VISIBILITY PHENOMENA IN HYPERCUBES

JAYADEV S. ATHREYA, CRISTIAN COBELI, ALEXANDRU ZAHARESCU

Abstract. We study the set of visible lattice points in multidimensional hypercubes. The problems we investigate mix together geometric, probabilistic and number theoretic tones. For example, we prove that almost all self-visible triangles with vertices in the lattice of points with integer coordinates in $\mathcal{W} = [0, N]^d$ are almost equilateral having all sides almost equal to $\sqrt{d}N/\sqrt{6}$, and the sine of the typical angle between rays from the visual spectra from the origin of $\mathcal{W}$ is, in the limit, equal to $\sqrt{7}/4$, as d and N/d tend to infinity. We also show that there exists an interesting number theoretic constant $\Lambda_{d,K}$, which is the limit probability of the chance that a K-polytope with vertices in the lattice $\mathcal{W}$ has all vertices visible from each other.

1. Introduction

Various phenomena related to distances in high dimensional spaces have attracted attention recently. For instance Gafni, Iosevich and Wyman [11] continue the study of the interesting connections with the unit distance problem in higher dimentions explored in [13–16, 19]. Problems linked to the distribution of distances between points in finite sets placed in metric spaces (particular Euclidean spaces) have been investigated by various authors from many different perspectives. A selection of such results, by no means complete, includes the works of general theoretical interest of Bäsel [3], Baileya, Borwein and Crandall [2], Burgstaller and Pillichshammer [5], Dunbar [9], Mathai, Moschopoulos and Pederzoli [18].

There are also many practical applications of these problems, in particular in high-dimensional data analysis. For example, the article of Aggarwal, Hinneburg and Keim [1] is related to data mining techniques, Li and Qiu [17] study probabilistic problems related to wireless communication networks, Srinivasa and Haenggi [21] are interested in wireless networks whose efficiency is strongly influenced by the nodal distances, while Bubeck and Sellke [4] prove a universal law of robustness that explains the necessity of overparametrization in deep neural networks.

In the present paper we study a few aspects related to visible lattice points in high dimensional hypercubes. Let $\mathcal{C} := [0, N]^d \subset \mathbb{R}^d$ be the d dimensional real cube of side length N, for some integers $d, N \geq 1$. Denote by $\mathcal{W} := \mathcal{C} \cap \mathbb{Z}^d$ the set of $(N + 1)^d$ points with integer coordinates in $\mathcal{C}$. We denote by $\mathfrak{d}(\boldsymbol{v}, \boldsymbol{w})$ the Euclidean distance between any two points $\boldsymbol{v} = (v_1, \ldots, v_d)$ and $\boldsymbol{w} = (w_1, \ldots, w_d)$.

The smallest distance between two points in $\mathcal{W}$ is equal to 1, which is always met between two neighbor points, while the largest is attained by the opposite end points of the longest diagonals. Such points are $\boldsymbol{v} = (0, \ldots, 0)$ and $\boldsymbol{w} = (N, \ldots, N)$ and the distance between them is $\mathfrak{d}(\boldsymbol{v}, \boldsymbol{w}) = \sqrt{d \cdot N^2} = Nd^{1/2}$. Then, it is natural to normalize $\mathfrak{d}(\boldsymbol{v}, \boldsymbol{w})$ to obtain *the normalized Euclidean distance* $\mathfrak{d}_d(\boldsymbol{v}, \boldsymbol{w})$, for which all normalized distances between points

Date: April 8, 2022.
2010 *Mathematics Subject Classification.* 11B99; 11K99; 11P21; 51M20; 52Bxx.
Key words and phrases. Hypercube, visible points, polytope, Euclidean distance.

in $\mathcal{W}$ will belong to the interval $[0,1]$. Thus,

$$\mathfrak{d}_d(\boldsymbol{v},\boldsymbol{w}) := \frac{1}{Nd^{1/2}}\left(\sum_{n=1}^{d}(w_n-v_n)^2\right)^{1/2}$$

and

$$\mathfrak{d}(\boldsymbol{v},\boldsymbol{w}) = Nd^{1/2}\,\mathfrak{d}_d(\boldsymbol{v},\boldsymbol{w}) = \left(\sum_{n=1}^{d}(w_n-v_n)^2\right)^{1/2}.$$

