2r} (-1)^{2r-\ell} \binom{2r}{\ell} x_i^\ell x_j^{2r-\ell}.$$

Therefore

$$B_{t^{2r}}(\boldsymbol{a}, \boldsymbol{b}) = \sum_{\ell=0}^{2r} (-1)^{2r-\ell} \binom{2r}{\ell} M_\ell(\boldsymbol{a}) M_{2r-\ell}(\boldsymbol{b}).$$

Since $\boldsymbol{a}, \boldsymbol{b} \in V_{r-1}$, we have $M_\ell(\boldsymbol{a}) = M_\ell(\boldsymbol{b}) = 0$ for every $\ell < r$. Hence every term vanishes except the middle term $\ell = r$, and so

$$B_{t^{2r}}(\boldsymbol{a}, \boldsymbol{b}) = (-1)^r \binom{2r}{r} M_r(\boldsymbol{a}) M_r(\boldsymbol{b}).$$

The two stated consequences now follow directly from this formula. If $\boldsymbol{b} \in V_r$, then $M_r(\boldsymbol{b}) = 0$, and hence $B_{t^{2r}}(\boldsymbol{a}, \boldsymbol{b}) = 0$ for every $\boldsymbol{a} \in V_{r-1}$. On the other hand, if $0 \neq \boldsymbol{a} \in V_{r-1} \backslash V_r$, then $M_r(\boldsymbol{a}) \neq 0$. Therefore

$$Q_{(-1)^{r+1}t^{2r}}(\boldsymbol{a}) = (-1)^{r+1} B_{t^{2r}}(\boldsymbol{a}, \boldsymbol{a}) = -\binom{2r}{r} M_r(\boldsymbol{a})^2 < 0.$$

□

The second ingredient is the positivity of $|x-y|^p$ on V_d when $d = \lfloor p/2 \rfloor$ is odd. Recall that $Q_f(\boldsymbol{a}) := B_f(\boldsymbol{a}, \boldsymbol{a})$.

Lemma 2.8. *Let* $d \geqslant 0$ *be an integer, and let* p *satisfy* $2d < p < 2d+2$. *If* d *is odd, then the quadratic form* $Q_{|t|^p}$ *is positive definite on* V_d, *that is,* $Q_{|t|^p}(\boldsymbol{a}) > 0$ *for every* $\boldsymbol{0} \neq \boldsymbol{a} \in V_d$.

Proof of Lemma 2.8. Define

$$R_d(u) := \cos u - \sum_{\ell=0}^{d} (-1)^\ell \frac{u^{2\ell}}{(2\ell)!}.$$

Since $\cos u = \sum_{\ell=0}^{\infty} (-1)^\ell \frac{u^{2\ell}}{(2\ell)!}$, we have $R_d(u) = O(u^{2d+2})$ as $u \to 0$. Also, $R_d(u) = O(u^{2d})$ as $|u| \to \infty$. Therefore, the integral

$$J_{p,d} := \int_0^\infty \frac{R_d(u)}{u^{1+p}}\, du$$

converges absolutely for $2d < p < 2d+2$.

Integrating by parts $2d$ times, we claim that all boundary terms vanish. Indeed, at the $(j+1)$-st integration by parts, where $0 \leqslant j \leqslant 2d-1$, the boundary term is a constant multiple of $R_d^{(j)}(u) u^{j-p}$. As $u \to 0$, we have

$$R_d^{(j)}(u) = O(u^{2d+2-j}),$$

and hence

$$R_d^{(j)}(u) u^{j-p} = O(u^{2d+2-p}) = o(1),$$

because $p < 2d+2$. As $u \to \infty$, we have

$$R_d^{(j)}(u) = O(u^{2d-j}),$$

and hence

$$R_d^{(j)}(u) u^{j-p} = O(u^{2d-p}) = o(1),$$

because $p > 2d$. Thus all boundary terms vanish. We therefore obtain

$$J_{p,d} = \frac{(-1)^d}{p(p-1)\cdots(p-2d+1)} \int_0^\infty \frac{\cos u - 1}{u^{1+(p-2d)}}\, du.$$

As $\cos u - 1 \leqslant 0$, we have

$$(-1)^{d+1} J_{p,d} > 0.$$

Set $c_{p,d} := \frac{1}{|J_{p,d}|} > 0$. Then by the change of variables $u = t|x|$, for every $x \in \mathbb{R}$, we have

$$|x|^p = c_{p,d}(-1)^{d+1} \int_0^\infty \frac{R_d(t|x|)}{t^{1+p}}\, dt = c_{p,d}(-1)^{d+1} \int_0^\infty \frac{\cos(tx) - \sum_{\ell=0}^{d}(-1)^\ell (tx)^{2\ell}/(2\ell)!}{t^{1+p}}\, dt. \tag{3}$$

For $\boldsymbol{a} = (a_1, \ldots, a_m) \in V_d$, and $\ell = 0, 1, \ldots, d$, we have

$$M_\ell(\boldsymbol{a}) = \sum_{i=1}^m a_i x_i^\ell = 0.$$

Applying (3) to $x_i - x_j$, summing over i, j, and using Lemma 2.7 to eliminate all polynomial terms, we obtain

$$Q_{|t|^p}(\boldsymbol{a}) = c_{p,d}(-1)^{d+1} \int_0^\infty \frac{\sum_{i,j=1}^m a_i a_j \cos(t(x_i - x_j))}{t^{1+p}}\, dt.$$

By Euler's formula $e^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$, we have

$$\sum_{i,j=1}^m a_i a_j e^{\mathrm{i}t(x_i - x_j)} = \left(\sum_{i=1}^m a_i e^{\mathrm{i}tx_i}\right)\left(\sum_{j=1}^m a_j e^{-\mathrm{i}tx_j}\right) = \left|\sum_{i=1}^m a_i e^{\mathrm{i}tx_i}\right|^2.$$

Taking real parts gives

$$\sum_{i,j=1}^m a_i a_j \cos(t(x_i - x_j)) = \left|\sum_{i=1}^m a_i e^{\mathrm{i}tx_i}\right|^2.$$

It follows that

$$Q_{|t|^p}(\boldsymbol{a}) = c_{p,d}(-1)^{d+1} \int_0^\infty \frac{\left|\sum_{i=1}^m a_i e^{\mathrm{i}tx_i}\right|^2}{t^{1+p}}\, dt. \tag{4}$$

Because d is odd, we have $(-1)^{d+1} = 1$, so the right hand side is nonnegative.

It remains to prove strict positivity. Let

$$F_{\boldsymbol{a}}(t) := \sum_{i=1}^m a_i e^{\mathrm{i}tx_i}.$$

This is an entire function. If $F_{\boldsymbol{a}} \equiv 0$, then all derivatives at 0 vanish, so

$$\sum_{i=1}^m a_i x_i^r = 0$$

for every $r \geqslant 0$. In particular, this holds for $0 \leqslant r \leqslant m-1$. Since $x_1, \ldots, x_m$ are distinct, the Vandermonde matrix is invertible, and hence $a_1 = \cdots = a_m = 0$. Thus $F_{\boldsymbol{a}} \not\equiv 0$ whenever $\boldsymbol{a} \neq \mathbf{0}$. The integrand in (4) is therefore not identically zero, and the integral is strictly positive. This proves the lemma. □

We are now ready to prove Lemma 2.6.

Proof of Lemma 2.6. If $m = 1$, the conclusion is trivial, so assume $m \geqslant 2$. Set $s := \min\{d, m-1\}$. We first show that V_r has codimension 1 inside V_{r-1}.

Claim 2.9. *For each* $r \in \{1, 2, \ldots, s\}$, $\dim(V_{r-1}/V_r) = 1$.

Proof of the claim. Notice that the linear functionals $M_0, \ldots, M_s$ are linearly independent on $\mathbb{R}^m$. To see this, suppose that $\sum_{r=0}^{s} c_r M_r = 0$. Then the polynomial $q(x) := \sum_{r=0}^{s} c_r x^r$ vanishes at the distinct real numbers $x_1, \ldots, x_m$. Since $s \leqslant m-1$, this forces $q \equiv 0$, and hence $c_0 = \cdots = c_s = 0$. Therefore each time we pass from V_{r-1} to V_r, we impose exactly one new independent linear condition, namely $M_r(\boldsymbol{a}) = 0$. Thus $\dim(V_{r-1}/V_r) = 1$ for every $1 \leqslant r \leqslant s$. ■

Also, if $s < m-1$, then necessarily $s = d$, while if $s = m-1 \leqslant d$, then $V_s = \{\mathbf{0}\}$.

We shall construct the correcting coefficients by descending induction. For $1 \leqslant r \leqslant s$, define

$$K_r(x, y) := \sum_{j=r}^{s} C_{2j}(A, p)(x-y)^{2j} - |x-y|^p.$$

We will choose the coefficients so that

$$Q_{K_r}(\boldsymbol{a}) < 0 \tag{5}$$

for every $0 \neq \boldsymbol{a} \in V_{r-1}$.

Start with $K_{s+1}(x, y) := -|x-y|^p$. We claim that $Q_{K_{s+1}}$ is negative definite on V_s. Indeed, if $s = d$, this follows from Lemma 2.8, since $d = 2k+1$ is odd. If $s = m-1 < d$, then $V_s = \{\mathbf{0}\}$.

Fix $r \in \{1, 2, \ldots, s\}$, and assume that the coefficients $C_{2(r+1)}(A, p), \ldots, C_{2s}(A, p)$ have already been chosen so that $Q_{K_{r+1}}(\boldsymbol{a}) < 0$ for all $\mathbf{0} \neq \boldsymbol{a} \in V_r$. Thus the bilinear form $B_{K_{r+1}}$ is nondegenerate on V_r: if $B_{K_{r+1}}(\boldsymbol{u}, \boldsymbol{v}) = 0$ for every $\boldsymbol{v} \in V_r$, then in particular $Q_{K_{r+1}}(\boldsymbol{u}) = 0$, forcing $\boldsymbol{u} = 0$.

Consider the subspace

$$W_r := \{\boldsymbol{w} \in V_{r-1} : B_{K_{r+1}}(\boldsymbol{w}, \boldsymbol{v}) = 0 \text{ for all } \boldsymbol{v} \in V_r\}.$$

Since $B_{K_{r+1}}|_{V_r}$ is nondegenerate, we have $W_r \cap V_r = \{\mathbf{0}\}$. Consider the linear map

$$\Phi : V_{r-1} \longrightarrow V_r^*$$

defined as follows: for each $\boldsymbol{w} \in V_{r-1}$, the image $\Phi(\boldsymbol{w})$ is the linear functional on V_r given by

$$\Phi(\boldsymbol{w})(\boldsymbol{v}) = B_{K_{r+1}}(\boldsymbol{w}, \boldsymbol{v})$$

for every $\boldsymbol{v} \in V_r$. Since $B_{K_{r+1}}$ is nondegenerate on V_r, the restriction of Φ to V_r is an isomorphism from V_r onto V_r^*. In particular,

$$\dim(\operatorname{im}\Phi) = \dim V_r.$$

Moreover, $\ker(\Phi) = W_r$. Hence, by rank-nullity and Claim 2.9,

$$\dim W_r = \dim V_{r-1} - \dim V_r = 1,$$

and therefore $V_{r-1} = W_r \oplus V_r$.

Choose a nonzero vector $\boldsymbol{w}_r \in W_r$. Since $W_r \cap V_r = \{\mathbf{0}\}$, we have $\boldsymbol{w}_r \notin V_r$, and so $M_r(\boldsymbol{w}_r) \neq 0$. By Lemma 2.7,

$$Q_{(-1)^{r+1}t^{2r}}(\boldsymbol{w}_r) = -\binom{2r}{r} M_r(\boldsymbol{w}_r)^2 < 0. \tag{6}$$

Also, by the same lemma, for every $\boldsymbol{w} \in W_r \subseteq V_{r-1}$ and every $\boldsymbol{v} \in V_r$, $B_{t^{2r}}(\boldsymbol{w}, \boldsymbol{v}) = 0$.

Since $\boldsymbol{w}_r$ is fixed, by (6), we can choose $\lambda_r > 0$ so large that

$$Q_{K_{r+1}}(\boldsymbol{w}_r) + \lambda_r\, Q_{(-1)^{r+1}t^{2r}}(\boldsymbol{w}_r) < 0.$$

Set $C_{2r}(A,p) := \lambda_r(-1)^{r+1}$, and define

$$K_r(x,y) := K_{r+1}(x,y) + \lambda_r(-1)^{r+1}(x-y)^{2r}.$$

We claim that Q_{K_r} is negative definite on V_{r-1}. Let $\mathbf{0} \neq \boldsymbol{u} \in V_{r-1}$. Write uniquely

$$\boldsymbol{u} = \boldsymbol{w} + \boldsymbol{v}$$

for $\boldsymbol{w} \in W_r,\ \boldsymbol{v} \in V_r$. Since W_r is $B_{K_{r+1}}$-orthogonal to V_r, and $B_{t^{2r}}(W_r, V_r) = 0$, we have

$$Q_{K_r}(\boldsymbol{u}) = Q_{K_r}(\boldsymbol{w}) + Q_{K_r}(\boldsymbol{v}).$$

Now:

- if $\boldsymbol{v} \neq \mathbf{0}$, then $Q_{K_r}(\boldsymbol{v}) = Q_{K_{r+1}}(\boldsymbol{v}) < 0$, because $\boldsymbol{v} \in V_r$ and $Q_{t^{2r}}(\boldsymbol{v}) = 0$ by Lemma 2.7;
- if $\boldsymbol{w} \neq \mathbf{0}$, then $\boldsymbol{w} = \alpha \boldsymbol{w}_r$ for some $\alpha \neq 0$, and hence

$$Q_{K_r}(\boldsymbol{w}) = \alpha^2 \Big(Q_{K_{r+1}}(\boldsymbol{w}_r) + \lambda_r\, Q_{(-1)^{r+1}t^{2r}}(\boldsymbol{w}_r) \Big) < 0.$$

Therefore $Q_{K_r}(\boldsymbol{u}) < 0$ whenever $\boldsymbol{u} \neq \mathbf{0}$, and so (5) holds. This completes the descending induction.

At the end, we obtain coefficients $C_{2r}(A,p)$ for $1 \leqslant r \leqslant s$ such that $Q_{K_1}(\boldsymbol{a}) < 0$ for all $\mathbf{0} \neq \boldsymbol{a} \in V_0$. Finally, set $C_{2r}(A,p) := 0$ for $s < r \leqslant d$. Then

$$K(x,y) := \sum_{r=1}^{d} C_{2r}(A,p)\,(x-y)^{2r} - |x-y|^p$$

satisfies $Q_K(\boldsymbol{a}) < 0$ for all $\mathbf{0} \neq \boldsymbol{a} \in V_0$. Since

$$V_0 = \Big\{ \boldsymbol{a} \in \mathbb{R}^m : \sum_{i=1}^{m} a_i = 0 \Big\},$$

this exactly says that K is strictly conditionally negative definite on A. □

We now pass from the polynomial correction to the trigonometric correction required in Theorem 2.2.

Proof of Theorem 2.2. If $|A| = 1$, the conclusion is trivial. We assume $|A| \geqslant 2$.

We first handle the case $d = 1$. If $p = 2$, take $b_0 = 1$ and $E_\theta \equiv 0$. If $2 < p < 4$, then Lemma 2.6 gives a coefficient a_2 such that the kernel

$$a_2(x-y)^2 - |x-y|^p$$

is conditionally negative definite on A. In this case take $b_0 = a_2$ and $E_\theta \equiv 0$. Since there are no oscillatory terms when $d = 1$, the sign assertion is vacuous. Thus we may assume $d \geqslant 2$.

We first choose an even polynomial correction. If $p \in (4k+2, 4k+4)$, then Lemma 2.6 gives real coefficients $a_2, a_4, \dots, a_{2d}$ such that the kernel

$$K_{\mathrm{poly}}(x,y) := \sum_{r=1}^{d} a_{2r}(x-y)^{2r} - |x-y|^p$$

is conditionally negative definite on A. If $p \in \{4k+2, 4k+4\}$, then $p = 2d$. In this case take

$$(a_2, a_4, \dots, a_{2d-2}, a_{2d}) = (0, 0, \dots, 0, 1),$$

so that

$$K_{\mathrm{poly}}(x,y) = (x-y)^{2d} - |x-y|^{2d} = 0.$$

Thus in all cases K_{poly} is conditionally negative definite on A.

We then approximate the polynomial part by a finite cosine sum. For $\theta > 0$, define $b_1(\theta), \ldots, b_{d-1}(\theta)$ as the unique solution of

$$\sum_{h=1}^{d-1} b_h(\theta) h^{2r} = (-1)^{r+1} \frac{(2r)!}{\theta^{2r}} a_{2r} \tag{7}$$

for $r = 2, \ldots, d$. The coefficient matrix $(h^{2r})_{2 \leqslant r \leqslant d, 1 \leqslant h \leqslant d-1}$ is invertible. Indeed, after writing $\lambda_h = h^2$, it becomes a Vandermonde matrix in the distinct numbers $\lambda_1, \ldots, \lambda_{d-1}$, up to multiplication of columns by positive factors. Define

$$b_0(\theta) := a_2 - \frac{\theta^2}{2} \sum_{h=1}^{d-1} b_h(\theta) h^2 \tag{8}$$

and set

$$T_\theta(t) := b_0(\theta) t^2 + \sum_{h=1}^{d-1} b_h(\theta)(1 - \cos(h\theta t)).$$

Let

$$\Delta := A - A = \{x - y : x, y \in A\}.$$

Since Δ is finite, all estimates below are uniform for $t \in \Delta$. For each $h \in \{1, 2, \ldots, d-1\}$, Taylor's formula gives

$$1 - \cos(h\theta t) = \sum_{r=1}^{d} (-1)^{r+1} \frac{(h\theta t)^{2r}}{(2r)!} + R_h(\theta, t),$$

where $R_h(\theta, t) = O(\theta^{2d+2})$ uniformly for $t \in \Delta$.

Let

$$s := \max\{r \in \{1, 2, \ldots, d\} : a_{2r} \neq 0\}.$$

If $s = 1$, then the right hand side of (7) is zero for every $r = 2, \ldots, d$, and hence $b_h(\theta) = 0$ for all h. In this case $T_\theta(t) = a_2 t^2$, and the approximation below is exact. We may therefore assume $s \geqslant 2$. Since $a_{2r} = 0$ for all $r > s$, equation (7) gives $b_h(\theta) = O(\theta^{-2s})$ for $1 \leqslant h \leqslant d-1$. Hence, uniformly for $t \in \Delta$,

$$\sum_{h=1}^{d-1} b_h(\theta) R_h(\theta, t) = O(\theta^{2d+2-2s}) = O(\theta^2).$$

By (7) and (8), the Taylor coefficients of $T_\theta(t)$ up to degree $2d$ agree with those of $\sum_{r=1}^{d} a_{2r} t^{2r}$. Therefore, defining

$$E_\theta(x,y) := \sum_{r=1}^{d} a_{2r}(x-y)^{2r} - T_\theta(x-y),$$

we obtain

$$|E_\theta(x,y)| \leqslant C_p(A)\theta^2$$

for all $x, y \in A$, after increasing $C_p(A)$ if necessary. Since T_θ is even, E_θ is symmetric.

Now define

$$K_\theta(x,y) := b_0(\theta)|x-y|^2 + \sum_{h=1}^{d-1} b_h(\theta)(1 - \cos(h\theta(x-y))) - |x-y|^p + E_\theta(x,y).$$

By the definition of E_θ, we have

$$K_\theta(x,y) = K_{\mathrm{poly}}(x,y)$$

for all $x, y \in A$. Hence K_θ is conditionally negative definite on A.

It remains to control the signs of the coefficients $b_1(\theta), \ldots, b_{d-1}(\theta)$. If $s = 1$, then all these coefficients are zero, so there is nothing to prove. Assume $s \geqslant 2$. Write (7) in matrix form as

$$M\boldsymbol{b}(\theta) = \boldsymbol{u}(\theta),$$

where $\boldsymbol{b}(\theta) := (b_1(\theta), \ldots, b_{d-1}(\theta))^\mathsf{T}$ and

$$\boldsymbol{u}(\theta) := \left((-1)^{r+1}\frac{(2r)!}{\theta^{2r}}a_{2r}\right)_{r=2}^{d}.$$

Since $a_{2r} = 0$ for $r > s$ and $a_{2s} \neq 0$, we have

$$\boldsymbol{u}(\theta) = (-1)^{s+1}(2s)!a_{2s}\theta^{-2s}\boldsymbol{e}_{s-1} + O(\theta^{-2s+2}),$$

where $\boldsymbol{e}_{s-1}$ denotes the $(s-1)$-st standard basis vector of $\mathbb{R}^{d-1}$. Therefore

$$b_h(\theta) = (-1)^{s+1}(2s)!a_{2s}(M^{-1})_{h,s-1}\theta^{-2s} + O(\theta^{-2s+2}). \tag{9}$$

We now determine the sign pattern of M^{-1}. For $1 \leqslant h \leqslant d-1$, put $\lambda_h := h^2$. Recall that the rows of M are indexed by $r = 2, \ldots, d$, while the columns are indexed by $h = 1, \ldots, d-1$, and

$$M_{r,h} = h^{2r} = \lambda_h^r.$$

Equivalently, after relabeling the row $r = q+1$, where $1 \leqslant q \leqslant d-1$, the corresponding entry is

$$h^{2r} = h^{2(q+1)} = \lambda_h^{q+1} = \lambda_h^{q-1}\lambda_h^2.$$

Thus, after this relabeling of the rows,

$$M = VD,$$

where $V = (\lambda_h^{q-1})_{1\leqslant q,h\leqslant d-1}$ and $D = \operatorname{diag}(\lambda_1^2, \ldots, \lambda_{d-1}^2)$.

Since D has positive diagonal entries, $M^{-1} = D^{-1}V^{-1}$ has the same sign pattern as V^{-1}. Write $V^{-1} = (\alpha_{h,q})_{1\leqslant h,q\leqslant d-1}$. The h-th Lagrange interpolation polynomial for the nodes $\lambda_1, \ldots, \lambda_{d-1}$ is

$$L_h(X) := \prod_{\substack{1\leqslant j\leqslant d-1\\ j\neq h}} \frac{X-\lambda_j}{\lambda_h-\lambda_j}.$$

Since the rows of V correspond to the monomials $1, X, \ldots, X^{d-2}$, we may write

$$L_h(X) = \alpha_{h,1} + \alpha_{h,2}X + \cdots + \alpha_{h,d-1}X^{d-2}.$$

Thus $\alpha_{h,q}$ is the coefficient of X^{q-1} in $L_h(X)$. The coefficient of X^{q-1} in

$$\prod_{\substack{1\leqslant j\leqslant d-1\\ j\neq h}} (X-\lambda_j)$$

has sign $(-1)^{d-1-q}$, because all λ_j are positive. On the other hand,

$$\operatorname{sgn}\left(\prod_{j\neq h}(\lambda_h-\lambda_j)\right) = (-1)^{d-1-h}.$$

Therefore

$$\operatorname{sgn}(\alpha_{h,q}) = (-1)^{d-1-q}(-1)^{d-1-h} = (-1)^{h+q}.$$

It follows that

$$\operatorname{sgn}((M^{-1})_{h,q}) = \operatorname{sgn}((V^{-1})_{h,q}) = (-1)^{h+q}.$$

Applying this with $q = s-1$ in (9), we get, for all sufficiently small θ,

$$\operatorname{sgn}(b_h(\theta)) = \operatorname{sgn}(a_{2s})(-1)^{s+1}(-1)^{h+s-1} = \operatorname{sgn}(a_{2s})(-1)^h.$$

Thus the nonzero coefficients $b_1(\theta), \ldots, b_{d-1}(\theta)$ have alternating signs.

If $p \in [4k+2, 4k+4)$, then $d = 2k+1$, so there are $d-1 = 2k$ oscillatory coefficients, and an alternating sign pattern gives at most k positive ones. If $p = 4k+4$, then $d = 2k+2$, $s = d$, and $a_{2s} = a_{2d} = 1$. Hence positive signs can occur only for even h, namely among $h = 2, 4, \ldots, 2k$, so again there are at most k positive coefficients.

Finally, choose $\theta_0 = \theta_0(p, A) > 0$ sufficiently small so that both the error estimate above and the sign conclusions hold for every $0 < \theta < \theta_0$. This proves the theorem. □

2.4.2 Proofs of remaining lemmas

We finish this subsection by proving the three elementary algebraic lemmas used in the rank argument.

Proof of Lemma 2.3. Let

$$M := \big(K(x_i, x_j)\big)_{i,j=1}^m.$$

For any $\boldsymbol{v} \in \mathbb{R}^m$, set $\boldsymbol{u} := C_m\boldsymbol{v}$. Since C_m is the orthogonal projection onto

$$\mathbf{1}^\perp := \{\boldsymbol{w} \in \mathbb{R}^m : \langle \boldsymbol{w}, \mathbf{1}\rangle = 0\},$$

we have $\boldsymbol{u} \in \mathbf{1}^\perp$. Therefore, by the conditional negative definiteness of K,

$$\boldsymbol{v}^\top(-C_mMC_m)\boldsymbol{v} = -\boldsymbol{u}^\top M\boldsymbol{u} \geqslant 0.$$

Thus $-C_mMC_m$ is positive semidefinite. □

Proof of Lemma 2.4. We show that G is positive definite on the $(m-1)$-dimensional subspace $\mathbf{1}^\perp$. Let $\boldsymbol{v} \in \mathbf{1}^\perp\backslash\{\mathbf{0}\}$. Since C_m is the orthogonal projection onto $\mathbf{1}^\perp$, we have $C_m\boldsymbol{v} = \boldsymbol{v}$. Hence

$$\boldsymbol{v}^\top G\boldsymbol{v} = \boldsymbol{v}^\top NN^\top\boldsymbol{v} + \boldsymbol{v}^\top C_m\boldsymbol{v} + \boldsymbol{v}^\top E\boldsymbol{v} = \|N^\top\boldsymbol{v}\|_2^2 + \|\boldsymbol{v}\|_2^2 + \boldsymbol{v}^\top E\boldsymbol{v}.$$

Since $\lambda_{\min}(E) > -1$, we have

$$\boldsymbol{v}^\top E\boldsymbol{v} \geqslant \lambda_{\min}(E)\|\boldsymbol{v}\|_2^2.$$

It follows that

$$\boldsymbol{v}^\top G\boldsymbol{v} \geqslant \|N^\top\boldsymbol{v}\|_2^2 + (1+\lambda_{\min}(E))\|\boldsymbol{v}\|_2^2 > 0.$$

Thus G is positive definite on $\mathbf{1}^\perp$. Consequently, G has rank at least $\dim \mathbf{1}^\perp = m-1$, as desired. □

Proof of Lemma 2.5. Let

$$H := \big(\cos(r\theta(u_i - u_j))\big)_{i,j=1}^m.$$

Since $D = J_m - H$ and $C_mJ_m = J_mC_m = 0$, we have

$$-C_mDC_m = C_mHC_m.$$

It is therefore enough to show that H is positive semidefinite and has rank at most 2.

Using $\cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$, we can write

$$H = \boldsymbol{c}\boldsymbol{c}^\top + \boldsymbol{s}\boldsymbol{s}^\top,$$

where $\boldsymbol{c} := (\cos(r\theta u_1), \ldots, \cos(r\theta u_m))^\top$ and $\boldsymbol{s} := (\sin(r\theta u_1), \ldots, \sin(r\theta u_m))^\top$. Hence H is positive semidefinite and $\operatorname{rank}(H) \leqslant 2$. Since C_mHC_m is a compression of H, it is also positive semidefinite and has rank at most 2. Therefore $-C_mDC_m$ is positive semidefinite and has rank at most 2. □

2.5 Affine independence and nonregular examples for $2 < p < 4$

We close this section with two simple observations about the equality case in the range $2 < p < 4$. The first one says that every largest equilateral set must be affinely independent. Thus, in this range, the extremal configuration is always an n-simplex in the affine sense.

Proposition 2.10. *Let $2 < p < 4$, and let $A = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_m\} \subseteq \mathbb{R}^n$ be an equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$. If $m = n+1$, then A is affinely independent. Equivalently, A is the vertex set of an n-simplex.*

Proof of Proposition 2.10. In the proof of Theorem 1.2, when $2 < p < 4$, we have $k = 0$. Thus there are no oscillatory correction terms and no error term. The centered identity reduces to

$$2b_0\widetilde{G}_2 = -C_m M_K C_m + C_m,$$

where $\widetilde{G}_2 := -\frac{1}{2}C_m D_2 C_m$ is the centered Gram matrix of $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$, and M_K is the matrix associated with the conditionally negative definite kernel K.

By Lemma 2.3, the matrix $-C_m M_K C_m$ is positive semidefinite. Hence we may write

$$-C_m M_K C_m = NN^{\mathsf{T}}$$

for some real matrix N. Therefore

$$2b_0\widetilde{G}_2 = NN^{\mathsf{T}} + C_m.$$

By Lemma 2.4, applied with $E = 0$, we get $\operatorname{rank}(2b_0\widetilde{G}_2) \geqslant m-1$. Since $m = n+1 \geqslant 2$, this also implies $b_0 \neq 0$, and hence

$$\operatorname{rank}(\widetilde{G}_2) \geqslant m-1.$$

On the other hand, $\widetilde{G}_2$ is the Gram matrix of m centered vectors in $\mathbb{R}^n$, so $\operatorname{rank}(\widetilde{G}_2) \leqslant n$. Since $m = n+1$, we have

$$\operatorname{rank}(\widetilde{G}_2) = n = m-1.$$

But the rank of the centered Gram matrix is exactly the dimension of the affine span of A. Hence the affine span of A has dimension $m-1$, which means that $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ are affinely independent. $\square$

The preceding proposition should not be confused with Euclidean regularity. It only says that a largest equilateral set is an affine simplex. In general, this simplex need not be regular with respect to the Euclidean metric.

Proposition 2.11. *Let $2 < p < 4$ and $n \geqslant 2$. Then there exists an equilateral set of size $n+1$ in $(\mathbb{R}^n, \|\cdot\|_p)$ which is not a regular simplex in the Euclidean sense.*

Proof of Proposition 2.11. Let $\boldsymbol{x}_0 := \frac{1}{2}\boldsymbol{e}_1$ and $\boldsymbol{x}_1 := -\frac{1}{2}\boldsymbol{e}_1$. Next, let $\alpha > 0$ be the unique real number satisfying

$$\big(\alpha + 2^{-1/p}\big)^p + (n-2)\alpha^p = 1 - 2^{-p}. \tag{10}$$

Such an α exists and is unique, because the function

$$h(t) := \big(t + 2^{-1/p}\big)^p + (n-2)t^p$$

is continuous and strictly increasing when $t \geqslant 0$, with $h(0) = 2^{-1} < 1 - 2^{-p}$, since $p > 1$, and $h(t) \to \infty$ as $t \to \infty$. For each $i = 1, \dots, n-1$, define

$$\boldsymbol{x}_{i+1} := \alpha \sum_{j=2}^{n} \boldsymbol{e}_j + 2^{-1/p}\boldsymbol{e}_{i+1}.$$

We claim that $A := \{\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$ is equilateral in $(\mathbb{R}^n, \|\cdot\|_p)$ with common distance 1. First, $\|\boldsymbol{x}_0 - \boldsymbol{x}_1\|_p^p = 1$. Next, if $i \neq j$ with $1 \leqslant i, j \leqslant n-1$, then $\boldsymbol{x}_{i+1} - \boldsymbol{x}_{j+1}$ has exactly two nonzero coordinates, namely $2^{-1/p}$ and $-2^{-1/p}$. Hence

$$\|\boldsymbol{x}_{i+1} - \boldsymbol{x}_{j+1}\|_p^p = 2 \cdot 2^{-1} = 1.$$

Finally, for each $i = 1, \dots, n-1$,

$$\boldsymbol{x}_0 - \boldsymbol{x}_{i+1} = \frac{1}{2}\boldsymbol{e}_1 - \alpha \sum_{j=2}^{n} \boldsymbol{e}_j - 2^{-1/p}\boldsymbol{e}_{i+1},$$

and hence, by (10),

$$\|\boldsymbol{x}_0 - \boldsymbol{x}_{i+1}\|_p^p = 2^{-p} + \big(\alpha + 2^{-1/p}\big)^p + (n-2)\alpha^p = 1.$$

The same computation applies to $\boldsymbol{x}_1$, and gives $\|\boldsymbol{x}_1 - \boldsymbol{x}_{i+1}\|_p^p = 1$. Therefore A is equilateral in $(\mathbb{R}^n, \|\cdot\|_p)$. It remains to show that A is not a regular simplex in the Euclidean sense. We have $\|\boldsymbol{x}_0 - \boldsymbol{x}_1\|_2^2 = 1$.

If $n \geqslant 3$, then for distinct $i, j \in \{1, 2, \dots, n-1\}$,

$$\|\boldsymbol{x}_{i+1} - \boldsymbol{x}_{j+1}\|_2^2 = 2 \cdot 2^{-2/p} = 2^{1-2/p} \neq 1.$$

Thus A is not Euclidean regular when $n \geqslant 3$. It remains to consider the case $n = 2$. In this case (10) gives $\alpha + 2^{-1/p} = (1 - 2^{-p})^{1/p}$. Therefore

$$\|\boldsymbol{x}_0 - \boldsymbol{x}_2\|_2^2 = \frac{1}{4} + \big(\alpha + 2^{-1/p}\big)^2 = \frac{1}{4} + (1 - 2^{-p})^{2/p}.$$

Since $p > 2$, we have $2/p < 1$, and hence $(1 - 2^{-p})^{2/p} > 1 - 2^{-p} > \frac{3}{4}$. It follows that

$$\|\boldsymbol{x}_0 - \boldsymbol{x}_2\|_2^2 > 1 = \|\boldsymbol{x}_0 - \boldsymbol{x}_1\|_2^2.$$

Thus A is not Euclidean regular also when $n = 2$. This completes the proof. $\square$

3 An almost linear bound for $e(\ell_p^n)$ when $p \in (4k, 4k+2)$

The proof of Theorem 1.3 requires a different argument from that used in the preceding section. We begin with a brief sketch, highlighting where the previous rank argument breaks down and explaining the mechanism that replaces it.

3.1 Sketch of the Proof of Theorem 1.3

Let $A \subseteq \mathbb{R}^n$ be an equilateral set with common ℓ_p-distance 1, where $p \in (4k, 4k+2)$. After translating A, we may assume that one point of A is the origin. This normalization places all coordinates in a bounded interval and gives a uniform coordinatewise p-mass bound, which will later make the approximation errors add up correctly over the n coordinates.

The finite-rank argument from the preceding section cannot be closed in the same way in the range $p \in (4k, 4k+2)$. The obstruction is a parity-driven sign reversal at the last level of the moment filtration. In the interior of the range treated in the preceding section, namely $p \in (4k+2, 4k+4)$, one has $d = \lfloor p/2 \rfloor = 2k+1$, and Lemma 2.8 gives the favorable positive definiteness of $Q_{|t|^p}$ on V_d. This sign was the key input which allowed the signed kernel $-|x-y|^p$ to be corrected, up to a controlled error term, by a quadratic term and finitely many oscillatory kernels. After double centering, the conditionally negative contribution lay on the same side as the p-distance term and could be absorbed into the positive-semidefinite side, leaving only a controlled number of unfavorable finite-rank modes to be counted.

In the complementary range $p \in (4k, 4k+2)$, however, we have $d = \lfloor p/2 \rfloor = 2k$, and the corresponding quadratic form $Q_{|t|^p}$ is negative definite on V_d rather than positive definite. Thus the sign needed for the preceding absorption mechanism is reversed. The previous finite-rank correction scheme therefore no longer produces a conditionally negative definite correction of the signed kernel $-|x-y|^p$ in a way compatible with the same rank count. This reversal suggests the appropriate replacement: the kernel to be corrected is now $|x-y|^p$ itself, rather than $-|x-y|^p$.

The replacement is a zero-anchored Hilbert-space representation. Write $p = 4k+\beta$, with $0 < \beta < 2$. Since the kernel $|u-v|^\beta$ is conditionally negative definite on the real line, it admits a Hilbert-space realization by squared distances. We use this conditionally negative kernel after taking $2k$-fold zero-anchored primitives. Integrating by parts then reduces the exponent p to the conditionally negative exponent β. This gives a one-dimensional representation in which the main sign-definite part of the p-distance quadratic form becomes a squared norm in a Hilbert space, up to finite-dimensional boundary correction terms. The anchoring at the origin is essential: it ensures that the boundary terms produced by the integrations remain finite-dimensional.

Applying this representation coordinate by coordinate, and then performing the centering step, gives the analogue of the matrix identity used in the preceding section. The centered simplex matrix is balanced against two contributions: a finite-dimensional correction term of rank at most $4kn$, coming from the boundary terms, and a Gram matrix coming from the Hilbert-space component.

The crucial difference is where this Hilbert-space term appears in the rank identity. In the previous argument, the conditionally negative contribution was on the same side as the p-distance term. Hence, after double centering, it could be absorbed into the positive-semidefinite side, while the rank estimate was applied to the opposite side. In the present range, the sign is reversed: the Hilbert-space Gram term appears on the side to which the rank estimate must be applied. Thus it is no longer harmless. Although its Gram matrix has finite rank for the finite set under consideration, that rank may be as large as the size of the set itself, so inserting it directly would destroy the rank count.

The rest of the proof is therefore an approximation argument. The resulting Hilbert map has a scaling structure, which allows its image to be approximated by dyadic shells. Using L shells gives an $O_p(L)$-dimensional approximation in one coordinate, and hence an $O_p(nL)$-dimensional approximation after taking the direct sum over all coordinates. A coordinatewise p-mass estimate controls the accumulation of the approximation errors over the n coordinates. Choosing $L = O_p(\log n)$, and then applying a suitable projection lower bound to the resulting centered Gram identity, yields $|A| \leqslant C_p n \log(2n)$.

3.2 A zero-anchored primitive and an exact bridge

The following zero-anchored kernel may be viewed as a compactly supported distributional $2k$-fold primitive of the point mass δ_x, with all lower-order boundary terms anchored at the origin. This anchoring is what later produces only finitely many correction terms.

Lemma 3.1. *Let $p \in (4k, 4k+2)$ for some integer $k \geqslant 1$, and set $\beta := p - 4k \in (0,2)$ and $n_0 := 2k-1$. For each $x \in [-1,1]$, define*

$$\psi_x^{(0)}(t) := \begin{cases} \dfrac{(x-t)_+^{n_0}}{n_0!} - \displaystyle\sum_{r=0}^{n_0} \frac{x^r}{r!}\,\frac{(-t)_+^{n_0-r}}{(n_0-r)!}, & x \geqslant 0, \\[2ex] \dfrac{(t-x)_+^{n_0}}{n_0!} - \displaystyle\sum_{r=0}^{n_0} \frac{(-x)^r}{r!}\,\frac{t_+^{n_0-r}}{(n_0-r)!}, & x \leqslant 0, \end{cases}$$

where $u_+^s := u^s \mathbf{1}_{\{u>0\}}$ for $u \in \mathbb{R}$ and $s \in \mathbb{N}_0$. Then the following hold.

(i) $\operatorname{supp} \psi_x^{(0)} \subseteq [\min\{x,0\}, \max\{x,0\}]$.