Denote by $\Omega \subset \mathcal{W}\times\mathcal{W}$ the set of pairs of points that are *visible from each other*, that is, there are no other lattice points in $\mathcal{W}$ between them on the straight line segment that joins them. Then, by definition, Ω is the set of all pairs $(\boldsymbol{v},\boldsymbol{w}) \in \mathcal{W}\times\mathcal{W}$ such that

$$\gcd(v_1-w_1,\dots,v_d-w_d) = 1. \tag{1.1}$$

We show that the normalized distance between almost any two points in $\mathcal{W}$ that are visible from each other is as close to $1/\sqrt{6} \approx 0.40825$ as one wishes, if the dimension d is sufficiently large and N is large enough with respect to d.

Theorem 1. *For any $\varepsilon > 0$, there exists an integer $C(\varepsilon) \geq 3$ such that for any integers $d \geq C(\varepsilon)$ and $N \geq C(\varepsilon)d$ we have:*

$$\frac{1}{\#\Omega}\cdot\#\left\{(\boldsymbol{v},\boldsymbol{w})\in\Omega \;:\; \mathfrak{d}_d(\boldsymbol{v},\boldsymbol{w}) \in \left[\frac{1}{\sqrt{6}}-\varepsilon, \frac{1}{\sqrt{6}}+\varepsilon\right]\right\} \geq 1-\varepsilon\,. \tag{1.2}$$

As a consequence of Theorem 1 we see that almost all triangles with vertices visible from each other are almost equilateral, almost all tetrahedrons with vertices visible from each other are almost regular and so on.

In general, for any $K \geq 2$ let us denote by Ω_K the set of K-polytopes with the property that any two of its vertices are visible from each other. Said differently, if we call *self-visible* a K-polytope with the property that from any of its vertices one can see all the others without any obstruction from any of the lattice points in $\mathcal{W}$, then

$$\Omega_K := \{P = (\boldsymbol{v}_1,\dots,\boldsymbol{v}_d) \in \mathcal{W}^K \;:\; P \text{ is self-visible}\}.$$

Then, essentially, Theorem 1 can be restated in the following form, which is, at the same time, a consequence and a more general form of it.

Corollary 1. *Let $K \geq 2$ be a fixed integer. Then, for any $\varepsilon > 0$, there exists an integer $C(K,\varepsilon) \geq 3$, such that for any integers $d \geq C(K,\varepsilon)$ and $N \geq C(K,\varepsilon)d$, the proportion of polytopes $P \in \Omega_K$ for which*

$$\mathfrak{d}_d(\boldsymbol{w}',\boldsymbol{w}'') \in \left[1/\sqrt{6}-\varepsilon, 1/\sqrt{6}+\varepsilon\right]$$

for all distinct $\boldsymbol{w}',\boldsymbol{w}'' \in P$ is greater than $1-\varepsilon$.

The next theorem answers the question of whether there is a limit probability that a K-polytope in $\mathcal{W}^K$ is self-visible.

Theorem 2. *Let $d \geq 2$, $N \geq 2$ and $2 \leq K \leq 2^d$ be integers. Then, the probability that a K-polytope is self-visible is*

$$\frac{\#\Omega_K}{\#\mathcal{W}^K} = \prod_{p\ prime}\left(1-\frac{1}{p^d}\right)\cdots\left(1-\frac{K-1}{p^d}\right) + O\left(\frac{dK}{N^{1/2}}\right) + O\left(\frac{2^dK^2}{\log^{d-1}N}\right), \tag{1.3}$$

and the implied constants in the big O terms are absolute.

The infinite product over all primes in (1.3) is convergent and defines an endless square of constants

$$\Lambda_{d,K} := \prod_{p \text{ prime}} \prod_{k=1}^{K-1} \left(1 - \frac{k}{p^d}\right),$$

which increase with d and decrease with K. They comprise some remarkable numbers for which Theorem 2 gives a probabilistic geometric interpretation. For instance, if $K = 2$, then $\Lambda_{2,2} = 1/\zeta(2) = 6/\pi^2 \approx 0.6079271$ and $\Lambda_{d,2} = \zeta(d)^{-1}$ for $d \geq 2$. If $d = 2$ and $K = 3$, then

$$\Lambda_{2,3} = \frac{6}{\pi^2}(2C_{FT} - 1) \approx 0.196138,$$

where

$$C_{FT} := \frac{1}{2}\left(1 + \frac{1}{\zeta(2)} \prod_{p} \left(1 - \frac{1}{p^2 - 1}\right)\right) \approx 0.661317$$

is the Feller-Tornier constant (see Feller and Tornier [10] and sequence A065493 from OEIS [20]), which is related to the prime zeta function.