(ii) *In the sense of distributions,*

$$\big(\psi_x^{(0)}\big)^{(2k)} = \delta_x - \sum_{r=0}^{2k-1} \frac{(-1)^r x^r}{r!}\,\delta_0^{(r)}.$$

(iii) *For every $\alpha > -1$ and every integer $r \in \{0,1,\ldots,2k-1\}$,*

$$\int_{\mathbb{R}} \psi_x^{(0)}(t)\ \mathrm{sgn}(t)^r |t|^\alpha\, dt = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+2k+1)}\,\mathrm{sgn}(x)^r |x|^{\alpha+2k}.$$

In particular, $\int_{\mathbb{R}} \psi_x^{(0)}(t)\,dt = \frac{|x|^{2k}}{(2k)!}$, *and*

$$\int_{\mathbb{R}} \psi_x^{(0)}(t)\,|t|^\beta\, dt = \frac{\Gamma(\beta+1)}{\Gamma(p-2k+1)}|x|^{p-2k}.$$

Proof of Lemma 3.1. Suppose first that $x \geqslant 0$. For $t < 0$, we have

$$(x-t)^{n_0} = \sum_{r=0}^{n_0} \binom{n_0}{r} x^r (-t)^{n_0-r},$$

hence the two terms in the definition of $\psi_x^{(0)}$ cancel identically on $(-\infty, 0)$. Also, for $t > x$, the first term vanishes and the second term is already zero. Therefore $\mathrm{supp}\,\psi_x^{(0)} \subseteq [0,x]$. Suppose now that $x \leqslant 0$. For $t > 0$, we have

$$(t-x)^{n_0} = \sum_{r=0}^{n_0} \binom{n_0}{r} (-x)^r t^{n_0-r},$$

so the two terms in the definition of $\psi_x^{(0)}$ cancel identically on $(0,\infty)$. Also, for $t < x$, the first term vanishes and the second term is already zero. Hence $\mathrm{supp}\,\psi_x^{(0)} \subseteq [x,0]$. This proves (i).

We next prove (ii). If $x \geqslant 0$, then $\frac{d^{2k}}{dt^{2k}} \frac{(x-t)_+^{n_0}}{n_0!} = \delta_x$. Also for $0 \leqslant r \leqslant n_0$, we have

$$\frac{d^{2k}}{dt^{2k}} \left(\frac{(-t)_+^{n_0-r}}{(n_0-r)!} \right) = (-1)^r \delta_0^{(r)}.$$

Therefore

$$\big(\psi_x^{(0)}\big)^{(2k)} = \delta_x - \sum_{r=0}^{2k-1} \frac{(-1)^r x^r}{r!}\delta_0^{(r)}.$$

If $x \leqslant 0$, then $\frac{d^{2k}}{dt^{2k}} \frac{(t-x)_+^{n_0}}{n_0!} = \delta_x$, and $\frac{d^{2k}}{dt^{2k}} \left(\frac{t_+^{n_0-r}}{(n_0-r)!} \right) = \delta_0^{(r)}$. Since $(-x)^r = (-1)^r x^r$, we again obtain

$$\big(\psi_x^{(0)}\big)^{(2k)} = \delta_x - \sum_{r=0}^{2k-1} \frac{(-1)^r x^r}{r!}\delta_0^{(r)}.$$

This proves (ii).

Finally, we prove (iii). If $x = 0$, then $\psi_0^{(0)} \equiv 0$, and the claim is trivial. Assume first that $x > 0$. By (i), $\psi_x^{(0)}$ is supported on $[0,x]$, where $\mathrm{sgn}(t)^r = 1$. Thus

$$\int_{\mathbb{R}} \psi_x^{(0)}(t)\ \mathrm{sgn}(t)^r |t|^\alpha\, dt = \int_0^x \frac{(x-t)^{2k-1}}{(2k-1)!}\, t^\alpha\, dt.$$

After the change of variables $t = xu$, this becomes

$$\frac{x^{\alpha+2k}}{(2k-1)!}\int_0^1 (1-u)^{2k-1}u^\alpha\,du = \frac{x^{\alpha+2k}}{(2k-1)!}B(\alpha+1,2k) = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+2k+1)}x^{\alpha+2k}.$$

If $x < 0$, then $\psi_x^{(0)}$ is supported on $[x, 0]$, where $\operatorname{sgn}(t)^r = \operatorname{sgn}(x)^r$. Hence

$$\int_{\mathbb{R}} \psi_x^{(0)}(t)\,\operatorname{sgn}(t)^r |t|^\alpha\,dt = \operatorname{sgn}(x)^r \int_x^0 \frac{(t-x)^{2k-1}}{(2k-1)!}\,|t|^\alpha\,dt.$$

Writing $x = -a$ with $a > 0$, and changing variables $t = -au$, we obtain

$$\operatorname{sgn}(x)^r \frac{a^{\alpha+2k}}{(2k-1)!}\int_0^1 (1-u)^{2k-1}u^\alpha\,du = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+2k+1)}\operatorname{sgn}(x)^r|x|^{\alpha+2k}.$$

This proves (iii). □

We now package the one-dimensional identity into a Hilbert-space bridge. The only Hilbert-space input used below is the standard construction associated with a positive semidefinite symmetric bilinear form: one quotients out its null space and completes the resulting inner product space. In the resulting real Hilbert space, finite Gram matrices are positive semidefinite and orthogonal projections onto finite-dimensional subspaces are available.

For $0 < \beta < 2$, the kernel $|x-y|^\beta$ is conditionally negative definite and hence admits a Hilbert-space realization as squared distances, or equivalently as squared norms of zero-mass linear combinations. We apply this realization after the zero-anchored $2k$-fold primitive construction; this is why the resulting feature map has homogeneity $\frac{p}{2} = 2k + \frac{\beta}{2}$.

Proposition 3.2. *Let $p \in (4k, 4k+2)$ for some integer $k \geqslant 0$, and set $\beta := p - 4k \in (0, 2)$. Then there exist:*

- *a real Hilbert space $\mathcal{H}_\beta$,*
- *a continuous map $\Phi_0 : [-1, 1] \to \mathcal{H}_\beta$,*
- *a map $\boldsymbol{w}_0 : [-1, 1] \to \mathbb{R}^{4k}$,*
- *symmetric positive semidefinite matrices $M_{+,0}, M_{-,0} \in \mathbb{R}^{4k\times 4k}$,*
- *bounded linear operators $T_a : \mathcal{H}_\beta \to \mathcal{H}_\beta$ for $0 < a \leqslant 1$,*

such that the following hold.

(i) *For every finite set $X = \{x_1, \ldots, x_m\} \subseteq [-1, 1]$ and every $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbf{1}^\perp$,*

$$-\frac{1}{2}\sum_{i,j=1}^m \lambda_i\lambda_j|x_i - x_j|^p = \eta_p\left\|\sum_{i=1}^m \lambda_i\Phi_0(x_i)\right\|_{\mathcal{H}_\beta}^2 + \left\|\sum_{i=1}^m \lambda_i M_{+,0}^{1/2}\boldsymbol{w}_0(x_i)\right\|_2^2 - \left\|\sum_{i=1}^m \lambda_i M_{-,0}^{1/2}\boldsymbol{w}_0(x_i)\right\|_2^2,$$

where $\eta_p := \frac{\Gamma(p+1)}{\Gamma(\beta+1)}$ and $\Gamma(s) := \int_0^\infty t^{s-1}e^{-t}\,dt$ for $s > 0$ is the Euler Gamma function.

(ii) *For all $0 < a \leqslant 1$ and all $x \in [-1, 1]$, $\Phi_0(ax) = T_a\Phi_0(x)$, and for $\xi \in \mathcal{H}_\beta$,*

$$\|T_a\xi\|_{\mathcal{H}_\beta} = a^{p/2}\|\xi\|_{\mathcal{H}_\beta}.$$

Proof of Proposition 3.2. If $k = 0$, then $0 < p < 2$. By Schoenberg's theorem [18]; see also [15, Lemma 7] for this exact formulation, the kernel $|x-y|^p$ is conditionally negative definite on $\mathbb{R}$. The Hilbert space representation theorem for conditionally negative definite kernels [5, Proposition 3.2] yields a real Hilbert space $\mathcal{H}_p$ and a continuous map $\Phi_0 : \mathbb{R} \to \mathcal{H}_p$ such that

$$|x-y|^p = \|\Phi_0(x) - \Phi_0(y)\|^2_{\mathcal{H}_p}$$

for $x, y \in \mathbb{R}$. Hence, for every $\boldsymbol{\lambda} \in \mathbf{1}^\perp$,

$$\left\| \sum_i \lambda_i \Phi_0(x_i) \right\|^2_{\mathcal{H}_p} = -\frac{1}{2} \sum_{i,j} \lambda_i \lambda_j \|\Phi_0(x_i) - \Phi_0(x_j)\|^2_{\mathcal{H}_p} = -\frac{1}{2} \sum_{i,j} \lambda_i \lambda_j |x_i - x_j|^p.$$

We take $\boldsymbol{w}_0$ to be the map into $\mathbb{R}^0$, and $M_{+,0}, M_{-,0}$ to be the corresponding 0×0 matrices. After translating the embedding, we may assume that $\Phi_0(0) = 0$, and we replace $\mathcal{H}_p$ by the closed linear span of $\Phi_0(\mathbb{R})$. By polarization and the homogeneity of $|x-y|^p$, the assignment $T_a \Phi_0(x) := \Phi_0(ax)$ extends linearly to $\operatorname{span} \Phi_0(\mathbb{R})$, is well-defined, and satisfies $\|T_a \xi\|_{\mathcal{H}_\beta} = a^{p/2} \|\xi\|_{\mathcal{H}_\beta}$. It therefore extends to a bounded linear operator on $\mathcal{H}_\beta$, still denoted by T_a. Thus the proposition holds for $k = 0$.

Assume now that $k \geqslant 1$. We use signed measures only as a compact notation for combining ordinary integrals and point masses: $f(t)\,dt$ contributes $\int \varphi(t) f(t)\,dt$, while δ_y contributes $\varphi(y)$. Let

$$\mathcal{M}_0 := \Big\{ \mu : \mu \text{ is a compactly supported signed Borel measure on } \mathbb{R},\ \mu(\mathbb{R}) = 0 \Big\}.$$

Define a symmetric bilinear form on $\mathcal{M}_0$ by

$$\langle \mu, \nu \rangle_\beta := -\frac{1}{2} \iint_{\mathbb{R}^2} |u - v|^\beta \, d\mu(u)\, d\nu(v).$$

Since $0 < \beta < 2$, the kernel $(u, v) \to |u - v|^\beta$ is conditionally negative definite on $\mathbb{R}$, and therefore $\langle \cdot, \cdot \rangle_\beta$ is positive semidefinite on $\mathcal{M}_0$. Quotienting by the null space and completing, we obtain a real Hilbert space $\mathcal{H}_\beta$.

For $x \in [-1, 1]$, let $\psi_x^{(0)}$ be as in Lemma 3.1, and define

$$\nu_x^{(0)} := \psi_x^{(0)}(t)\,dt - \frac{|x|^{2k}}{(2k)!} \delta_0.$$

By Lemma 3.1 (i) and (iii), $\nu_x^{(0)} \in \mathcal{M}_0$. Define $\Phi_0(x) := [\nu_x^{(0)}] \in \mathcal{H}_\beta$. Now let $X = \{x_1, \ldots, x_m\} \subseteq [-1, 1]$, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbf{1}^\perp$, and define

$$Q_{p,X}(\boldsymbol{\lambda}) := -\frac{1}{2} \sum_{i,j=1}^m \lambda_i \lambda_j |x_i - x_j|^p.$$

Set $F_\lambda^{(0)}(t) := \sum_{i=1}^m \lambda_i \psi_{x_i}^{(0)}(t)$, $\mu_\lambda := \sum_{i=1}^m \lambda_i \delta_{x_i}$, $M_r(\boldsymbol{\lambda}) := \sum_{i=1}^m \lambda_i x_i^r$ for $0 \leqslant r \leqslant 2k$,

$$S_r(\boldsymbol{\lambda}) := \sum_{i=1}^m \lambda_i \operatorname{sgn}(x_i)^r |x_i|^{p-r}$$

for $1 \leqslant r \leqslant 2k$. By Lemma 3.1(ii),

$$(F_\lambda^{(0)})^{(2k)} = \mu_\lambda - \sum_{r=0}^{2k-1} \frac{(-1)^r M_r(\boldsymbol{\lambda})}{r!} \delta_0^{(r)}.$$

Since $\boldsymbol{\lambda} \in \mathbf{1}^{\perp}$, we have $M_0(\boldsymbol{\lambda}) = 0$, and therefore

$$(F_{\lambda}^{(0)})^{(2k)} = \mu_{\lambda} - \sum_{r=1}^{2k-1} \frac{(-1)^r M_r(\boldsymbol{\lambda})}{r!} \delta_0^{(r)}. \tag{11}$$

Define also

$$\nu_{\lambda}^{(0)} := F_{\lambda}^{(0)}(t)\, dt - \frac{M_{2k}(\boldsymbol{\lambda})}{(2k)!} \delta_0.$$

By Lemma 3.1 (i) and (iii), $\nu_{\lambda}^{(0)} \in \mathcal{M}_0$, and $\sum_{i=1}^{m} \lambda_i \Phi_0(x_i) = [\nu_{\lambda}^{(0)}]$. Let $\| \cdot \|_{\mathcal{H}_\beta}$ denote the norm in the Hilbert space $\mathcal{H}_\beta$, induced by the inner product

$$\langle \mu, \nu \rangle_\beta := -\frac{1}{2} \iint\limits_{\mathbb{R}^2} |u - v|^\beta \, d\mu(u)\, d\nu(v).$$

Since $\sum_{i=1}^{m} \lambda_i \Phi_0(x_i) = [\nu_{\lambda}^{(0)}] \in \mathcal{H}_\beta$, we have, by the definition of this Hilbert-space norm,

$$\left\| \sum_{i=1}^{m} \lambda_i \Phi_0(x_i) \right\|_{\mathcal{H}_\beta}^2 = -\frac{1}{2} \iint |u - v|^\beta \, d\nu_{\lambda}^{(0)}(u)\, d\nu_{\lambda}^{(0)}(v). \tag{12}$$

We now compute $Q_{p,X}(\boldsymbol{\lambda})$ by separating the bulk part from the boundary terms at the origin. Write $D_\lambda := (F_{\lambda}^{(0)})^{(2k)}$, and $B_\lambda := \sum_{r=1}^{2k-1} \frac{(-1)^r M_r(\boldsymbol{\lambda})}{r!} \delta_0^{(r)}$. Then (11) says precisely that $\mu_\lambda = D_\lambda + B_\lambda$. Thus,

$$Q_{p,X}(\boldsymbol{\lambda}) = -\frac{1}{2} \left\langle (D_\lambda + B_\lambda) \otimes (D_\lambda + B_\lambda), |x - y|^p \right\rangle.$$

Here $S \otimes T$ is understood in the standard distributional sense, with T acting first in the y-variable and then S in the x-variable. We decompose this expression as

$$Q_{p,X}(\boldsymbol{\lambda}) = Q^{\text{bulk}} + Q^{\text{cross}} + Q^{\text{bdry}}, \tag{13}$$

where

$$Q^{\text{bulk}} := -\frac{1}{2} \langle D_\lambda \otimes D_\lambda, |x - y|^p \rangle, \tag{14}$$

$$Q^{\text{cross}} := -\frac{1}{2} \left\langle D_\lambda \otimes B_\lambda + B_\lambda \otimes D_\lambda, |x - y|^p \right\rangle, \tag{15}$$

$$Q^{\text{bdry}} := -\frac{1}{2} \langle B_\lambda \otimes B_\lambda, |x - y|^p \rangle. \tag{16}$$

The following three paragraphs compute these three terms.

Bulk–bulk term Q^{bulk}. Since

$$\partial_x^{2k} \partial_y^{2k} \left(-\frac{1}{2} |x - y|^p \right) = \eta_p \left(-\frac{1}{2} |x - y|^\beta \right),$$

by (14), integration by parts yields

$$Q^{\text{bulk}} = -\frac{\eta_p}{2} \iint\limits_{\mathbb{R}^2} F_{\lambda}^{(0)}(u) F_{\lambda}^{(0)}(v) |u - v|^\beta \, du\, dv.$$

On the other hand, Lemma 3.1 (iii) gives

$$\int_{\mathbb{R}} F_\lambda^{(0)}(t)\,dt = \frac{M_{2k}(\boldsymbol{\lambda})}{(2k)!}$$

and

$$\int_{\mathbb{R}} |t|^\beta F_\lambda^{(0)}(t)\,dt = \frac{\Gamma(\beta+1)}{\Gamma(p-2k+1)}\,S_{2k}(\boldsymbol{\lambda}).$$

Expanding (12), and using $\eta_p := \frac{\Gamma(p+1)}{\Gamma(\beta+1)}$, we obtain

$$Q^{\text{bulk}} = \eta_p \left\| \sum_{i=1}^m \lambda_i \Phi_0(x_i) \right\|_{\mathcal{H}_\beta}^2 - \binom{p}{2k} M_{2k}(\boldsymbol{\lambda}) S_{2k}(\boldsymbol{\lambda}). \tag{17}$$

Bulk–boundary term Q^{cross}. For $1 \leqslant r \leqslant 2k-1$, define $h_r(x) := \operatorname{sgn}(x)^r |x|^{p-r}$. We also write $(p)_r := p(p-1)\cdots(p-r+1)$ for the falling factorial. Moreover, we use $\langle \delta_0^{(r)}, \cdot \rangle_y$ to denote that the distribution $\delta_0^{(r)}$ acts on the y-variable, with x fixed. Therefore, for $x \neq 0$,

$$\left\langle \delta_0^{(r)},\, |x-y|^p \right\rangle_y = (-1)^r \partial_y^r |x-y|^p\big|_{y=0} = (p)_r h_r(x).$$

The same identity holds at $x=0$ by continuous extension, as $p-r>0$. Since $|x-y|^p$ is symmetric, the two cross terms in (15) are equal. Hence

$$Q^{\text{cross}} = -\langle D_\lambda \otimes B_\lambda, |x-y|^p \rangle = -\sum_{r=1}^{2k-1} \frac{(-1)^r M_r(\boldsymbol{\lambda})}{r!} \left\langle (F_\lambda^{(0)})^{(2k)},\, (p)_r h_r \right\rangle.$$

Notice that $p-r>2k$, so $h_r \in C^{2k}(\mathbb{R})$, and

$$h_r^{(2k)}(x) = (p-r)_{2k}\ \operatorname{sgn}(x)^r |x|^{p-r-2k}.$$

Hence, by integration by parts and Lemma 3.1(iii) with $\alpha = p-r-2k > -1$,

$$\left\langle (F_\lambda^{(0)})^{(2k)},\, h_r \right\rangle = S_r(\boldsymbol{\lambda}).$$

Since $\frac{(p)_r}{r!} = \binom{p}{r}$, summing over $r = 1,\ldots,2k-1$ gives

$$Q^{\text{cross}} = -\sum_{r=1}^{2k-1} (-1)^r \binom{p}{r} M_r(\boldsymbol{\lambda}) S_r(\boldsymbol{\lambda}). \tag{18}$$

Boundary–boundary term Q^{bdry}. Since $B_\lambda := \sum_{r=1}^{2k-1} \frac{(-1)^r M_r(\boldsymbol{\lambda})}{r!} \delta_0^{(r)}$, by (16), every summand is a constant multiple of

$$\left\langle \delta_0^{(r)} \otimes \delta_0^{(s)},\, |x-y|^p \right\rangle = (-1)^{r+s} \partial_x^r \partial_y^s |x-y|^p\big|_{(0,0)},$$

with $1 \leqslant r,s \leqslant 2k-1$. Since $r+s \leqslant 4k-2 < p$, this derivative exists and equals 0 at $(0,0)$. Hence

$$Q^{\text{bdry}} = 0. \tag{19}$$

By (13), (17), (18), and (19), we conclude that

$$Q_{p,X}(\boldsymbol{\lambda}) = \eta_p \left\| \sum_{i=1}^m \lambda_i \Phi_0(x_i) \right\|_{\mathcal{H}_\beta}^2 - \sum_{r=1}^{2k} (-1)^r \binom{p}{r} M_r(\boldsymbol{\lambda}) S_r(\boldsymbol{\lambda}). \tag{20}$$

We now encode the finite-dimensional correction. Define $u_r(x) := x^r$ for $1 \leqslant r \leqslant 2k$, $v_r(x) := \operatorname{sgn}(x)^r |x|^{p-r}$ for $1 \leqslant r \leqslant 2k$. Set

$$\boldsymbol{w}_0(x) := \big(u_1(x), \dots, u_{2k}(x), v_1(x), \dots, v_{2k}(x)\big) \in \mathbb{R}^{4k}.$$

Define a symmetric matrix $A_{p,k} \in \mathbb{R}^{4k \times 4k}$ by

$$A_{p,k}(r, 2k+r) = A_{p,k}(2k+r, r) = -\frac{1}{2}(-1)^r \binom{p}{r}$$

for $1 \leqslant r \leqslant 2k$, and all other entries equal to 0. Notice that if $\mathbf{m}_{w_0}(\boldsymbol{\lambda}) := \sum_{i=1}^m \lambda_i \boldsymbol{w}_0(x_i)$, then by construction

$$\mathbf{m}_{w_0}(\boldsymbol{\lambda})^\top A_{p,k}\, \mathbf{m}_{w_0}(\boldsymbol{\lambda}) = -\sum_{r=1}^{2k} (-1)^r \binom{p}{r} M_r(\boldsymbol{\lambda}) S_r(\boldsymbol{\lambda}).$$

Thus (20) becomes

$$Q_{p,X}(\boldsymbol{\lambda}) = \eta_p \left\| \sum_{i=1}^m \lambda_i \Phi_0(x_i) \right\|_{\mathcal{H}_\beta}^2 + \mathbf{m}_{w_0}(\boldsymbol{\lambda})^\top A_{p,k}\, \mathbf{m}_{w_0}(\boldsymbol{\lambda}).$$

Take a spectral decomposition $A_{p,k} = M_{+,0} - M_{-,0}$ with $M_{+,0}, M_{-,0} \succeq 0$. Then

$$Q_{p,X}(\boldsymbol{\lambda}) = \eta_p \left\| \sum_{i=1}^m \lambda_i \Phi_0(x_i) \right\|_{\mathcal{H}_\beta}^2 + \left\| \sum_{i=1}^m \lambda_i M_{+,0}^{1/2} \boldsymbol{w}_0(x_i) \right\|_2^2 - \left\| \sum_{i=1}^m \lambda_i M_{-,0}^{1/2} \boldsymbol{w}_0(x_i) \right\|_2^2.$$

This proves (i).

We now prove (ii). We first define the scaling operation at the level of measures. For $0 < a \leqslant 1$ and $\mu \in \mathcal{M}_0$, let $\widetilde{T}_a \mu$ be the signed measure determined by

$$\int_{\mathbb{R}} \varphi(t)\, d(\widetilde{T}_a\mu)(t) = a^{2k} \int_{\mathbb{R}} \varphi(au)\, d\mu(u)$$

for every test function φ. In words, we scale the underlying variable by $u \mapsto au$, and multiply the mass by a^{2k}. This operation preserves compact support and total mass zero, so $\widetilde{T}_a \mu \in \mathcal{M}_0$.

For $\mu, \nu \in \mathcal{M}_0$, the definition gives

$$\begin{aligned}
\langle \widetilde{T}_a\mu, \widetilde{T}_a\nu \rangle_\beta &= -\frac{1}{2} \iint\limits_{\mathbb{R}^2} |u-v|^\beta \, d(\widetilde{T}_a\mu)(u)\, d(\widetilde{T}_a\nu)(v) \\
&= a^{4k} \left(-\frac{1}{2} \iint\limits_{\mathbb{R}^2} |au - av|^\beta \, d\mu(u)\, d\nu(v) \right) \\
&= a^{4k+\beta} \langle \mu, \nu \rangle_\beta = a^p \langle \mu, \nu \rangle_\beta.
\end{aligned}$$

In particular, if μ belongs to the null space of the semidefinite form $\langle \cdot, \cdot \rangle_\beta$, then so does $\widetilde{T}_a\mu$. Hence $\widetilde{T}_a$ descends to the quotient space. On the quotient, and then on its Hilbert-space completion, we denote the induced operator by T_a. The identity above gives $\|T_a \xi\|_{\mathcal{H}_\beta} = a^{p/2} \|\xi\|_{\mathcal{H}_\beta}$ for every $\xi \in \mathcal{H}_\beta$.

It remains to verify that this operator has the required action on the vectors $\Phi_0(x)$. By direct inspection of the definition of $\psi_x^{(0)}$, for every $x \in [-1,1]$ and $0 < a \leqslant 1$,

$$\psi_{ax}^{(0)}(t) = a^{2k-1} \psi_x^{(0)}(t/a).$$

Equivalently, for every test function φ,

$$\int_{\mathbb{R}} \varphi(t)\psi_{ax}^{(0)}(t)\,dt = a^{2k}\int_{\mathbb{R}} \varphi(au)\psi_x^{(0)}(u)\,du.$$

Moreover, $\frac{|ax|^{2k}}{(2k)!}\delta_0$ is obtained from $\frac{|x|^{2k}}{(2k)!}\delta_0$ by the same scaling rule, because the mass is multiplied by a^{2k} and the point 0 remains fixed. Therefore $\nu_{ax}^{(0)} = \widetilde{T}_a\nu_x^{(0)}$. Passing to equivalence classes in $\mathcal{H}_\beta$, we obtain $\Phi_0(ax) = T_a\Phi_0(x)$.

Finally, we prove that Φ_0 is continuous. Using the identity established in (i) with the two points x, y and coefficients $(1, -1)$, we obtain

$$|x-y|^p = \eta_p\|\Phi_0(x) - \Phi_0(y)\|_{\mathcal{H}_\beta}^2 + \|M_{+,0}^{1/2}(\boldsymbol{w}_0(x) - \boldsymbol{w}_0(y))\|_2^2 - \|M_{-,0}^{1/2}(\boldsymbol{w}_0(x) - \boldsymbol{w}_0(y))\|_2^2.$$

Thus

$$\eta_p\|\Phi_0(x) - \Phi_0(y)\|_{\mathcal{H}_\beta}^2 \leqslant |x-y|^p + \left|\|M_{+,0}^{1/2}(\boldsymbol{w}_0(x) - \boldsymbol{w}_0(y))\|_2^2 - \|M_{-,0}^{1/2}(\boldsymbol{w}_0(x) - \boldsymbol{w}_0(y))\|_2^2\right|.$$

Since $\boldsymbol{w}_0$ is continuous on $[-1,1]$ and $M_{+,0}, M_{-,0}$ are fixed finite-dimensional matrices, the finite-dimensional correction tends to 0 as $y \to x$. Hence

$$\|\Phi_0(x) - \Phi_0(y)\|_{\mathcal{H}_\beta} \to 0$$

as $y \to x$. This completes the proof. □

3.3 Projection and dyadic shell approximation

In what follows, we assume that $p \in (4k, 4k+2)$ for some integer $k \geqslant 0$. Set $\beta := p - 4k \in (0,2)$ and $\eta_p := \frac{\Gamma(p+1)}{\Gamma(\beta+1)}$, where $\Gamma(s) := \int_0^\infty t^{s-1}e^{-t}\,dt$ for $s > 0$ is the Euler Gamma function. We also fix once and for all one choice of the objects $\mathcal{H}_\beta$, Φ_0, $\boldsymbol{w}_0$, $M_{+,0}$, $M_{-,0}$, and T_a for $0 < a \leqslant 1$ provided by Proposition 3.2.

The following proposition gives a lower bound on the maximal squared error after projection onto an s-dimensional subspace.

Proposition 3.3. *Let $B = \{x_1, \ldots, x_m\}$ be a finite set, let $\mathcal{H}$ be a real Hilbert space, and let $\phi : B \to \mathcal{H}$ be a map. Define $R := \big(\langle\phi(x_i), \phi(x_j)\rangle_{\mathcal{H}}\big)_{i,j=1}^m$, and $C_m := I_m - \frac{1}{m}J_m$. Assume that $a > 0$ and that there exists an $m \times m$ matrix X with rank r such that $\frac{1}{a}C_m + X = R$. Then for every orthogonal projection $P : \mathcal{H} \to \mathcal{S}$ onto an s-dimensional subspace $\mathcal{S} \subseteq \mathcal{H}$, one has*

$$\max_{x \in B}\|\phi(x) - P\phi(x)\|_{\mathcal{H}}^2 \geqslant \frac{m-r-1-s}{am}. \tag{21}$$

Proof of Proposition 3.3. Let $Q := I_{\mathcal{H}} - P$, so that Q is the orthogonal projection onto $\mathcal{S}^\perp$. For each $i \in \{1, 2, \ldots, m\}$, write $\phi_i := \phi(x_i)$. Define $R_P := \big(\langle P\phi_i, P\phi_j\rangle_{\mathcal{H}}\big)_{i,j=1}^m$, and $R_Q := \big(\langle Q\phi_i, Q\phi_j\rangle_{\mathcal{H}}\big)_{i,j=1}^m$. Since P and Q are complementary orthogonal projections, we have $R = R_P + R_Q$, therefore, $R_Q = \frac{1}{a}C_m + X - R_P$. Next, we set $M := R_P - X$. Then

$$\operatorname{rank}(M) \leqslant \operatorname{rank}(R_P) + \operatorname{rank}(X) \leqslant s + r.$$

Consider the subspace

$$U := \ker(M) \cap \mathbf{1}^\perp.$$

Since $\dim(\mathbf{1}^\perp) = m-1$ and $\dim\ker(M) = m - \operatorname{rank}(M) \geqslant m - (s+r)$, we obtain

$$\dim U \geqslant m - 1 - (s+r).$$

If $\boldsymbol{u} \in U$, then $M\boldsymbol{u} = 0$ and $\boldsymbol{u} \in \mathbf{1}^{\perp}$, so $C_m \boldsymbol{u} = \boldsymbol{u}$. Therefore

$$R_Q \boldsymbol{u} = \left(\frac{1}{a} C_m - M \right) \boldsymbol{u} = \frac{1}{a} \boldsymbol{u}.$$

Thus every vector in U is an eigenvector of R_Q with eigenvalue $\frac{1}{a}$. Since R_Q is a Gram matrix, it is positive semidefinite. Hence

$$\operatorname{Tr}(R_Q) \geqslant \frac{\dim U}{a} \geqslant \frac{m-1-(s+r)}{a}.$$

On the other hand, $\operatorname{Tr}(R_Q) = \sum_{i=1}^{m} \|Q\phi_i\|_{\mathcal{H}}^2$. Therefore

$$\max_{1 \leqslant i \leqslant m} \|Q\phi_i\|_{\mathcal{H}}^2 \geqslant \frac{1}{m} \sum_{i=1}^{m} \|Q\phi_i\|_{\mathcal{H}}^2 = \frac{\operatorname{tr}(R_Q)}{m} \geqslant \frac{m-r-1-s}{am}.$$

Since $Q = I_{\mathcal{H}} - P$, this is exactly (21). $\square$

The next lemma approximates the fixed map Φ_0 on dyadic shells. Its proof uses only the continuity of Φ_0 and the scaling property in Proposition 3.2 (ii).

Lemma 3.4. *Fix $\delta > 0$. Then for every integer $L \geqslant 1$, there exists a subspace $\mathcal{V}_L \subseteq \mathcal{H}_\beta$ such that $\dim(\mathcal{V}_L) \leqslant C_{p,\delta} L$ and for all $x \in [-1,1]$,*

$$\operatorname{dist}\big(\Phi_0(x), \mathcal{V}_L\big)^2 \leqslant 2^p \delta^2 |x|^p + C_p 2^{-Lp}, \tag{22}$$

where $\operatorname{dist}(\xi, V) := \inf_{z \in V} \|\xi - z\|_{\mathcal{H}_\beta}$.

Proof of Lemma 3.4. Consider the compact intervals $I_+ := [1/2, 1]$ and $I_- := [-1, -1/2]$. Since Φ_0 is continuous, the sets $\Phi_0(I_+)$ and $\Phi_0(I_-)$ are compact in $\mathcal{H}_\beta$. Hence there exist finite sets $U_+ = \{u_1^+, u_2^+, \dots, u_{N_+}^+\} \subseteq I_+$ and $U_- = \{u_1^-, u_2^-, \dots, u_{N_-}^-\} \subseteq I_-$ such that $\operatorname{dist}\big(\Phi_0(x), V_+\big) \leqslant \delta$ for all $x \in I_+$, and $\operatorname{dist}\big(\Phi_0(x), V_-\big) \leqslant \delta$ for all $x \in I_-$, where $V_+ := \operatorname{span}\{\Phi_0(u_1^+), \Phi_0(u_2^+), \dots, \Phi_0(u_{N_+}^+)\}$ and $V_- := \operatorname{span}\{\Phi_0(u_1^-), \Phi_0(u_2^-), \dots, \Phi_0(u_{N_-}^-)\}$.

For each $j = 0, 1, \dots, L-1$, define $V_{j,+} := T_{2^{-j}} V_+$ and $V_{j,-} := T_{2^{-j}} V_-$. Now set

$$\mathcal{V}_L := \sum_{j=0}^{L-1} (V_{j,+} + V_{j,-}).$$

Then $\dim(\mathcal{V}_L) \leqslant (N_+ + N_-)L =: C_{p,\delta} L$ for some $C_{p,\delta} > 0$. Let $x \in [-1,1]$. If $x \in (2^{-j-1}, 2^{-j}]$ for some $0 \leqslant j \leqslant L-1$, write $x = 2^{-j}u$, $u \in [1/2, 1]$. Choose $v \in V_+$ with $\|\Phi_0(u) - v\|_{\mathcal{H}_\beta} \leqslant \delta$. Then $T_{2^{-j}} v \in V_{j,+} \subseteq \mathcal{V}_L$, and by the Proposition 3.2 (ii),

$$\operatorname{dist}\big(\Phi_0(x), \mathcal{V}_L\big)^2 \leqslant \|T_{2^{-j}}(\Phi_0(u) - v)\|_{\mathcal{H}_\beta}^2 = 2^{-jp} \|\Phi_0(u) - v\|_{\mathcal{H}_\beta}^2 \leqslant 2^{-jp} \delta^2.$$

Since $x \in (2^{-j-1}, 2^{-j}]$, we have $2^{-jp} \leqslant 2^p |x|^p$, and therefore

$$\operatorname{dist}\big(\Phi_0(x), \mathcal{V}_L\big)^2 \leqslant 2^p \delta^2 |x|^p.$$

If $x \in [-2^{-j}, -2^{-j-1})$ for some $0 \leqslant j \leqslant L-1$, write $x = 2^{-j}u$, $u \in [-1, -1/2]$. Choose $v \in V_-$ with $\|\Phi_0(u) - v\|_{\mathcal{H}_\beta} \leqslant \delta$. Then $T_{2^{-j}} v \in V_{j,-} \subseteq \mathcal{V}_L$, and similarly

$$\operatorname{dist}\big(\Phi_0(x), \mathcal{V}_L\big)^2 \leqslant 2^{-jp} \delta^2 \leqslant 2^p \delta^2 |x|^p.$$

Finally, if $|x| \leqslant 2^{-L}$, then $0 \in \mathcal{V}_L$, so

$$\operatorname{dist}\big(\Phi_0(x), \mathcal{V}_L\big)^2 \leqslant \|\Phi_0(x)\|_{\mathcal{H}_\beta}^2.$$

If $x \neq 0$, write $x = |x|\sigma$ with $\sigma \in \{-1, 1\}$. Then

$$\|\Phi_0(x)\|_{\mathcal{H}_\beta} = \|T_{|x|}\Phi_0(\sigma)\|_{\mathcal{H}_\beta} = |x|^{p/2}\|\Phi_0(\sigma)\|_{\mathcal{H}_\beta}.$$

Therefore

$$\|\Phi_0(x)\|_{\mathcal{H}_\beta}^2 \leqslant C_p|x|^p \leqslant C_p 2^{-Lp}.$$

The same bound holds for $x = 0$, since $\Phi_0(0) = 0$ by Proposition 3.2 (ii). This proves (22). $\square$

Lemma 3.5. *Let $A = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ be an equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$, and normalize the common distance to be* 1. *After translation, assume that $\boldsymbol{x}_1 = \mathbf{0}$. For $s \in [-1, 1]$, define*

$$\Theta(s) := \big(\sqrt{\eta_p}\,\Phi_0(s),\ M_{+,0}^{1/2}\boldsymbol{w}_0(s)\big) \in \mathcal{H}_\beta \oplus \mathbb{R}^{4k},$$

and

$$U(s) := M_{-,0}^{1/2}\boldsymbol{w}_0(s) \in \mathbb{R}^{4k}.$$

For each $\boldsymbol{x}_i = (x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(n)}) \in A$, set $\Theta_i := \bigoplus_{t=1}^n \Theta(x_i^{(t)})$ and $U_i := \bigoplus_{t=1}^n U(x_i^{(t)})$. Define $R := C_m\big(\langle\Theta_i, \Theta_j\rangle\big)_{i,j=1}^m C_m$ and $X := C_m\big(\langle U_i, U_j\rangle\big)_{i,j=1}^m C_m$ with $C_m := I_m - \frac{1}{m}J_m$. Then X is positive semidefinite, $\operatorname{rank}(X) \leqslant 4kn$, *and $R = \frac{1}{2}C_m + X$.*

Proof of Lemma 3.5. Let $G_\Theta := \big(\langle\Theta_i, \Theta_j\rangle\big)_{i,j=1}^m$ and $G_U := \big(\langle U_i, U_j\rangle\big)_{i,j=1}^m$. Then $R = C_m G_\Theta C_m$ and $X = C_m G_U C_m$. For every $\boldsymbol{a} \in \mathbb{R}^m$,

$$\boldsymbol{a}^\top X\boldsymbol{a} = (C_m\boldsymbol{a})^\top G_U(C_m\boldsymbol{a}) = \left\|\sum_{i=1}^m (C_m\boldsymbol{a})_i U_i\right\|^2 \geqslant 0.$$

Thus $X \succeq 0$. Also,

$$\operatorname{rank}(X) \leqslant \operatorname{rank}(G_U) \leqslant \dim\big((\mathbb{R}^{4k})^{\oplus n}\big) = 4kn.$$

It remains to prove the identity $R = \frac{1}{2}C_m + X$. Let $\boldsymbol{a} \in \mathbb{R}^m$ and set $\boldsymbol{\lambda} := C_m\boldsymbol{a}$. Then $\boldsymbol{\lambda} \in \mathbf{1}^\perp$. Applying Proposition 3.2 (i) coordinatewise and summing over $t = 1, 2, \ldots, n$, we get

$$\boldsymbol{\lambda}^\top(G_\Theta - G_U)\boldsymbol{\lambda} = -\frac{1}{2}\sum_{i,j=1}^m \lambda_i\lambda_j\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_p^p.$$

Since A is equilateral with common distance 1, and since $\sum_i \lambda_i = 0$,

$$-\frac{1}{2}\sum_{i,j=1}^m \lambda_i\lambda_j\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_p^p = \frac{1}{2}\|\boldsymbol{\lambda}\|_2^2.$$

Therefore

$$\boldsymbol{a}^\top(R - X)\boldsymbol{a} = (C_m\boldsymbol{a})^\top(G_\Theta - G_U)(C_m\boldsymbol{a}) = \frac{1}{2}\|C_m\boldsymbol{a}\|_2^2 = \boldsymbol{a}^\top\left(\frac{1}{2}C_m\right)\boldsymbol{a}.$$

Since this holds for every $\boldsymbol{a} \in \mathbb{R}^m$, we conclude that $R - X = \frac{1}{2}C_m$, as desired. $\square$

3.4 Proof of the logarithmic upper bound

We are now ready to prove Theorem 1.3.