Let us note that if $N = 0$ then $\mathcal{W}^K$ and Ω_K are reduced to a single point. If $N = 1$, all the points in $\mathcal{W}$ are vertices. Then all points in $\mathcal{W}^K$ are self-visible, because there are no intermediary points in the hypercube's lattice which blind the view from one point to another. Then, if $N = 1$, $\frac{\#\Omega_K}{\#\mathcal{W}^K} = 1$ for $2 \leq K \leq 2^d$.

We also remark that for any $N > 1$, if $K > 2^d$ then in (1.3) both non-error terms, the one from the left side and the main term on the right side, are equal to zero. In this case (1.3) holds true with no error terms.

Let $\vec{\boldsymbol{v}}$ denote the ray that starts at the origin $\mathbf{0} = (0, \ldots, 0) \in \mathcal{W}$ and passes through $\boldsymbol{v} \in \mathcal{W}$. Denote by Ψ the set of pairs of rays $(\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}})$ with $(\boldsymbol{v}, \boldsymbol{w}) \in \Omega$. The next result shows that if d and N/d are sufficiently large then almost all angles between rays from the origin towards points that are visible to each other have the sine almost equal to $\sqrt{7}/4$.

Theorem 3. *For any $\varepsilon > 0$, there exists an integer $C(\varepsilon) \geq 3$ such that for any integers $d \geq C(\varepsilon)$ and $N \geq C(\varepsilon)d$ we have:*

$$\frac{1}{\#\Psi} \cdot \# \left\{ (\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}) \in \Psi \ : \ \sin\left(\widehat{\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}}\right) \in \left[\frac{\sqrt{7}}{4} - \varepsilon, \frac{\sqrt{7}}{4} + \varepsilon\right] \right\} \geq 1 - \varepsilon\,. \tag{1.4}$$

Note that the statement in Theorem 3 does not depend on normalization, because the angles are preserved regardless of any scaling.

For any two subsets $\mathscr{M}', \mathscr{M}'' \subseteq \mathcal{W}$, let $\sigma(\mathscr{M}', \mathscr{M}'')$ be the visual *spectrum*, which we define to be the set of sines of all angles between distinct rays that start from the origin towards the points in $\mathscr{M}'$ and $\mathscr{M}''$, that is, denoting identically a point $\mathfrak{m}$ and the ray from the origin towards $\mathfrak{m}$,

$$\sigma(\mathscr{M}', \mathscr{M}'') := \left\{\sin(\mathfrak{m}', \mathfrak{m}'') \ : \ \mathfrak{m}' \in \mathscr{M}', \mathfrak{m}'' \in \mathscr{M}'', \ \mathfrak{m}' \neq \mathfrak{m}''\right\}. \tag{1.5}$$

If $\mathscr{M}' = \mathscr{M}'' = \mathscr{M}$, we write shortly $\sigma(\mathscr{M})$ instead of $\sigma(\mathscr{M}, \mathscr{M})$. The angles between rays from the origin to points in $\mathcal{W}$ cover a larger and larger set of possibilities as the dimension increases and in the limit, as $d \to \infty$, there is a limit set of the spectrum $\sigma(\mathscr{W})$, which is the interval $[0, 1]$. On top of that, choosing elements of $\sigma(\mathcal{W})$ is a random variable, which, by Theorem 3, has a limit probability density function $f(t)$ that is discrete and concentrated in a single point, and $f(t) = \delta\big(t - \frac{\sqrt{7}}{4}\big)$, for $0 \leq t \leq 1$, where δ is the Dirac distribution.

Since most points in $\mathcal{W}$ are visible from each other, the results in Theorems 1, 3 and Corollary 1 may prove useful to check particularities related to randomness of large set of

data. A related result about points on hyperspheres appearred in a theoretical context in the theory of neural network (see Bubeck and Sellke [4]).