Proof of Theorem 1.3. Let $A = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ be an equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$. By scaling, we may assume that the common distance is 1. After translation, we may assume that $\boldsymbol{x}_1 = \boldsymbol{0}$. Then, for every $\boldsymbol{x}_i = (x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(n)}) \in A$, we have

$$\sum_{t=1}^{n} |x_i^{(t)}|^p = \|\boldsymbol{x}_i\|_p^p \leqslant 1. \tag{23}$$

Define Θ_i, U_i, R, X as in Lemma 3.5. By that lemma, $R = \frac{1}{2}C_m + X$ with $X \succeq 0$ and

$$\operatorname{rank}(X) \leqslant 4kn. \tag{24}$$

Fix $\delta > 0$ so that

$$\eta_p 2^p \delta^2 \leqslant \frac{1}{8}. \tag{25}$$

For an integer $L \geqslant 1$, let $\mathcal{V}_L$ be given by Lemma 3.4, and define

$$Z_0(L) := \mathcal{V}_L \oplus M_{+,0}^{1/2}\mathbb{R}^{4k} \subseteq \mathcal{H}_\beta \oplus \mathbb{R}^{4k}.$$

Since $L \geqslant 1$, we have $\dim Z_0(L) \leqslant C_p' L$ for some constant $C_p' > 0$. Set $Z(L) := \bigoplus_{t=1}^{n} Z_0(L)$. Then

$$\dim Z(L) \leqslant C_p' nL. \tag{26}$$

Fix $i \in [m]$. Since $M_{+,0}^{1/2}\mathbb{R}^{4k} \subseteq Z_0(L)$, only the Hilbert component of each coordinate block contributes to the distance from Θ_i to $Z(L)$. Hence Lemma 3.4 gives

$$\begin{aligned}
\operatorname{dist}(\Theta_i, Z(L))^2 &\leqslant \sum_{t=1}^{n} \eta_p \operatorname{dist}\big(\Phi_0(x_i^{(t)}), \mathcal{V}_L\big)^2 \\
&\leqslant \sum_{t=1}^{n} \Big(\eta_p 2^p \delta^2 |x_i^{(t)}|^p + \eta_p C_p 2^{-Lp}\Big).
\end{aligned}$$

Using (23) and (25), we obtain

$$\operatorname{dist}(\Theta_i, Z(L))^2 \leqslant \frac{1}{8} + C_p'' n 2^{-Lp}$$

for some constant $C_p'' > 0$. Choose $L := \max\left\{1, \left\lceil \frac{1}{p}\log_2(8C_p'' n)\right\rceil\right\}$. Then $C_p'' n 2^{-Lp} \leqslant \frac{1}{8}$, and therefore

$$\max_{1 \leqslant i \leqslant m} \operatorname{dist}(\Theta_i, Z(L))^2 \leqslant \frac{1}{4}. \tag{27}$$

Let $\overline{\Theta} = \frac{1}{m}\sum_{j=1}^{m} \Theta_j$, and define $W := \operatorname{span}\big(\overline{\Theta},\, Z(L)\big)$. Then $\dim W \leqslant \dim Z(L) + 1$. Moreover, for each i,

$$\operatorname{dist}\big(\Theta_i - \overline{\Theta},\, W\big) \leqslant \operatorname{dist}(\Theta_i, Z(L)).$$

Indeed, if $z \in Z(L)$, then $z - \overline{\Theta} \in W$, and $\big(\Theta_i - \overline{\Theta}\big) - \big(z - \overline{\Theta}\big) = \Theta_i - z$.

By the definition of R in Lemma 3.5, R is the Gram matrix of the centered family $\big\{\Theta_i - \overline{\Theta}\big\}_{i=1}^{m}$. We now apply Proposition 3.3 to this family, with $a = 2$,$r = 4kn$, and $s = \dim W$. Using (24), we get

$$\max_{1 \leqslant i \leqslant m} \operatorname{dist}\left(\Theta_i - \frac{1}{m}\sum_{j=1}^{m}\Theta_j,\, W\right)^2 \geqslant \frac{m - 4kn - 1 - \dim W}{2m}.$$

Combining this with (27), we obtain

$$\frac{m - 4kn - 1 - \dim W}{2m} \leqslant \frac{1}{4}.$$

Since $\dim W \leqslant \dim Z(L) + 1$, it follows that

$$\frac{m - 4kn - \dim Z(L) - 2}{2m} \leqslant \frac{1}{4}.$$

Therefore $m \leqslant 2\big(4kn + \dim Z(L) + 2\big)$. Using (26), we conclude that

$$m \leqslant 2\big(4kn + C'_p nL + 2\big).$$

Finally, by the choice of L, $L \leqslant C_p \log(2n)$ for some constant $C_p > 0$. Hence

$$m \leqslant C_p n \log(2n),$$

after enlarging C_p if necessary. This completes the proof. □

4 Equilateral sets in $(\mathbb{T}^n, d_p)$: improved upper bounds

In this section, we prove Theorem 1.5. Compared with the Euclidean problem, the toroidal problem has an additional obstruction: in one dimension the cyclic distance $d_{\mathbb{T}}(x, y) = \min\{|x-y|, 1-|x-y|\}$ is only piecewise a real-line distance. It switches branches at the antipodal point $y = x + 1/2$, where $d_{\mathbb{T}}(x, y)^p$ has a corner. This loss of smoothness produces an additional wrap-around contribution in the associated kernel, so the sign-controlled kernel decompositions used in $\mathbb{R}^n$ do not apply directly. Our proof first shows that the small-distance case reduces to the Euclidean setting, so that it remains to treat equilateral sets with large common distance. We then reduce the argument to a one-dimensional estimate for the anchored torus kernel. This kernel is split into a real-line part and a genuinely toroidal correction, the latter capturing precisely the corner contribution near the antipodal point. The real-line part is controlled by the Hilbert-space approximation developed earlier, while the toroidal correction is reduced to one-sided corner kernels and estimated by low-rank approximation with small positive trace error. After summing over the coordinates, this gives the required rank comparison and proves the theorem.

4.1 Equilateral sets in $(\mathbb{R}^n, \|\cdot\|_p)$ and $(\mathbb{T}^n, d_p)$

We begin with a simple embedding observation showing that the torus problem is at least as hard as the corresponding Euclidean problem. We also explain why on the torus it is enough to treat the regime of large common distance.

Lemma 4.1. *Let $n \geqslant 1$ be an integer. For $1 \leqslant p < \infty$, the following statements hold:*

(i) *If $A \subseteq (\mathbb{T}^n, d_p)$ is an equilateral set with common distance $\lambda < 1/3$, then A is isometric to an equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$.*

(ii) *If $B \subseteq (\mathbb{R}^n, \|\cdot\|_p)$ is an equilateral set with common distance $\lambda \leqslant 1/2$, then B embeds isometrically into $(\mathbb{T}^n, d_p)$.*

Proof. For (i), translate A so that $\boldsymbol{x}^* = (1/2, \dots, 1/2) \in A$. Let $\boldsymbol{x} = (x_1, \dots, x_n) \in A \backslash \{\boldsymbol{x}^*\}$. Since $d_{\mathbb{T}}(x_i, 1/2) \leqslant d_p(\boldsymbol{x}, \boldsymbol{x}^*) = \lambda < 1/3$ for every $i \in [n]$, we have $x_i \in (1/6, 5/6)$. Now let $\boldsymbol{x}, \boldsymbol{y} \in A$ be distinct. Then for every $i \in [n]$, $1 - |x_i - y_i| > 1/3$, whereas $d_{\mathbb{T}}(x_i, y_i) \leqslant d_p(\boldsymbol{x}, \boldsymbol{y}) < 1/3$. Thus the second branch in the definition of $d_{\mathbb{T}}$ cannot occur, and so $d_{\mathbb{T}}(x_i, y_i) = |x_i - y_i|$ for every $i \in [n]$. Therefore A, viewed inside the cube $(1/6, 5/6)^n \subseteq \mathbb{R}^n$, has the same pairwise distances as in $(\mathbb{T}^n, d_p)$. This proves (i).

For (ii), let $\pi : \mathbb{R}^n \to \mathbb{T}^n$ be the coordinatewise quotient map. Let $\boldsymbol{x}, \boldsymbol{y} \in B$ be distinct. Since $|x_i - y_i| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|_p = \lambda \leqslant 1/2$ for every $i \in [n]$, we have $d_{\mathbb{T}}(\pi(x_i), \pi(y_i)) = |x_i - y_i|$. It follows that $d_p(\pi(\boldsymbol{x}), \pi(\boldsymbol{y})) = \|\boldsymbol{x} - \boldsymbol{y}\|_p$. Thus $\pi|_B$ is an isometric embedding of B into $(\mathbb{T}^n, d_p)$. $\square$

As an immediate consequence of Lemma 4.1 (ii), every equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$ can be rescaled and embedded isometrically into $(\mathbb{T}^n, d_p)$ with the same cardinality. Hence

$$e(\ell_p^n) \leqslant e_p(\mathbb{T}^n).$$

Lemma 4.1 also shows that the case of small common distance can be reduced to the case where the common distance is $1/3$. Indeed, let $A \subseteq (\mathbb{T}^n, d_p)$ be an equilateral set with common distance $0 < \lambda < 1/3$. By Lemma 4.1 (i), A is isometric to an equilateral set $B \subseteq (\mathbb{R}^n, \|\cdot\|_p)$ with the same common distance λ. Rescaling B by the factor $(3\lambda)^{-1}$, we obtain an equilateral set in $(\mathbb{R}^n, \|\cdot\|_p)$ with common distance $1/3$. By Lemma 4.1 (ii), this rescaled set embeds isometrically into $(\mathbb{T}^n, d_p)$. Thus A has the same cardinality as an equilateral set in $(\mathbb{T}^n, d_p)$ with common distance $1/3$. Consequently, in proving upper bounds for equilateral sets in $(\mathbb{T}^n, d_p)$, we may and will assume that $\lambda \geqslant \frac{1}{3}$.

4.2 Reduction to a one-dimensional statement

We keep the notation $d_{\mathbb{T}}$ and d_p from the Introduction. Throughout this section we use the centred realization of the circle. Equivalently, identifying $\mathbb{T}$ with $[-1/2, 1/2)$, we shall use the 1-periodic function $\rho(s) := \min_{m \in \mathbb{Z}} |s - m|$ for $s \in \mathbb{R}$. Thus $\rho(s) = |s|$ for $s \in [-1/2, 1/2)$, and, by periodicity, we also regard ρ as a function on $\mathbb{T}$. Hence, for $x, y \in \mathbb{T}$,

$$d_{\mathbb{T}}(x, y) = \rho(x - y).$$

Consequently, for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{T}^n$ and $1 \leqslant p < \infty$,

$$d_p(\boldsymbol{x}, \boldsymbol{y})^p = \sum_{j=1}^{n} \rho(x_j - y_j)^p.$$

For $1 \leqslant p < \infty$ and $x, y \in \mathbb{T}$, define the anchored kernel

$$B_p(x, y) := \frac{1}{2}\big(\rho(x)^p + \rho(y)^p - \rho(x - y)^p\big).$$

For a real symmetric matrix M, we write

$$\operatorname{tr}_+(M) := \sum_{\lambda_i(M) > 0} \lambda_i(M),$$

where eigenvalues are counted with multiplicity. We also set

$$\vartheta_p := \begin{cases} 0, & 1 \leqslant p \leqslant 2, \\ \dfrac{p-2}{3p-2}, & p > 2. \end{cases}$$

Note that for $p > 2$,

$$\frac{1}{1 - \vartheta_p} = \frac{3}{2} - \frac{1}{p}.$$

The key auxiliary tool is the following one-dimensional proposition.

Proposition 4.2. *Let $p \geqslant 1$. Let $X = \{x_1, \ldots, x_m\} \subseteq \mathbb{T}$ and let $N \geqslant 2$. Define $M_{ij} := B_p(x_i, x_j)$ for $1 \leqslant i, j \leqslant m$. For every $\varepsilon > 0$, there exists $C_{p,\varepsilon} > 0$ such that there are real symmetric matrices $L, R \in \mathbb{R}^{m \times m}$ satisfying*

(i) $M \preceq L + R$.

(ii) $\operatorname{rank} L \leqslant C_{p,\varepsilon}\big(m^{\vartheta_p} + \log(2m) + \log(2N)\big)$.

(iii) $\operatorname{tr}_+(R) \leqslant \varepsilon \sum_{i=1}^m \rho(x_i)^p + C_{p,\varepsilon}\frac{m}{N}$.

We first show how Proposition 4.2 implies Theorem 1.5.

Proof of Theorem 1.5 assuming Proposition 4.2. Let

$$A = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\} \subseteq (\mathbb{T}^n, d_p)$$

be an equilateral set with common distance λ. By translation, assume $\boldsymbol{x}_0 = \mathbf{0}$.

As explained at the end of Subsection 4.1, we may assume that $\lambda \geqslant \frac{1}{3}$.

For each coordinate $t \in \{1, 2, \ldots, n\}$, put $(B_t)_{ij} := B_p(x_i^{(t)}, x_j^{(t)})$ for $1 \leqslant i, j \leqslant m$. Since $d_p(\boldsymbol{x}_i, \mathbf{0}) = \lambda$ and $d_p(\boldsymbol{x}_i, \boldsymbol{x}_j) = \lambda$ for $i \neq j$, we get

$$\sum_{t=1}^n B_t = \frac{\lambda^p}{2}(I_m + J_m). \tag{28}$$

Fix $\varepsilon_* > 0$, to be chosen later. Choose $N := \lceil kn \rceil + 2$, where $k = k(p, \varepsilon_*)$ is sufficiently large that $C_{p,\varepsilon_*}\frac{n}{N} \leqslant \varepsilon_* 4^{-p}$. Apply Proposition 4.2 to each B_t. Thus there are symmetric matrices $L_t, R_t \in \mathbb{R}^{m \times m}$ such that $B_t \preceq L_t + R_t$, $\operatorname{rank} L_t \leqslant C_{p,\varepsilon_*}\big(m^{\vartheta_p} + \log(2m) + \log(2N)\big)$, and

$$\operatorname{tr}_+(R_t) \leqslant \varepsilon_* \sum_{i=1}^m \rho(x_i^{(t)})^p + C_{p,\varepsilon_*}\frac{m}{N}.$$

Set $L_B := \sum_{t=1}^n L_t$ and $R_B := \sum_{t=1}^n R_t$. Then, by (28),

$$\frac{\lambda^p}{2}(I_m + J_m) \preceq L_B + R_B.$$

Moreover, by subadditivity of tr_+,

$$\operatorname{tr}_+(R_B) \leqslant \varepsilon_* \sum_{i=1}^m \sum_{t=1}^n \rho(x_i^{(t)})^p + C_{p,\varepsilon_*}\frac{mn}{N} \leqslant 2\varepsilon_* m\lambda^p,$$

where we used $\lambda \geqslant 1/3$ and the choice of N.

Let $s = \operatorname{rank} L_B$, and let P be the orthogonal projection onto $\ker L_B$. Then $\operatorname{rank} P \geqslant m - s$, and since $PL_BP = 0$,

$$PR_BP = P(L_B + R_B)P \succeq P\frac{\lambda^p}{2}(I_m + J_m)P \succeq \frac{\lambda^p}{2}P.$$

For every real symmetric matrix R and every orthogonal projection P, one has

$$\operatorname{tr}_+(R) \geqslant \operatorname{tr}(PRP).$$

Therefore

$$2\varepsilon_* m\lambda^p \geqslant \operatorname{tr}_+(R_B) \geqslant \operatorname{tr}(PR_BP) \geqslant \frac{\lambda^p}{2}(m - s).$$

Taking $\varepsilon_* = 1/16$, we obtain $m - s \leqslant \frac{m}{4}$, and hence $m \leqslant 2s$.

On the other hand,

$$s = \operatorname{rank} L_B \leqslant nC_{p,\varepsilon_*}\big(m^{\vartheta_p} + \log(2m) + \log(2N)\big).$$

Thus

$$m \leqslant C_p n\big(m^{\vartheta_p} + \log(2m) + \log(2n)\big) \tag{29}$$

for some constant $C_p > 0$.

If $1 \leqslant p \leqslant 2$, then $\vartheta_p = 0$. Hence $m \leqslant C_p n \log(2m) + C_p n \log(2n)$. The elementary bootstrap estimate gives

$$m \leqslant C_p n \log(2n).$$

Now suppose $p > 2$. If $m \leqslant n^{1/(1-\vartheta_p)}$, then the desired bound already follows. Otherwise $m > n^{1/(1-\vartheta_p)}$. Since $\vartheta_p > 0$, we have

$$\log(2m) + \log(2n) \leqslant C_p m^{\vartheta_p}.$$

Thus (29) gives $m \leqslant C_p n m^{\vartheta_p}$. Consequently,

$$m \leqslant C_p n^{1/(1-\vartheta_p)} = C_p n^{3/2-1/p}.$$

Since $|A| = m+1$, the theorem follows after adjusting the constant C_p. □

4.3 Preliminaries for the one-dimensional proposition

We first record a linear algebra observation.

Lemma 4.3. *Let $A, D \in \mathbb{R}^{m\times m}$ be symmetric matrices and let $C_m := I_m - \frac{1}{m}J_m$. If $C_m A C_m \preceq C_m D C_m$, then there exists $\boldsymbol{b} \in \mathbb{R}^m$ such that*

$$A \preceq D + \mathbf{1}\boldsymbol{b}^{\mathsf{T}} + \boldsymbol{b}\mathbf{1}^{\mathsf{T}}.$$

Moreover, if $C_m A C_m = C_m D C_m$, then $\boldsymbol{b}$ can be chosen so that equality holds.

Proof of Lemma 4.3. Put $S := D - A$. The assumption says that $C_m S C_m \succeq 0$ on $\mathbf{1}^\perp$. Decompose

$$\mathbb{R}^m = \mathbf{1}^\perp \oplus \operatorname{span}\{\mathbf{1}\}.$$

The matrix $\mathbf{1}\boldsymbol{b}^{\mathsf{T}} + \boldsymbol{b}\mathbf{1}^{\mathsf{T}}$ does not change the restriction to $\mathbf{1}^\perp$. By choosing the component of $\boldsymbol{b}$ inside $\mathbf{1}^\perp$, we can cancel the off-diagonal block between $\mathbf{1}^\perp$ and $\operatorname{span}\{\mathbf{1}\}$. Then choosing the component of $\boldsymbol{b}$ in the direction of $\mathbf{1}$ sufficiently large makes the remaining scalar block nonnegative. Hence

$$S + \mathbf{1}\boldsymbol{b}^{\mathsf{T}} + \boldsymbol{b}\mathbf{1}^{\mathsf{T}} \succeq 0,$$

which is the desired domination. If $C_m S C_m = 0$, the same block computation allows us to choose $\boldsymbol{b}$ so that the resulting matrix is identically zero. □

Lemma 4.4. *Let $p \geqslant 1$. For every $\varepsilon > 0$, there exists $C_{p,\varepsilon} > 0$ such that the following holds. Let $x_1, \ldots, x_m \in [-1/2, 1/2]$, and let $N \geqslant 2$. Define*

$$(B_{\mathbb{R}})_{ij} := \frac{1}{2}\big(|x_i|^p + |x_j|^p - |x_i - x_j|^p\big).$$

Then there are symmetric matrices $L_{\mathbb{R}}, R_{\mathbb{R}} \in \mathbb{R}^{m\times m}$ such that

(i) $B_{\mathbb{R}} \preceq L_{\mathbb{R}} + R_{\mathbb{R}}$.

(ii) $\operatorname{rank} L_{\mathbb{R}} \leqslant C_{p,\varepsilon} \log(2N)$.

(iii) $R_{\mathbb{R}} \succeq 0$ *and* $\operatorname{tr}(R_{\mathbb{R}}) \leqslant \varepsilon \sum_{i=1}^m |x_i|^p + C_{p,\varepsilon} \frac{m}{N}$.

Proof of Lemma 4.4. If $p = 2h$ is an even integer, then $\frac{1}{2}\big(|x|^{2h} + |y|^{2h} - |x-y|^{2h}\big)$ is a polynomial in x and y of degree bounded in terms of p. Hence $B_{\mathbb{R}}$ has rank at most C_p. Taking $L_{\mathbb{R}} := B_{\mathbb{R}}$ and $R_{\mathbb{R}} := 0$ gives the desired conclusion. Therefore, it suffices to consider the case $p \in (2k, 2k+2)$.

First suppose $p \in (4q+2, 4q+4)$ for some integer $q \geqslant 0$. Set $d := \left\lfloor \frac{p}{2} \right\rfloor = 2q+1$. Let $A_0 := \{x_1, \ldots, x_m\} \subseteq \mathbb{R}$, where repetitions ignored. By Lemma 2.6, there are real constants $C_2(A_0,p), \ldots, C_{2d}(A_0,p)$ such that

$$K(x,y) := \sum_{r=1}^{d} C_{2r}(A_0,p)(x-y)^{2r} - |x-y|^p$$

is conditionally negative definite on A_0. Define $K_{ij} := K(x_i, x_j)$, and $P_{ij} := -\frac{1}{2}\sum_{r=1}^{d} C_{2r}(A_0,p)(x_i - x_j)^{2r}$. Since $C_m\mathbf{1} = 0$, we have

$$C_m B_{\mathbb{R}} C_m = \frac{1}{2} C_m K C_m + C_m P C_m.$$

By Lemma 2.3, $-C_m K C_m \succeq 0$, and hence

$$C_m B_{\mathbb{R}} C_m \preceq C_m P C_m.$$

Lemma 4.3 therefore gives a vector $\boldsymbol{b} \in \mathbb{R}^m$ such that

$$B_{\mathbb{R}} \preceq P + \mathbf{1}\boldsymbol{b}^{\mathsf{T}} + \boldsymbol{b}\mathbf{1}^{\mathsf{T}}.$$

The matrix P has rank at most $2d+1$, and consequently

$$\operatorname{rank}\big(P + \mathbf{1}\boldsymbol{b}^{\mathsf{T}} + \boldsymbol{b}\mathbf{1}^{\mathsf{T}}\big) \leqslant C_p.$$

Thus we may take $L_{\mathbb{R}} := P + \mathbf{1}\boldsymbol{b}^{\mathsf{T}} + \boldsymbol{b}\mathbf{1}^{\mathsf{T}}$ and $R_{\mathbb{R}} := 0$. This proves the lemma in this case.

It remains to consider the case $p \in (4q, 4q+2)$ for some integer $q \geqslant 0$, with $p \geqslant 1$. Let $\Phi_0, \boldsymbol{w}_0, M_{+,0}, M_{-,0}$ be the objects supplied by Proposition 3.2, and put

$$W := \begin{bmatrix} \boldsymbol{w}_0(x_1) & \boldsymbol{w}_0(x_2) & \cdots & \boldsymbol{w}_0(x_m) \end{bmatrix} \in \mathbb{R}^{4q \times m}.$$

For every $\boldsymbol{\lambda} \in \mathbf{1}^{\perp}$, Proposition 3.2 (i) gives

$$\boldsymbol{\lambda}^{\mathsf{T}} B_{\mathbb{R}} \boldsymbol{\lambda} = c_{p,q} \left\| \sum_{i=1}^m \lambda_i \Phi_0(x_i) \right\|_{\mathcal{H}_\beta}^2 + \boldsymbol{\lambda}^{\mathsf{T}} W^{\mathsf{T}} (M_{+,0} - M_{-,0}) W \boldsymbol{\lambda}.$$

By Lemma 4.3, in the equality case, there exists $\boldsymbol{b} \in \mathbb{R}^m$ such that

$$B_{\mathbb{R}} = c_{p,q}\big(\langle \Phi_0(x_i), \Phi_0(x_j)\rangle_{\mathcal{H}_\beta}\big)_{i,j=1}^m + W^{\mathsf{T}}(M_{+,0} - M_{-,0})W + \mathbf{1}\boldsymbol{b}^{\mathsf{T}} + \boldsymbol{b}\mathbf{1}^{\mathsf{T}}.$$

We approximate the Hilbert part by dyadic shells. By compactness of $\Phi_0([1/4, 1/2] \cup [-1/2, -1/4])$, there is a finite-dimensional subspace $V_0 \subseteq \mathcal{H}_\beta$ such that

$$\operatorname{dist}(\Phi_0(u), V_0)^2 \leqslant 4^{-p}\varepsilon$$

for every u with $1/4 \leqslant |u| \leqslant 1/2$. Let $L_0 := \left\lceil \frac{1}{p}\log_2 N \right\rceil + 3$ and $V_N := \sum_{\ell=0}^{L_0} T_{2^{-\ell}} V_0$. Then

$$\dim V_N \leqslant C_{p,\varepsilon} \log(2N).$$

Claim 4.5. $\operatorname{dist}(\Phi_0(x), V_N)^2 \leqslant \varepsilon|x|^p + C_{p,\varepsilon}N^{-1}$ *for* $|x| \leqslant \frac{1}{2}$.

Proof of the claim. If $2^{-\ell-2} < |x| \leqslant 2^{-\ell-1}$ for some $0 \leqslant \ell \leqslant L_0$, then $x = 2^{-\ell}u$ for some u with $1/4 < |u| \leqslant 1/2$. By Proposition 3.2 (ii),

$$\operatorname{dist}(\Phi_0(x), V_N)^2 \leqslant 2^{-\ell p} \operatorname{dist}(\Phi_0(u), V_0)^2 \leqslant 4^{-p}\varepsilon\, 2^{-\ell p} \leqslant \varepsilon|x|^p.$$

If $|x| \leqslant 2^{-L_0-2}$, then $0 \in V_N$, and By Proposition 3.2 (ii) also gives

$$\operatorname{dist}(\Phi_0(x), V_N)^2 \leqslant \|\Phi_0(x)\|^2_{\mathcal{H}_\beta} \leqslant C_p|x|^p \leqslant C_p 2^{-L_0 p} \leqslant C_p N^{-1}.$$

This proves the claim. ■

Let P_N be the orthogonal projection onto V_N. Define

$$(L_{\mathbb{R}})_{ij} := c_{p,q}\langle P_N\Phi_0(x_i), P_N\Phi_0(x_j)\rangle_{\mathcal{H}_\beta} + \big(W^\top(M_{+,0} - M_{-,0})W\big)_{ij} + b_i + b_j,$$

and set $R_{\mathbb{R}} := B_{\mathbb{R}} - L_{\mathbb{R}}$. Then

$$(R_{\mathbb{R}})_{ij} = c_{p,q}\langle (I - P_N)\Phi_0(x_i), (I - P_N)\Phi_0(x_j)\rangle_{\mathcal{H}_\beta},$$

so $R_{\mathbb{R}} \succeq 0$. Moreover,

$$\operatorname{rank} L_{\mathbb{R}} \leqslant \dim V_N + C_p \leqslant C_{p,\varepsilon}\log(2N),$$

and

$$\operatorname{tr}(R_{\mathbb{R}}) = c_{p,q}\sum_{i=1}^m \|(I - P_N)\Phi_0(x_i)\|^2_{\mathcal{H}_\beta} \leqslant \varepsilon\sum_{i=1}^m |x_i|^p + C_{p,\varepsilon}\frac{m}{N}.$$

This completes the proof. □

4.4 The toroidal correction

We next treat the genuinely toroidal correction.

Lemma 4.6. *Let $\psi : [0,1] \to \mathbb{R}$ be absolutely continuous, let $\psi(0) = 0$, and assume that ψ' has bounded variation on $[0,1]$. Then there exists a constant $C_\psi < \infty$ such that the following holds. For every $0 < \tau \leqslant 1$, every $u_1, \ldots, u_r, v_1, \ldots, v_s \in [0,\tau]$, and every integer $N \geqslant 1$, the matrix $M_{ij} := \psi((v_j - u_i)_+)$ admits a decomposition $M = E + H$ such that $\|E\|_{S_1} \leqslant C_\psi\frac{\tau\sqrt{rs}}{N}$, and $\operatorname{rank} H \leqslant N$. Moreover, after increasing C_ψ if necessary, one has $|\psi(t)| \leqslant C_\psi t$ for $0 \leqslant t \leqslant 1$.*

Proof of Lemma 4.6. Choose a representative of ψ' of bounded variation, and put $L_\psi := \|\psi'\|_{L^\infty(0,1)}$ and $V_\psi := \operatorname{Var}_{[0,1]}(\psi')$. Since $\psi(0) = 0$, the Lipschitz bound $|\psi(t)| \leqslant C_\psi t$ follows from absolute continuity once $C_\psi \geqslant L_\psi$.

Fix $0 < \tau \leqslant 1$. Define a function g_τ on $[-2,2]$ by

$$g_\tau(t) = \begin{cases} 0, & -2 \leqslant t \leqslant 0,\\ \psi(\tau t), & 0 \leqslant t \leqslant 1,\\ (2-t)\psi(\tau), & 1 \leqslant t \leqslant 2.\end{cases}$$

Extend g_τ to a continuous 4-periodic function on $\mathbb{R}$. On $[-1,1]$, we have $g_\tau(t) = \psi((\tau t)_+)$. The distributional second derivative of this periodic function is a finite signed periodic measure, and

$$\|g_\tau''\|_{\mathrm{TV}(\mathbb{R}/4\mathbb{Z})} \leqslant C\tau(L_\psi + V_\psi),$$

for an absolute constant C. Also, $\|g_\tau\|_{L^\infty([-2,2])} \leqslant L_\psi \tau$. Let $g_\tau(t) = \sum_{\ell\in\mathbb{Z}} \widehat{g}_\tau(\ell) e^{\pi i \ell t/2}$ be the Fourier series on the period-four circle. For $\ell \neq 0$, two integrations by parts in the sense of distributions give

$$|\widehat{g}_\tau(\ell)| \leqslant \frac{C\tau(L_\psi + V_\psi)}{\ell^2}.$$

Moreover, $|\widehat{g}_\tau(0)| \leqslant L_\psi \tau$. Hence, for all $k \geqslant 0$,

$$\sum_{|\ell|>k} |\widehat{g}_\tau(\ell)| \leqslant \frac{B_\psi \tau}{k+1},$$

where B_ψ depends only on ψ.

For $u, v \in [0, \tau]$, the number $(v-u)/\tau$ belongs to $[-1, 1]$. Hence

$$\psi((v-u)_+) = g_\tau\left(\frac{v-u}{\tau}\right) = \sum_{\ell\in\mathbb{Z}} \widehat{g}_\tau(\ell) e^{-\pi i \ell u/(2\tau)} e^{\pi i \ell v/(2\tau)}.$$

Let $k := \left\lfloor \frac{N-1}{2} \right\rfloor$, and let H be the real matrix obtained by keeping only the modes $|\ell| \leqslant k$. The term $\ell = 0$ has rank at most 1, and the modes ℓ and $-\ell$ together have real rank at most 2. Thus

$$\operatorname{rank} H \leqslant 2k + 1 \leqslant N.$$

The remainder is a sum of separated rank-one matrices. Since each separated vector has Euclidean norm $\sqrt{r}$ or $\sqrt{s}$, we get

$$\|E\|_{S_1} \leqslant \sum_{|\ell|>k} |\widehat{g}_\tau(\ell)| \sqrt{rs} \leqslant C_\psi \frac{\tau\sqrt{rs}}{N}.$$

This proves the lemma. $\square$

For $0 \leqslant t \leqslant 1$, define

$$\rho_p(t) := 2^{-p}\big((1+t)^p - (1-t)^p\big).$$

Lemma 4.7. *Let $m \geqslant 1$ be an integer, and let r, s be nonnegative integers such that $r + s \leqslant m$. Let $u_1, \ldots, u_r, v_1, \ldots, v_s \in [0,1]$, and define $M \in \mathbb{R}^{r\times s}$ by $M_{ij} := \rho_p\big((v_j - u_i)_+\big)$. Then, for $p \geqslant 1$ and $\eta \in (0,1]$, there exists a constant $C_{p,\eta}$ depending only on p and η, such that M admits a decomposition $M = E + H$ with $\|E\|_{S_1} \leqslant \eta\left(r + \sum_{j=1}^s v_j^p\right)$ and $\operatorname{rank} H \leqslant C_{p,\eta}\big(m^{\vartheta_p} + \log(2m)\big)$.*

Proof of Lemma 4.7. The function ρ_p satisfies the assumptions of Lemma 4.6. Indeed, $\rho_p(0) = 0$, ρ_p is absolutely continuous, and, on $[0,1)$,

$$\rho_p'(t) = p2^{-p}\big((1+t)^{p-1} + (1-t)^{p-1}\big).$$

This derivative extends to a function of bounded variation on $[0,1]$. Let C_{ρ_p} be the constant from Lemma 4.6. We then split the proof into the three regimes $1 \leqslant p < 2$, $p = 2$, and $p > 2$.

First assume $1 \leqslant p < 2$. Put

$$\tau_* := \min\left\{1, \frac{\eta}{16 C_{\rho_p} m}\right\}.$$

All columns with $v_j \leqslant \tau_*$ are put into the error. Since $|\rho_p(t)| \leqslant C_{\rho_p} t$, we have

$$\|E_{\mathrm{bot}}\|_{S_1} \leqslant \sum_{j: v_j \leqslant \tau_*} \sum_{i=1}^r |\rho_p((v_j - u_i)_+)| \leqslant C_{\rho_p} \tau_* rs \leqslant \frac{\eta r}{16}.$$

For the remaining columns, use dyadic column scales $C_\ell \subseteq (\tau_*, 1]$. For each scale, denote its upper endpoint by τ_ℓ, so that $\frac{\tau_\ell}{2} < v_j \leqslant \tau_\ell$ for every $j \in C_\ell$, except possibly for the first or last truncated scale, which only changes the absolute constants. Set $R_\ell := \{i : u_i < \tau_\ell\}$, $r_\ell := |R_\ell|$, $s_\ell := |C_\ell|$, and $B_\ell := \sum_{j \in C_\ell} v_j^p$. The block $R_\ell \times C_\ell$ contains all nonzero entries from this scale. Moreover,

$$s_\ell \tau_\ell^p \leqslant 2^p \sum_{j \in C_\ell} v_j^p = 2^p B_\ell.$$

Apply Lemma 4.6 to this block with $N_0 := \lceil C_p C_{\rho_p} \eta^{-1} \rceil$. Then

$$\|E_\ell\|_{S_1} \leqslant C_{\rho_p} \frac{\tau_\ell \sqrt{r_\ell s_\ell}}{N_0}.$$

Now

$$\tau_\ell \sqrt{r_\ell s_\ell} = \sqrt{\tau_\ell^{2-p} r_\ell} \sqrt{s_\ell \tau_\ell^p} \leqslant C_p \big(\tau_\ell^{2-p} r_\ell + B_\ell \big).$$

Since $2 - p > 0$, the dyadic sum is summable:

$$\sum_\ell \tau_\ell^{2-p} r_\ell = \sum_{i=1}^{r} \sum_{\ell : u_i < \tau_\ell} \tau_\ell^{2-p} \leqslant C_p r.$$

Also, $\sum_\ell B_\ell \leqslant \sum_{j=1}^s v_j^p$. Thus, after choosing the constant in N_0 sufficiently large,

$$\sum_\ell \|E_\ell\|_{S_1} \leqslant \frac{\eta}{2} \left(r + \sum_{j=1}^{s} v_j^p \right).$$

Together with the bottom scale, this gives the required error estimate. The rank is at most a constant depending on p, η times the number of scales. Since $\tau_*^{-1} \leqslant C_{p,\eta} m$, we get

$$\operatorname{rank} H \leqslant C_{p,\eta} \log(2m).$$

Next assume $p = 2$. Then $\rho_2(t) = t$. Put $\tau_* := \min\left\{1, \frac{\eta}{16m}\right\}$. The columns with $v_j \leqslant \tau_*$ are put into the error, giving

$$\|E_{\mathrm{bot}}\|_{S_1} \leqslant \tau_* r s \leqslant \frac{\eta r}{16}.$$

For each non-bottom dyadic column scale C_ℓ, denote its upper and lower endpoints by τ_ℓ and $\underline{\tau}_\ell$, and put $s_\ell := |C_\ell|$ and $B_\ell := \sum_{j \in C_\ell} v_j^2$. Split the relevant rows as $R_\ell^- := \{i : u_i \leqslant \underline{\tau}_\ell\}$ and $R_\ell^+ := \{i : \underline{\tau}_\ell < u_i < \tau_\ell\}$. Rows with $u_i \geqslant \tau_\ell$ give zero entries on C_ℓ. If $i \in R_\ell^-$ and $j \in C_\ell$, then $u_i < v_j$, and hence $(v_j - u_i)_+ = v_j - u_i$. Thus the block $R_\ell^- \times C_\ell$ has rank at most 2, and is put entirely into the low-rank part.

On $R_\ell^+ \times C_\ell$, apply Lemma 4.6 to $\psi(t) = t$ with $N_0 := \lceil 2 C_{\rho_p} \eta^{-1} \rceil$. Writing $r_\ell^+ := |R_\ell^+|$, we obtain

$$\|E_\ell\|_{S_1} \leqslant C_{\rho_p} \frac{\tau_\ell \sqrt{r_\ell^+ s_\ell}}{N_0}.$$

Since

$$\tau_\ell \sqrt{r_\ell^+ s_\ell} \leqslant \frac{1}{2} r_\ell^+ + \frac{1}{2} s_\ell \tau_\ell^2 \leqslant \frac{1}{2} r_\ell^+ + 2 B_\ell,$$

and since the row bands R_ℓ^+ are disjoint, we have $\sum_\ell r_\ell^+ \leqslant r$. Also $\sum_\ell B_\ell \leqslant \sum_j v_j^2$. Hence the total error is bounded by $\eta \left(r + \sum_j v_j^2 \right)$. The rank contribution per scale is at most $2 + N_0$, so $\operatorname{rank} H \leqslant C_\eta \log(2m)$.

Finally assume $p > 2$. Put $\theta := \vartheta_p = \frac{p-2}{3p-2}$, $M_0 := m^\theta$ and $\tau_0 := m^{-2/(3p-2)}$. Then $\tau_0^{1-p/2} = M_0$ and $\tau_0\sqrt{\frac{m}{M_0}} = M_0$. Let

$$C_0 := \{j : 0 \leqslant v_j \leqslant \tau_0\}$$

be the bottom scale. For the remaining columns use dyadic scales

$$C_\ell := \{j : 2^{\ell-1}\tau_0 < v_j \leqslant 2^\ell \tau_0\},$$

for $\ell \geqslant 1$, with the last scale truncated at 1. The number of non-bottom scales is at most $C_p \log(2m)$.

First consider the bottom scale. Put $R_0 := \{i : u_i < \tau_0\}$, $r_0 := |R_0|$ and $s_0 := |C_0|$. If $r_0 \leqslant M_0$, put the whole bottom block into the low-rank part, and its rank is at most M_0. If $r_0 > M_0$, apply Lemma 4.6 with $N_0 := \lceil C_p C_{\rho_p} \eta^{-1} M_0 \rceil$. Since $s_0 \leqslant m$ and $r_0 > M_0$,

$$\tau_0\sqrt{r_0 s_0} \leqslant \tau_0 r_0 \sqrt{\frac{m}{r_0}} \leqslant \tau_0 r_0 \sqrt{\frac{m}{M_0}} = M_0 r_0.$$

Therefore the bottom error is at most $\eta r_0/4$, after choosing the constant in N_0 large enough, and the bottom rank is $O_{p,\eta}(M_0)$.

Now consider non-bottom scales $\ell \geqslant 1$. Define $R_\ell := \{i : u_i < \tau_\ell\}$, $r_\ell := |R_\ell|$, $s_\ell := |C_\ell|$ and $B_\ell := \sum_{j \in C_\ell} v_j^p$. As before, $s_\ell \tau_\ell^p \leqslant C_p B_\ell$. Set $\gamma := \frac{p-2}{3}$, $S := \sum_{\ell \geqslant 1} \tau_\ell^{-\gamma}$ and $\alpha_\ell := \frac{\tau_\ell^{-\gamma}}{S}$. If there are no non-bottom scales, there is nothing to prove. Otherwise $\sum_\ell \alpha_\ell = 1$. Apply Lemma 4.6 on the ℓ-th block with $N_\ell := \left\lceil C_p C_{\rho_p} \eta^{-1} \frac{\tau_\ell^{1-p/2}}{\sqrt{\alpha_\ell}} \right\rceil$. Then

$$\|E_\ell\|_{S_1} \leqslant C_{\rho_p} \frac{\tau_\ell \sqrt{r_\ell s_\ell}}{N_\ell} \leqslant C_p \eta \sqrt{\alpha_\ell r_\ell} \sqrt{s_\ell \tau_\ell^p}.$$

Using $2ab \leqslant a^2 + b^2$, we obtain

$$\|E_\ell\|_{S_1} \leqslant \frac{\eta}{4} \alpha_\ell r_\ell + C_p \eta B_\ell.$$

Summing in ℓ, and using $\sum_\ell \alpha_\ell r_\ell \leqslant r$ and $\sum_\ell B_\ell \leqslant \sum_j v_j^p$, the total non-bottom error is bounded by $\frac{\eta}{2}\left(r + \sum_j v_j^p\right)$, after adjusting constants.