In different contexts in nature, it happens and it is not uncommon for a property that is proved to be valid for almost overall objects in a certain universe to be difficult or even impossible to build or to indicate just a single instance that satisfy it. However, in the context of the hypercube lattice $\mathcal{W}$, we can extract some distinguished polytopes that offer a cross-section view of its inner structure.

Let $\mathscr{C} = \{\mathfrak{c}_0, \mathfrak{c}_1, \dots, \mathfrak{c}_d\}$ to be the set of points whose coordinates are the rows of the circular symmetric matrix

$$M_{\mathscr{C}} = \begin{bmatrix} 0 & 1 & \cdots & d-1 & d \\ 1 & 2 & \cdots & d & 0 \\ 2 & 3 & \cdots & 0 & 1 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ d-1 & d & \cdots & d-3 & d-2 \\ d & 0 & \cdots & d-2 & d-1 \end{bmatrix} \tag{1.6}$$

and let $\mathscr{G} := \{\mathfrak{g}_0, \dots, \mathfrak{g}_{p-1}\}$ be the set of points whose components are the rows of the matrix

$$M_{\mathscr{G}} = \begin{bmatrix} 0 & 1^{-1} & \cdots & (p-2)^{-1} & (p-1)^{-1} \\ 1^{-1} & 2^{-1} & \cdots & (p-1)^{-1} & 0 \\ 2^{-1} & 3^{-1} & \cdots & 0 & 1^{-1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ (p-2)^{-1} & (p-1)^{-1} & \cdots & (p-4)^{-1} & (p-3)^{-1} \\ (p-1)^{-1} & 0 & \cdots & (p-3)^{-1} & (p-2)^{-1} \end{bmatrix} \tag{1.7}$$

Here p is prime, the classes of the representatives of the inverses in $M_{\mathscr{G}}$ are taken from $\{1, \dots, p-1\}$ and for symmetry, by convention, we may set 0 to be 'the inverse' of 0.

Theorem 4. *Let $d \geq 2$ and p be prime. Then we have:*

1. Any point in $\mathscr{C} \cup \mathscr{G}$ is visible from the origin. Any two points in $\mathscr{G}$ are visible from each other. If $d = p-1$ and p is large enough any two points in $\mathscr{C} \cup \mathscr{G}$ are visible from each other.

2. The limit set of the normalized distances between points in $\mathscr{C}$ is the interval $[0, 1/2]$, as d tends to infinity.

3. The limit set of the normalized distances between points in $\mathscr{G}$ consists of the single point $\{1/\sqrt{6}\}$, as p tends to infinity.

4. If $d = p-1$, the limit set of the normalized distances between points in $\mathscr{C}$ and points in $\mathscr{G}$ is also $\{1/\sqrt{6}\}$, as p tends to infinity.

5. The limit of the spectrum $\sigma(\mathscr{C})$ is the interval $[0, \sqrt{39}/8]$, as d tends to infinity.

6. The limit of the spectrum $\sigma(\mathscr{G})$ consists of the single point $\{\sqrt{7}/4\}$, as p tends to infinity.

7. If $d = p-1$, the limit of the spectrum $\sigma(\mathscr{C}, \mathscr{G})$ consists of the single point $\{\sqrt{7}/4\}$, also, as p tends to infinity.

Note the decimal approximation of the size of the spectra in Theorem 4:
$\sqrt{7}/4 \approx 0.661438$, $\arcsin(\sqrt{7}/4) \approx 0.72273$ radians or $\approx 41.40962°$; and
$\sqrt{39}/8 \approx 0.78063$, $\arcsin(\sqrt{39}/8) \approx 0.89566$ radians or $\approx 51.31781°$.

As one can see from Theorem 4, about a quarter of all distances $\mathfrak{d}_d(\boldsymbol{v}, \boldsymbol{w})$ with $\boldsymbol{v}, \boldsymbol{w} \in \mathscr{C} \cup \mathscr{G}$ are singular, being different from the other three quarters which in the limit are all

equal to $1/\sqrt{6}$. However, in Subsection 2.4 we prove that if $d = p - 1$ then $A_p(\mathscr{C} \cup \mathscr{G})$, the average of all normalized distances between points in $\mathscr{C} \cup \mathscr{G}$, is still the same $1/\sqrt{6}$, in the limit as $p \to \infty$.