It remains to estimate the rank. We have

$$\sum_\ell N_\ell \leqslant C_p \log(2m) + C_{p,\eta} \sum_\ell \frac{\tau_\ell^{1-p/2}}{\sqrt{\alpha_\ell}}.$$

Since $\alpha_\ell^{-1/2} = S^{1/2} \tau_\ell^{\gamma/2}$ and $1 - \frac{p}{2} + \frac{\gamma}{2} = -\gamma$, we get

$$\sum_\ell \frac{\tau_\ell^{1-p/2}}{\sqrt{\alpha_\ell}} = S^{1/2} \sum_\ell \tau_\ell^{-\gamma} = S^{3/2}.$$

The dyadic sum is dominated by its first term: $S \leqslant C_p \tau_0^{-\gamma}$. Therefore

$$S^{3/2} \leqslant C_p \tau_0^{-3\gamma/2} = C_p \tau_0^{-(p-2)/2} = C_p m^{(p-2)/(3p-2)} = C_p M_0.$$

Thus the non-bottom rank is at most $C_{p,\eta}\big(M_0 + \log(2m)\big)$. Including the bottom scale gives the same bound. This completes the proof. □

Proposition 4.8. *Fix $p \geqslant 1$ and $\varepsilon \in (0,1]$. There exists a constant $C_{p,\varepsilon}$ depending only on p and ε, such that the following holds. Let $X = \{z_1, \ldots, z_m\} \subseteq [-1/2, 1/2]$, and define $W_p(x,y) := |x-y|^p - \rho(x-y)^p$, and $W_p(X) := \big(W_p(z_\alpha, z_\beta)\big)_{\alpha,\beta=1}^m$. Then $W_p(X)$ admits a decomposition $W_p(X) = F + G$ with $F, G \in \mathbb{R}^{m\times m}$ symmetric and satisfying $\operatorname{tr}_+(F) \leqslant \varepsilon \sum_{\ell=1}^m |z_\ell|^p$ and $\operatorname{rank}(G) \leqslant C_{p,\varepsilon}\big(m^{\vartheta_p} + \log(2m)\big)$.*

Proof of Proposition 4.8. Reorder the sample so that the positive points come first and the negative points second. Points equal to 0 give zero rows and columns in $W_p(X)$, and may be ignored.

Write the positive points as $x_i = \frac{a_i}{2}$, with $a_i \in (0,1]$, and the negative points as $y_j = -\frac{b_j}{2}$, with $b_j \in (0,1]$.

Same-sign pairs give no contribution, since then $|x-y| \leqslant 1/2$ and hence $\rho(x-y) = |x-y|$.

For an opposite-sign pair $x = a/2 > 0$, $y = -b/2 < 0$, one has

$$|x-y| = \frac{a+b}{2}.$$

If $a+b \leqslant 1$, then again $\rho(x-y) = |x-y|$, so $W_p(x,y) = 0$. If $a+b > 1$, then

$$\rho(x-y) = 1 - \frac{a+b}{2} = \frac{2-a-b}{2}.$$

Therefore, in all cases,

$$W_p(x,y) = \rho_p\big((a+b-1)_+\big),$$

where $\rho_p(t) = 2^{-p}\big((1+t)^p - (1-t)^p\big)$. Thus, after the above ordering,

$$W_p(X) = \begin{pmatrix} 0 & D \\ D^\mathsf{T} & 0 \end{pmatrix},$$

where $D_{ij} = \rho_p\big((a_i + b_j - 1)_+\big)$.

Split the positive and negative indices into

$$I_{\mathrm{hi}} := \{i : a_i \geqslant 1/2\}, \qquad I_{\mathrm{lo}} := \{i : a_i < 1/2\},$$

and

$$J_{\mathrm{hi}} := \{j : b_j \geqslant 1/2\}, \qquad J_{\mathrm{lo}} := \{j : b_j < 1/2\}.$$

On $I_{\mathrm{lo}} \times J_{\mathrm{lo}}$, one has $a_i + b_j < 1$, and hence the corresponding block of D vanishes.

First consider the block $I_{\mathrm{hi}} \times \{1,2,\dots,s\}$. Put $u_i := 1 - a_i$ and $v_j := b_j$. Then $a_i + b_j - 1 = v_j - u_i$, and the block has the form $\rho_p\big((v_j - u_i)_+\big)$. Applying Lemma 4.7 with parameter $\eta \in (0,1]$, we get a decomposition of this block into $E_1 + H_1$, with

$$\|E_1\|_{S_1} \leqslant \eta \left(|I_{\mathrm{hi}}| + \sum_j b_j^p \right)$$

and $\operatorname{rank} H_1 \leqslant C_{p,\eta}\big(m^{\vartheta_p} + \log(2m)\big)$. Since $a_i \geqslant 1/2$ on I_{hi}, $|I_{\mathrm{hi}}| \leqslant 2^p \sum_{i \in I_{\mathrm{hi}}} a_i^p$. Thus

$$\|E_1\|_{S_1} \leqslant C_p \eta \left(\sum_i a_i^p + \sum_j b_j^p \right).$$

Next consider the block $I_{\mathrm{lo}} \times J_{\mathrm{hi}}$. Transpose it and put $u_j := 1 - b_j$ and $v_i := a_i$. Then $a_i + b_j - 1 = v_i - u_j$, so the transposed block again has the form $\rho_p\big((v_i - u_j)_+\big)$. Lemma 4.7 gives a decomposition $E_2 + H_2$, with

$$\|E_2\|_{S_1} \leqslant \eta \left(|J_{\mathrm{hi}}| + \sum_{i \in I_{\mathrm{lo}}} a_i^p \right)$$

and $\operatorname{rank} H_2 \leqslant C_{p,\eta}\big(m^{\vartheta_p} + \log(2m)\big)$. Since $b_j \geqslant 1/2$ on J_{hi}, $|J_{\mathrm{hi}}| \leqslant 2^p \sum_{j \in J_{\mathrm{hi}}} b_j^p$, and therefore

$$\|E_2\|_{S_1} \leqslant C_p \eta \left(\sum_i a_i^p + \sum_j b_j^p \right).$$

Combining the two rectangular pieces, we obtain $D = E_D + H_D$, with $\|E_D\|_{S_1} \leqslant C_p\eta \left(\sum_i a_i^p + \sum_j b_j^p\right)$, and $\operatorname{rank} H_D \leqslant C_{p,\eta}\big(m^{\vartheta_p} + \log(2m)\big)$.

Now embed this rectangular decomposition into the full symmetric matrix:

$$F := \begin{pmatrix} 0 & E_D \\ E_D^{\mathsf{T}} & 0 \end{pmatrix}, \qquad G := \begin{pmatrix} 0 & H_D \\ H_D^{\mathsf{T}} & 0 \end{pmatrix}.$$

Then

$$W_p(X) = F + G.$$

The nonzero eigenvalues of F are precisely the singular values of E_D, with both signs. Hence

$$\operatorname{tr}_+(F) = \|E_D\|_{S_1}.$$

Also, $\operatorname{rank} G \leqslant 2 \operatorname{rank} H_D$. Consequently,

$$\operatorname{tr}_+(F) \leqslant C_p\eta \left(\sum_i a_i^p + \sum_j b_j^p\right),$$

and $\operatorname{rank} G \leqslant C_{p,\eta}\big(m^{\vartheta_p} + \log(2m)\big)$.

Finally, since $a_i = 2|x_i|$ and $b_j = 2|y_j|$, we have

$$\sum_i a_i^p + \sum_j b_j^p = 2^p \sum_{\ell=1}^{m} |z_\ell|^p.$$

Choosing $\eta = \varepsilon/(C_p 2^p)$, we obtain

$$\operatorname{tr}_+(F) \leqslant \varepsilon \sum_{\ell=1}^{m} |z_\ell|^p.$$

The rank estimate becomes

$$\operatorname{rank} G \leqslant C_{p,\varepsilon}\big(m^{\vartheta_p} + \log(2m)\big).$$

This completes the proof. □

4.5 Proof of the one-dimensional proposition

Proof of Proposition 4.2. Let $\varepsilon_0 := \frac{\varepsilon}{2}$. We identify $\mathbb{T}$ with $[-1/2, 1/2)$, and write x_i also for the representative of x_i in this interval. Thus $\rho(x_i) = |x_i|$. For $x, y \in [-1/2, 1/2)$, define

$$B_{\mathbb{R}}(x,y) := \frac{1}{2}\big(|x|^p + |y|^p - |x-y|^p\big),$$

and

$$W(x,y) := \frac{1}{2}\big(|x-y|^p - \rho(x-y)^p\big).$$

Then

$$B_p(x,y) = B_{\mathbb{R}}(x,y) + W(x,y),$$

and hence the matrix M decomposes as $M = B_{\mathbb{R}} + W$.

By Lemma 4.4, applied with ε_0, there are real symmetric matrices $L_{\mathbb{R}}, R_{\mathbb{R}}$ such that $B_{\mathbb{R}} \preceq L_{\mathbb{R}} + R_{\mathbb{R}}$, $\operatorname{rank} L_{\mathbb{R}} \leqslant C_{p,\varepsilon} \log(2N)$, $R_{\mathbb{R}} \succeq 0$, and

$$\operatorname{tr}_+(R_{\mathbb{R}}) = \operatorname{tr}(R_{\mathbb{R}}) \leqslant \varepsilon_0 \sum_{i=1}^{m} |x_i|^p + C_{p,\varepsilon}\frac{m}{N}.$$

By Proposition 4.8, applied with ε_0, the matrix

$$W_p(X) = \big(|x_i - x_j|^p - \rho(x_i - x_j)^p\big)_{i,j=1}^m$$

admits a decomposition $W_p(X) = F_0 + G_0$ such that $\operatorname{tr}_+(F_0) \leqslant \varepsilon_0 \sum_{i=1}^m |x_i|^p$, and

$$\operatorname{rank} G_0 \leqslant C_{p,\varepsilon}\big(m^{\vartheta_p} + \log(2m)\big).$$

Since $W = \frac{1}{2} W_p(X)$, we may write $W = F + G$ with $\operatorname{tr}_+(F) \leqslant \varepsilon_0 \sum_{i=1}^m |x_i|^p$ and

$$\operatorname{rank} G \leqslant C_{p,\varepsilon}\big(m^{\vartheta_p} + \log(2m)\big).$$

Finally set $L := L_{\mathbb{R}} + G$ and $R := R_{\mathbb{R}} + F$. Then

$$M = B_{\mathbb{R}} + W \preceq L_{\mathbb{R}} + R_{\mathbb{R}} + F + G = L + R.$$

Moreover,

$$\operatorname{rank} L \leqslant \operatorname{rank} L_{\mathbb{R}} + \operatorname{rank} G \leqslant C_{p,\varepsilon}\big(m^{\vartheta_p} + \log(2m) + \log(2N)\big).$$

By subadditivity of tr_+,

$$\operatorname{tr}_+(R) \leqslant \operatorname{tr}_+(R_{\mathbb{R}}) + \operatorname{tr}_+(F) \leqslant \varepsilon \sum_{i=1}^m \rho(x_i)^p + C_{p,\varepsilon} \frac{m}{N}.$$

This proves the proposition. $\square$

5 Lower bounds for $e_p(\mathbb{T}^n)$

In this section, we present two constructions that extend and strengthen the known lower bounds for equilateral sets in $\mathbb{T}^n$. The first extends Alon's construction from the case where $2n+1$ is prime to the more general case where $2n+1$ is an odd prime power; in this setting, it yields $2n+1$ points in $\mathbb{T}^n$ forming an equilateral set for every $1 \leqslant p \leqslant \infty$. This construction is given in Subsection 5.1.

The second construction yields larger equilateral sets for suitably chosen exponents. We show that there exist infinitely many primes q for which one can choose $p_q > (q-1)/4$ such that, for every integer $m \geqslant 1$, if $n = \frac{q^m - 1}{4}$, then $e_{p_q}(\mathbb{T}^n) \geqslant 4n + 1$. This construction is given in Subsection 5.2.

5.1 Proof of Theorem 1.6

Let $V := \mathbb{F}_r^m$ be an m-dimensional vector space over $\mathbb{F}_r$, and then $|V| = r^m = q$. We write

$$V^* := \operatorname{Hom}_{\mathbb{F}_r}(V, \mathbb{F}_r)$$

for the dual space of V, namely the set of all $\mathbb{F}_r$-linear maps $\varphi : V \to \mathbb{F}_r$. Elements of V^* are called *linear functionals* on V. It is a standard fact that V^* is again an m-dimensional vector space over $\mathbb{F}_r$. In particular, $|V^*| = r^m = q$. For these basic facts on dual spaces, see for example [17].

Since r is odd, every nonzero functional $\varphi \in V^*$ is distinct from $-\varphi$. Therefore the set $V^* \backslash \{0\}$ is partitioned into disjoint pairs $\{\varphi, -\varphi\}$. Notice that $|V^* \backslash \{0\}| = q - 1 = 2n$, there are exactly n such pairs. Choose one representative from each pair and denote them by

$$\varphi_1, \ldots, \varphi_n \in V^* \backslash \{0\}.$$

We also identify $\mathbb{F}_r$ with the set of integers $\{0, 1, \ldots, r-1\}$, and then with the corresponding subset of $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ via $a \to \frac{a}{r}$. Now define a map $\eta : V \to \mathbb{T}^n$ by

$$\eta(\boldsymbol{x}) := \left(\frac{\varphi_1(\boldsymbol{x})}{r}, \frac{\varphi_2(\boldsymbol{x})}{r}, \ldots, \frac{\varphi_n(\boldsymbol{x})}{r} \right).$$

We claim that $\eta(V)$ is an equilateral set in $(\mathbb{T}^n, d_p)$. Let $\boldsymbol{x}, \boldsymbol{y} \in V$ be distinct, and write $\boldsymbol{u} := \boldsymbol{x} - \boldsymbol{y} \neq \mathbf{0}$. Consider the evaluation map $E_{\boldsymbol{u}} : V^* \to \mathbb{F}_r$, with $E_{\boldsymbol{u}}(\varphi) := \varphi(\boldsymbol{u})$. Since $\boldsymbol{u} \neq \mathbf{0}$, this is a nonzero linear functional on the m-dimensional space V^*. Hence $E_{\boldsymbol{u}}$ is surjective, and every fiber has cardinality $|E_{\boldsymbol{u}}^{-1}(a)| = r^{m-1}$ for each $a \in \mathbb{F}_r$. In particular, among all functionals in V^*,

- exactly r^{m-1} satisfy $\varphi(\boldsymbol{u}) = a$ for each $a \in \mathbb{F}_r$;
- hence, among the nonzero functionals in $V^* \backslash \{0\}$, exactly $r^{m-1} - 1$ satisfy $\varphi(\boldsymbol{u}) = 0$, and for each nonzero $a \in \mathbb{F}_r^\times$, exactly r^{m-1} satisfy $\varphi(\boldsymbol{u}) = a$.

Now pass from nonzero functionals to the chosen representatives $\varphi_1, \ldots, \varphi_n$, that is, to the quotient by the involution $\varphi \to -\varphi$. For each integer $t \in \left\{1, 2, \ldots, \frac{r-1}{2}\right\}$, the set of nonzero field elements $\{\pm t\}$ consists of two distinct elements of $\mathbb{F}_r$. Since each of the values t and $-t$ occurs exactly r^{m-1} times among all nonzero functionals, it follows that among the chosen representatives $\varphi_1, \ldots, \varphi_n$, exactly r^{m-1} of them satisfy $\varphi_j(\boldsymbol{u}) \in \{t, -t\}$. Also, exactly $\frac{r^{m-1}-1}{2}$ of the chosen representatives satisfy $\varphi_j(\boldsymbol{u}) = 0$.

By the definitions, for a coordinate j with $\varphi_j(\boldsymbol{u}) \in \{t, -t\}$, the j-th coordinate difference is either t/r or $1 - t/r$, and hence the circular distance is exactly $\frac{t}{r}$. Therefore

$$d_p\big(\eta(\boldsymbol{x}), \eta(\boldsymbol{y})\big)^p = r^{m-1} \sum_{t=1}^{(r-1)/2} \left(\frac{t}{r}\right)^p.$$

Thus $\eta(V)$ is indeed an equilateral set in $(\mathbb{T}^n, d_p)$.

It remains only to note that η is injective. Indeed, if $\boldsymbol{x} \neq \boldsymbol{y}$, then $\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{y} \neq \mathbf{0}$, and hence there exists some $\varphi \in V^*$ with $\varphi(\boldsymbol{u}) \neq 0$. Therefore the pair $\{\varphi, -\varphi\}$ contributes one chosen representative φ_j with $\varphi_j(\boldsymbol{u}) \neq 0$, so the j-th coordinates of $\eta(\boldsymbol{x})$ and $\eta(\boldsymbol{y})$ are different. Thus $|\eta(V)| = |V| = r^m = q = 2n + 1$. This finishes the proof.

5.2 Proof of Theorem 1.7

For an odd prime q and an integer u, write $\|u\|_q \in \left\{0, 1, \ldots, \frac{q-1}{2}\right\}$ for the unique integer such that $\|u\|_q \equiv \pm u \pmod q$. We also write χ_q for the quadratic character modulo q, so that $\chi_q(0) = 0$, $\chi_q(a) = 1$ if a is a nonzero quadratic residue modulo q, and $\chi_q(a) = -1$ otherwise.

We first construct a q-point equilateral seed from quadratic residues.

Lemma 5.1. *Let q be a prime with $q \equiv 1 \pmod 4$. Set $M := \frac{q-1}{2}$ and define $R_q := \{a \in \{1, 2, \ldots, M\} : \chi_q(a) = 1\}$. For $s > 0$, let $F_q(s) := \sum_{a=1}^{M} \chi_q(a) a^s$. If $F_q(p) = 0$ for some $p > 0$, then*

$$B_q := \left\{ \boldsymbol{b}_j = \left(\frac{aj}{q}\right)_{a \in R_q} : j \in \mathbb{F}_q \right\} \subseteq \mathbb{T}^{(q-1)/4}$$

is a q-point equilateral set in $(\mathbb{T}^{(q-1)/4}, d_p)$.

Proof of Lemma 5.1. Since $q \equiv 1 \pmod 4$, we have $\chi_q(-1) = 1$. Hence the two elements a and $-a$ have the same quadratic character. Therefore, among the representatives $1, 2, \ldots, M$, exactly half are quadratic residues. Thus

$$|R_q| = \frac{M}{2} = \frac{q-1}{4}.$$

Let $N_q := \{1, 2, \ldots, M\} \backslash R_q$. Notice that the points $\boldsymbol{b}_j$ are distinct. Indeed, if $\boldsymbol{b}_j = \boldsymbol{b}_k$, then for every $a \in R_q$ we have $a(j-k) \equiv 0 \pmod q$. Since $a \not\equiv 0 \pmod q$, this implies $j = k$. Hence $|B_q| = q$. Now take $j \neq k$, and put $h := j - k \in \mathbb{F}_q^\times$. Then

$$d_p(\boldsymbol{b}_j, \boldsymbol{b}_k)^p = q^{-p} \sum_{a \in R_q} \|ah\|_q^p.$$

If $\chi_q(h) = 1$, multiplication by h preserves quadratic residues. Since $\chi_q(-1) = 1$, it also preserves the quadratic character on each pair $\{\pm u\}$. Thus $\{\|ah\|_q : a \in R_q\} = R_q$. If $\chi_q(h) = -1$, multiplication by h swaps residues and nonresidues on the pairs $\{\pm u\}$, and therefore $\{\|ah\|_q : a \in R_q\} = N_q$. Consequently, $d_p(\boldsymbol{b}_j, \boldsymbol{b}_k)^p$ equals either $q^{-p} \sum_{u \in R_q} u^p$ or $q^{-p} \sum_{u \in N_q} u^p$. Notice that

$$F_q(p) = \sum_{u \in R_q} u^p - \sum_{u \in N_q} u^p = 0,$$

which yields B_q is equilateral. $\square$

The next lemma lifts a q-point seed to a larger equilateral set using the standard projective simplex code.