In Section 2.5 we complement the phenomena observed in Theorem 4 with yet another polytope $\mathscr{B}$ whose vertices are visible from each other, even though they lie all almost aligned on a straight line. This set wraps around the diameter of $\mathcal{W}$ being composed by close neighbors of the equally spaced points on the longest diagonal of the cube. The limit set of the normalized distances between the points in this example equals the full interval $[0, 1]$, if their number tends to infinity. The same goes for the limit of the spectrum $\sigma(\mathscr{B})$, which is the interval $[0, 1]$, also. Taken together, merged into a geometric spindle shape, the polytopes $\mathscr{C}, \mathscr{G}$ and $\mathscr{B}$ combine their arithmetic and probabilistic properties keeping in balance the spinning top intrinsic qualities of the still lattice hypercube $\mathcal{W}$.

The paper is organized as follows. In Section 2 we present the special polytopes $\mathscr{C}, \mathscr{G}$ and $\mathscr{B}$, checking the visibility and the mutual distances between their vertices, which proves Theorem 4. In Section 3 we turn to the visibility in the whole lattice and calulate the size of Ω. In Section 4 we find the average of the distances between points in $\mathcal{W}$ that are visible from each other and in Section 5 we estimate the second moment about their mean. We use these results in Section 6 to obtain effective results that in particular prove Theorems 1, 3 and Corollary 1. In Section 7 we discuss at large the problem of self-visible K-polytopes and prove Theorem 2. We conclude in Section 8 with a possible good place to start. It is a short heuristics that might be useful to adjust with the little peculiarities that appear from higher dimensions. It is an intuitive touch on the matter, although it is done in a continuous, where visibility has no meaning, unlike the discrete universe which we explore beyond.

2. Three distinguished polytopes. Their edges and diagonals.

Here we discuss three examples of polytopes with a number of vertices of order almost equal in size with the dimension of the hypercube. The polytopes $\mathscr{C}$ and $\mathscr{G}$ whose vertices are the rows of the matrices $M_{\mathscr{C}}$ and $M_{\mathscr{G}}$ introduced by (1.6) and (1.7) are as similar in construction as they are very different in shape. The first has the distances between its vertices well spread over a long interval, while the second has all the distances between the vertices approximately equal, being as 'equilateral' as it could be, as $d \to \infty$.

2.1. **Proof of Theorem 4 – visibility.**

Notice first that by the definition of $\mathscr{C}$ and $\mathscr{G}$ and that of visibility (1.1), all points in $\mathscr{C} \cup \mathscr{G}$ are visible from the origin.

2.1.1. *Any two points belonging to either $\mathscr{C}$ or $\mathscr{G}$ are visible from each other.* Let us suppose $d = p - 1$ and let $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{W}$. Also, suppose that the coordinates of $\boldsymbol{v}$ are a permutation of $\{0, 1, \dots, d\}$ and the coordinates of $\boldsymbol{w}$ are a circular rotation of the coordinates of $\boldsymbol{v}$. Then, if $\boldsymbol{v}$ and $\boldsymbol{w}$ were not visible from each other, then it would exist an integer $b \geq 2$ such that the following congruences would hold:

$$v_0 - w_0 \equiv 0 \,(\mathrm{mod}\, b)\,;\ v_1 - w_1 \equiv 0 \,(\mathrm{mod}\, b)\,;\ \dots\,;\ v_d - w_d \equiv 0 \,(\mathrm{mod}\, b)\,. \tag{2.1}$$

But since p is prime and $v_0, \dots, v_d$ are all distinct, as $w_0, \dots, w_d$ also are, and since they belong to the same set of numbers, $\{0, 1, \dots, d\}$, which appear each in exactly two of the congruences in (2.1), we find that

$$v_0 \equiv v_1 \equiv \cdots \equiv v_d \equiv r \,(\mathrm{mod}\, b)\,,$$

for some $r \in \{0, 1, \dots, b-1\}$. But this is impossible, unless $b = 1$. Therefore, two points that belong to one and the same set, be it either $\mathscr{C}$ or $\mathscr{G}$ are visible from each other.