Lemma 5.2. *Let q be a prime, let $p > 0$, and let $B = \{b_\alpha : \alpha \in \mathbb{F}_q\} \subseteq \mathbb{T}^r$ be a q-point equilateral set in $(\mathbb{T}^r, d_p)$, with common nonzero distance δ. Then for every integer $m \geqslant 1$, the space $\left(\mathbb{T}^{r(q^m-1)/(q-1)}, d_p\right)$ contains an equilateral set of cardinality q^m.*

Proof of Lemma 5.2. Let $L := \frac{q^m-1}{q-1}$. Choose representatives $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_L$ of the one-dimensional subspaces of $\mathbb{F}_q^m$. For $\boldsymbol{u} \in \mathbb{F}_q^m$, define

$$c(\boldsymbol{u}) := (\boldsymbol{u} \cdot \boldsymbol{v}_1, \ldots, \boldsymbol{u} \cdot \boldsymbol{v}_L) \in \mathbb{F}_q^L.$$

We first recall the standard distance property of this simplex code. If $\boldsymbol{u} \neq \boldsymbol{u}'$, put $\boldsymbol{w} := \boldsymbol{u} - \boldsymbol{u}' \neq \boldsymbol{0}$. The coordinates i for which $\boldsymbol{u} \cdot \boldsymbol{v}_i = \boldsymbol{u}' \cdot \boldsymbol{v}_i$ are exactly those for which

$$\boldsymbol{v}_i \in \boldsymbol{w}^\perp := \{\boldsymbol{x} \in \mathbb{F}_q^m : \boldsymbol{w} \cdot \boldsymbol{x} = 0\}.$$

Since $\boldsymbol{w} \neq \boldsymbol{0}$, the hyperplane $\boldsymbol{w}^\perp$ has dimension $m-1$, and hence contains $\frac{q^{m-1}-1}{q-1}$ one-dimensional subspaces. Therefore the number of coordinates on which $c(\boldsymbol{u})$ and $c(\boldsymbol{u}')$ differ is

$$L - \frac{q^{m-1}-1}{q-1} = q^{m-1}.$$

Now define

$$\Phi(\boldsymbol{u}) := \left(\boldsymbol{b}_{\boldsymbol{u}\cdot\boldsymbol{v}_1}, \ldots, \boldsymbol{b}_{\boldsymbol{u}\cdot\boldsymbol{v}_L}\right) \in (\mathbb{T}^r)^L = \mathbb{T}^{rL}.$$

If $\boldsymbol{u} \neq \boldsymbol{u}'$, then the previous paragraph shows that exactly q^{m-1} blocks of $\Phi(\boldsymbol{u})$ and $\Phi(\boldsymbol{u}')$ are different. In each different block, the two entries are distinct points of the seed B, and therefore contribute δ^p to the p-th power of the distance. The remaining blocks contribute 0. Hence

$$d_p(\Phi(\boldsymbol{u}), \Phi(\boldsymbol{u}'))^p = q^{m-1}\delta^p,$$

independently of the choice of distinct $\boldsymbol{u}, \boldsymbol{u}' \in \mathbb{F}_q^m$. Thus $\{\Phi(\boldsymbol{u}) : \boldsymbol{u} \in \mathbb{F}_q^m\}$ is equilateral in $\mathbb{T}^{rL}$.

Moreover, if $\boldsymbol{u} \neq \boldsymbol{u}'$, then $c(\boldsymbol{u})$ and $c(\boldsymbol{u}')$ differ in q^{m-1} coordinates. In particular, $\Phi(\boldsymbol{u}) \neq \Phi(\boldsymbol{u}')$. Therefore the constructed equilateral set has cardinality $|\mathbb{F}_q^m| = q^m$. This completes the proof. $\square$

We now show that infinitely many primes give rise to an admissible exponent p_q.

Proposition 5.3. *There exists a constant Q_0 with the following property. Let $q \geqslant Q_0$ be a prime satisfying $q \equiv 1 \pmod 4$ and $\chi_q(3) = \chi_q(5) = \chi_q(7) = \chi_q(11) = -1$. Set $M := \frac{q-1}{2}$ and $F_q(s) := \sum_{a=1}^{M} \chi_q(a) a^s$. Then there exists $p_q > \frac{q-1}{4}$ such that $F_q(p_q) = 0$. Consequently, for every integer $m \geqslant 1$, if $n = \frac{q^m-1}{4}$, then $e_{p_q}(\mathbb{T}^n) \geqslant 4n + 1$.*

Proof of Proposition 5.3. For $x > 0$, define $H_q(x) := M^{-Mx}F_q(Mx)$. Then

$$H_q(x) = \sum_{a=1}^{M} \chi_q(a)\left(\frac{a}{M}\right)^{Mx}.$$

Reindex by writing $a = M - j$, where $0 \leqslant j \leqslant M - 1$. Since

$$M - j = \frac{q-1}{2} - j \equiv -\frac{2j+1}{2} \pmod{q},$$

and since $q \equiv 1 \pmod 4$, we have $\chi_q(-1) = 1$. Set $\varepsilon_q := \chi_q(2) \in \{\pm 1\}$. Since $\chi_q(2)^{-1} = \chi_q(2)$, we get

$$\chi_q(M - j) = \chi_q(-1)\chi_q(2)^{-1}\chi_q(2j+1) = \varepsilon_q\chi_q(2j+1).$$

Therefore

$$\varepsilon_q H_q(x) = \sum_{j=0}^{M-1} \chi_q(2j+1)\left(1 - \frac{j}{M}\right)^{Mx}.$$

We evaluate this expression at $x = \frac{1}{2}$. The assumptions give $\chi_q(1) = 1$, $\chi_q(3) = \chi_q(5) = \chi_q(7) = \chi_q(11) = -1$ and $\chi_q(9) = \chi_q(3)^2 = 1$. Hence the first six coefficients $\chi_q(2j+1)$, for $j = 0, 1, \ldots, 5$, are $1, -1, -1, -1, 1, -1$. Thus

$$\varepsilon_q H_q\left(\frac{1}{2}\right) \leqslant \sum_{j=0}^{5} s_j\left(1 - \frac{j}{M}\right)^{M/2} + \sum_{j=6}^{M-1}\left(1 - \frac{j}{M}\right)^{M/2},$$

where $(s_0, s_1, s_2, s_3, s_4, s_5) = (1, -1, -1, -1, 1, -1)$. For each fixed j, $\left(1 - \frac{j}{M}\right)^{M/2} \to e^{-j/2}$ when $M \to \infty$, and for all $0 \leqslant j \leqslant M - 1$, $\left(1 - \frac{j}{M}\right)^{M/2} \leqslant e^{-j/2}$. It follows that

$$\limsup_{q\to\infty} \varepsilon_q H_q\left(\frac{1}{2}\right) \leqslant C,$$

where $C := 1 - e^{-1/2} - e^{-1} - e^{-3/2} + e^{-2} - e^{-5/2} + \sum_{j=6}^{\infty} e^{-j/2}$. Since $\sum_{j=6}^{\infty} e^{-j/2} = \frac{e^{-3}}{1-e^{-1/2}}$, we have

$$C = 1 - e^{-1/2} - e^{-1} - e^{-3/2} + e^{-2} - e^{-5/2} + \frac{e^{-3}}{1 - e^{-1/2}}.$$

Set $t := e^{-1/2}$. Then

$$C = 1 - t - t^2 - t^3 + t^4 - t^5 + \frac{t^6}{1-t} = \frac{N(t)}{1-t},$$

where $N(t) = 1 - 2t + 2t^4 - 2t^5 + 2t^6$. Notice that $3/5 < t < 61/100$ and we can also compute that $N'(u) = 2(6u^5 - 5u^4 + 4u^3 - 1)$. Then one can check that $N(t)$ is strictly decreasing on $[0, 61/100]$, and so

$$N(t) < N\left(\frac{3}{5}\right) = -\frac{47}{15625} < 0.$$

Since $1 - t > 0$, we obtain $C < 0$. Consequently, there exists Q_0 such that for every prime $q \geqslant Q_0$ satisfying the hypotheses, $\varepsilon_q H_q\left(\frac{1}{2}\right) < 0$. On the other hand, as $x \to \infty$, every term in the expression for $\varepsilon_q H_q(x)$ with $j \geqslant 1$ tends to 0, while the $j = 0$ term is equal to 1. Thus

$$\lim_{x\to\infty} \varepsilon_q H_q(x) = 1 > 0.$$

Since $\varepsilon_q H_q(x)$ is continuous on $(0,\infty)$, the intermediate value theorem gives some $x_q > \frac{1}{2}$ such that $H_q(x_q) = 0$. Now set $p_q := Mx_q$. Then $F_q(p_q) = 0$ and $p_q > M \cdot \frac{1}{2} = \frac{q-1}{4}$. Finally, Lemma 5.1 gives a q-point equilateral set in $\mathbb{T}^{\frac{q-1}{4}}$ with respect to d_{p_q}. Applying Lemma 5.2 with $r = \frac{q-1}{4}$ gives, for every $m \geqslant 1$, an equilateral set of cardinality q^m in $\mathbb{T}^{\frac{q-1}{4}\cdot\frac{q^m-1}{q-1}} = \mathbb{T}^{\frac{q^m-1}{4}}$. If $n = \frac{q^m-1}{4}$, then $q^m = 4n+1$. Hence $e_{p_q}(\mathbb{T}^n) \geqslant 4n+1$.

□

It remains to show that infinitely many primes satisfy the Legendre-symbol conditions used above.

Lemma 5.4. *There exist infinitely many primes q such that $q \equiv 1 \pmod 4$ and $\chi_q(3) = \chi_q(5) = \chi_q(7) = \chi_q(11) = -1$.*

Proof of Lemma 5.4. For each $\ell \in \{3,5,7,11\}$, choose a quadratic nonresidue $r_\ell \in (\mathbb{Z}/\ell\mathbb{Z})^\times$. By the Chinese remainder theorem, there exists an integer r such that $r \equiv 1 \pmod 4$ and $r \equiv r_\ell \pmod \ell$ for each $\ell = 3,5,7,11$. In particular, $\gcd(r, 4\cdot 3\cdot 5\cdot 7\cdot 11) = 1$. By Dirichlet's theorem on primes in arithmetic progressions, there are infinitely many primes $q \equiv r \pmod{4620}$. For every such prime q, we have $q \equiv 1 \pmod 4$. Hence quadratic reciprocity gives

$$\chi_q(\ell) = \chi_\ell(q) = \chi_\ell(r_\ell) = -1$$

for $\ell = 3,5,7,11$. This proves the lemma. □

Proposition 5.3 and Lemma 5.4 together yields Theorem 1.7.

Remark 5.5. For the sake of keeping the proof concise, we have restricted ourselves to a convenient infinite family of sufficiently large primes q, which already suffices to show that the values of p for which $e_p(\mathbb{T}^n) \geqslant 4n+1$ holds for infinitely many n can be chosen arbitrarily large. In fact, if one only asks for the existence of some $p_q > 1$, rather than the stronger lower bound $p_q > (q-1)/4$, then a similar argument shows that, for every prime $q \equiv 1 \pmod 8$, there exists $p_q > 1$ such that $e_{p_q}(\mathbb{T}^n) \geqslant 4n+1$ for every $m \geqslant 1$, where $n = \frac{q^m-1}{4}$.

6 Concluding remarks

We have proved that Kusner's conjecture holds for every $2 \leqslant p \leqslant 4$ and every dimension n. More generally, for every integer $k \geqslant 0$ and $p \in [4k+2, 4k+4]$, we obtain the linear bound $e(\ell_p^n) \leqslant (2k+1)n+1$, while for the complementary intervals $p \in (4k, 4k+2)$ with $p \geqslant 1$, we prove the almost linear estimate $e(\ell_p^n) \leqslant C_p n\log(2n)$.

The main difficulty in Kusner's conjecture is that the usual finite-dimensional algebraic methods are naturally adapted to even exponents. When p is an even integer, the kernel $|x-y|^p$ is polynomial, and one can hope to control the corresponding distance matrix by rank considerations. Away from the even integers, however, the kernel is no longer finite-rank in any direct sense. This is the central obstruction behind the problem: one needs a way to extract enough finite-dimensional structure from a non-polynomial kernel.

Our approach shows that this obstruction can be overcome on the intervals $[4k+2, 4k+4]$. The key point is not simply to approximate $|x-y|^p$, but to decompose it with the correct signs. More precisely, one constructs, in one dimension, a sign-controlled correction: the signed kernel $-|x-y|^p$ is modified by a quadratic term and finitely many oscillatory kernels, up to a controlled error, so that the semidefinite part has the favorable sign while the oscillatory part has controlled rank. Once the distance matrix is double-centered, the conditionally negative definite contribution becomes positive semidefinite and can be absorbed into the favorable side of the rank comparison. This leaves only a controlled number of unfavorable finite-rank modes to be counted. This sign-sensitive decomposition is what makes a linear bound possible beyond the isolated even cases.

The complementary intervals $(4k, 4k+2)$ appear to be genuinely harder. In this range, the same finite-rank correction no longer has the required sign pattern. Our proof instead keeps an infinite-dimensional Hilbert-space component and then approximates it on dyadic shells. The logarithmic factor in $C_p n \log(2n)$ arises from this multiscale approximation. Thus removing the logarithm seems to require a new way to control all scales at once, or a stronger use of the global equilateral structure.

Question 6.1. Let $p \in (4k, 4k+2)$ for some integer $k \geqslant 1$. Is it true that $e(\ell_p^n) \leqslant C_p n$ for all $n \geqslant 1$?

A positive answer would show that the equilateral number is linear throughout the whole super-Euclidean range $p > 2$. At present, the main obstacle is the lack of a finite-dimensional decomposition with the correct sign pattern on these intervals.

We also prove new upper bounds for the toroidal problem, namely $e_p(\mathbb{T}^n) \leqslant C_p n \log(2n)$ for $p \in [1,2]$, and $e_p(\mathbb{T}^n) \leqslant C_p n^{3/2-1/p}$ for every fixed $p > 2$. The toroidal problem introduces a different difficulty. Even after anchoring at the origin, the cyclic distance creates a wrap-around correction which has no analogue in $\mathbb{R}^n$. This correction is supported near the two opposite-sign corners of the fundamental domain $[-1/2, 1/2)^2$ for $\mathbb{T} \times \mathbb{T}$, and is not controlled by the Euclidean kernel alone. Our proof handles it by reducing the problem to one-sided rectangular kernels and then decomposing these kernels into a low-rank part plus a small positive-trace error. This mechanism is strong enough to improve Alon's polynomial bounds for all finite p, but it still falls short of the expected linear estimate.

Question 6.2. For every fixed $p \geqslant 1$, is there a constant C_p such that $e_p(\mathbb{T}^n) \leqslant C_p n$ for all $n \geqslant 1$?

Even obtaining such a linear bound in the endpoint cases $p = 1$ and $p = 2$ seems to require new ideas. Our bound $e_p(\mathbb{T}^n) \leqslant C_p n \log(2n)$ for $p \in [1,2]$ is within a logarithmic factor of the linear bound asked for above. It remains unclear, however, whether this logarithmic loss is merely an artifact of the method or a genuine feature of the toroidal setting.

Finally, the methods developed here suggest a broader principle. In several places, a global equilateral condition is transformed, after centering, into a rank or positive-trace comparison through a carefully chosen kernel decomposition. It would be interesting to see whether similar sign-sensitive kernel decompositions can be used in other extremal problems where polynomial methods capture only the even exponent or finite-rank cases.

References

[1] N. Alon. Equilateral sets in the torus. `https://web.math.princeton.edu/~nalon/PDFS/equitor.pdf`, 2026.

[2] N. Alon and P. Pudlák. Equilateral sets in l_p^n. *Geom. Funct. Anal.*, 13(3):467–482, 2003.

[3] I. Balla, F. Dräxler, P. Keevash, and B. Sudakov. Equiangular lines and spherical codes in Euclidean space. *Invent. Math.*, 211(1):179–212, 2018.

[4] H.-J. Bandelt, V. Chepoi, and M. Laurent. Embedding into rectilinear spaces. *Discrete Comput. Geom.*, 19(4):595–604, 1998.

[5] C. Berg, J. P. R. Christensen, and P. Ressel. *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions*, volume 100 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984.

[6] D. Bilyk, N. Nagel, and I. Ruohoniemi. Minimizing point configurations for tensor product energies on the torus. *arXiv preprint*, arXiv: 2510.25442, 2025.

[7] B. Bukh. Bounds on equiangular lines and on related spherical codes. *SIAM J. Discrete Math.*, 30(1):549–554, 2016.

[8] R. Chen, F. Gui, J. Tang, and N. Xiong. Few distance sets in ℓ_p spaces and ℓ_p product spaces. *European J. Combin.*, 102:Paper No. 103459, 14, 2022.

[9] S. Gershgorin. Über die Abgrenzung der Eigenwerte einer Matrix. *Bull. Acad. Sci. URSS*, 1931(6):749–754, 1931.

[10] R. K. Guy. Unsolved Problems: An Olla-Podrida of Open Problems, Often Oddly Posed. *Amer. Math. Monthly*, 90(3):196–200, 1983.

[11] Z. Jiang, J. Tidor, Y. Yao, S. Zhang, and Y. Zhao. Equiangular lines with a fixed angle. *Ann. of Math. (2)*, 194(3):729–743, 2021.

[12] S. V. Konyagin. On systems of vectors with equal distances. *Trudy Matematicheskogo Tsentra imeni N. I. Lobachevskogo*, 43:204–209, 2011. In Russian; English translation available.

[13] J. Koolen, M. Laurent, and A. Schrijver. Equilateral dimension of the rectilinear space. volume 21, pages 149–164. 2000. Special issue dedicated to Dr. Jaap Seidel on the occasion of his 80th birthday (Oisterwijk, 1999).

[14] P. W. H. Lemmens and J. J. Seidel. Equiangular lines. *J. Algebra*, 24:494–512, 1973.

[15] P. W. Nowak. Group actions on banach spaces and a geometric characterization of a-t-menability. *Topology Appl.*, 153(18):3409–3412, 2006.

[16] C. M. Petty. Equilateral sets in Minkowski spaces. *Proc. Amer. Math. Soc.*, 29:369–374, 1971.

[17] S. Roman. *Advanced linear algebra*, volume 135 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1992.

[18] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938.

[19] C. Smyth. *The conjectures of Rudich, Tardos, and Kusner*. ProQuest LLC, Ann Arbor, MI, 2001. Thesis (Ph.D.)–Rutgers The State University of New Jersey - New Brunswick.

[20] C. Smyth. Equilateral sets in ℓ_p^d. In J. Pach, editor, *Thirty Essays on Geometric Graph Theory*, pages 483–487. Springer, New York, 2013.

[21] K. J. Swanepoel. A problem of Kusner on equilateral sets. *Arch. Math. (Basel)*, 83(2):164–170, 2004.

[22] K. J. Swanepoel. Equilateral sets and a Schütte theorem for the 4-norm. *Canad. Math. Bull.*, 57(3):640–647, 2014.

[23] J. H. van Lint and J. J. Seidel. Equilateral point sets in elliptic geometry. *Indag. Math.*, 28:335–348, 1966. Nederl. Akad. Wetensch. Proc. Ser. A 69.