2.1.2. *Any two points $\mathfrak{c} \in \mathscr{C}$ and $\mathfrak{g} \in \mathscr{G}$ are visible from each other.* Suppose $d = p-1$ and let $\mathfrak{c} \in \mathscr{C}$ and $\mathfrak{g} \in \mathscr{G}$. Denote by $g = \gcd(v_1 - w_1, \dots, v_d - w_d)$ the expression that needs to be checked in the condition (1.1). Since g is invariant under the same circular rotation applied to both $\mathfrak{c} \in \mathscr{C}$ and $\mathfrak{g} \in \mathscr{G}$, we may assume that

$$\begin{aligned} \mathfrak{g} &= (0,\ 1^{-1},\ \dots, (p-1-h)^{-1}, (p-h)^{-1}, (p+1-h)^{-1}, \dots, (p-1)^{-1}\ ); \\ \mathfrak{c} &= (h, h+1, \dots,\quad p-1,\quad 0,\quad 1,\quad \dots,\ h-1)\,, \end{aligned} \tag{2.2}$$

for some $h \in \{0, 1, \dots, p-1\}$. As usual, the coordinates are taken as their representatives modulo p in the interval $[0, p-1]$. Since $1^{-1} \equiv 1 \pmod p$ and $(p-1)^{-1} \equiv p-1 \pmod p$, the difference between the second components of $\mathfrak{c}$ and $\mathfrak{g}$ is $(h+1) - 1^{-1} = h$ and the difference between the last components is $(h-1) - (p-1)^{-1} = h - p$. Because p is prime, these differences are relatively prime, so that $\mathfrak{c}$ and $\mathfrak{g}$ are visible from each other unless $h = 0$.

Suppose now that $h = 0$ in (2.2). Then, if $\mathfrak{c}$ is not visible from $\mathfrak{g}$, there exists a prime number $2 \le q \le p$ such that $q \mid (n - n^{-1})$ for $1 \le n \le p-1$. Let us notice that for any $0 \le r \le p-1$

$$\#\left\{1 \le n \le p-1\ :\ \left|n - n^{-1}\right| = r\right\} \le 4.$$

This is because $n - n^{-1} \equiv \pm r \pmod p$ is equivalent to $n^2 \mp rn - 1 \equiv 0 \pmod p$, congruence that has at most two solutions for each sign. Even more precise, if $r \neq 0$, if the congruence $n - n^{-1} \equiv r \pmod p$ has solution a then it has solution $-a^{-1}$, also. These solutions are always distinct unless $a \equiv -a^{-1} \pmod p$, which happens only if $p \equiv 1 \pmod 4$, in which case $a = \left(\frac{p-1}{2}\right)!$. If $r = 0$, there are exactly three values of $n \in \{0, 1, \dots, p-1\}$ for which $n - n^{-1} \equiv r \pmod p$, namely $0, 1, p-1$. In conclusion, putting together these observations, while counting separately in the cases $p = 2$, $p \equiv 1 \pmod 4$ and $p \equiv 3 \pmod 4$, we obtain in all cases the same number of distinct absolute values of differences

$$\#\left\{\left|n - n^{-1}\right|\ :\ 1 \le n \le p-1\right\} = \left\lfloor \frac{p}{4} \right\rfloor + 1. \tag{2.3}$$

Then, since

$$\left\{\left|n - n^{-1}\right|\ :\ 1 \le n \le p-1\right\} \subset \{0, 1, \dots, p-1\}$$

and by our assumption a prime q divides all differences $n - n^{-1}$, it follows that q has to be either 2 or 3.

If $q = 2$, the equality (2.3) says that the number of pairs (n, n^{-1}), $1 \le n \le p-1$ of the same parity is $\lfloor p/4 \rfloor + 1$. But this is not in agreement with Lehmer's conjecture [12, Problem F12], which is proved also for shorter general arithmetic progressions [8, Theorem 1]). A particular case of that result shows that if $I, J \subset \{1, \dots, p-1\}$ are arithmetic progressions of ratios $d_1, d_2 \ge 1$, then

$$\#\left\{(a, b) \in I \times J\ :\ ab \equiv 1 \pmod p\right\} = \frac{\#I \cdot \#J}{p} + O\left(p^{1/2} \log^2 p\right). \tag{2.4}$$

Then, if $d_1 = d_2 = 2$, counting the pairs (n, n^{-1}) with either both even or both odd components, we see that their total number is $p/2 + O\left(p^{1/2} \log^2 p\right)$, which contradicts (2.3). Likewise, in the remaining case $q = 3$, with $d_1 = d_2 = 3$, counting the pairs (n, n^{-1}) whose components both give the same remainder 0, 1 or 2 when dividing by 3, we find that their total number is $p/3 + O\left(p^{1/2} \log^2 p\right)$, which is also different from (2.3). In conclusion, $\mathfrak{c}$ and $\mathfrak{g}$ are visible from each other, which concludes the proof of the first part of Theorem 4.

2.2. **Proof of Theorem 4 – the distances.**

2.2.1. *Distances between points of $\mathscr{C}$.* Remark that the same rotation applied to the coordinates of two points, does not change the distance between them, which, in particular, shows that we have

$$\mathfrak{d}(\mathfrak{c}_k, \mathfrak{c}_l) = \mathfrak{d}(\mathfrak{c}_0, \mathfrak{c}_{l-k}), \text{ for any } 0 \le k \le l \le d. \tag{2.5}$$

This means that the range of values of all distances between any two points in $\mathscr{C}$ is covered by the distances between $\mathfrak{c}_0$ and each of $\mathfrak{c}_1, \ldots, \mathfrak{c}_d$. A straightforward calculation shows that in closed form these are:

$$\begin{aligned}
&\mathfrak{d}(\mathfrak{c}_0, \mathfrak{c}_1) = \big((d-0)\cdot 1^2 + 1\cdot(d-0)^2\big)^{1/2}\\
&\mathfrak{d}(\mathfrak{c}_0, \mathfrak{c}_2) = \big((d-1)\cdot 2^2 + 2\cdot(d-1)^2\big)^{1/2}\\
&\mathfrak{d}(\mathfrak{c}_0, \mathfrak{c}_3) = \big((d-2)\cdot 3^2 + 3\cdot(d-2)^2\big)^{1/2}\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\mathfrak{d}(\mathfrak{c}_0, \mathfrak{c}_d) = \big((d-(d-1))\cdot d^2 + d\cdot(d-(d-1))^2\big)^{1/2}.
\end{aligned} \tag{2.6}$$

Not all of these numbers are distinct, because of the symmetry of the parabola: $x \mapsto f(x)$, where $f(x) = (d-(x-1))\cdot x^2 + x\cdot(d-(x-1))^2 = (d+1)x(d+1-x)$, for $x \in [0, d+1]$. The maximum of $f(x)$ is attained for $x = (d+1)/2$ and it is equal to $(d+1)^3/4$. Also, $f(x) = f(d+1-x)$ and the values of $f(x)$ on the integers between 0 and $d+1$ cover quite uniformly the interval $[0, (d+1)^3/4]$ as d becomes sufficiently large. Precisely, for any $y \in [0, 1/2]$ and any $\varepsilon > 0$, there are $\mathfrak{c}', \mathfrak{c}'' \in \mathscr{C}$ such that $y - \varepsilon < \mathfrak{d}_d(\mathfrak{c}', \mathfrak{c}'') < y + \varepsilon$. We summarize in the following proposition these remarks on the polytope $\mathscr{C}$.

Proposition 2.2.1. *Let $\mathscr{C} = \{\mathfrak{c}_0, \mathfrak{c}_1, \ldots, \mathfrak{c}_d\}$ be the set of points whose coordinates are the rows of matrix (1.6). Then, the set of normalized distances between any two points in $\mathscr{C}$ is equal to*

$$\mathcal{D}(\mathscr{C}) := \left\{ \frac{\sqrt{(d+1)x(d+1-x)}}{d^{3/2}} \ : \ 0 \le x \le \left\lfloor \frac{d+1}{2} \right\rfloor \right\} \tag{2.7}$$

and the set $\mathcal{D}(\mathscr{C})$ is dense in the interval $[0, 1/2]$ as d tends to infinity.

2.2.2. *Distances between points of $\mathscr{G}$.* Let p be a prime number and $N = d = p-1$. The polytope $\mathscr{G} = \{\mathfrak{g}_0, \ldots, \mathfrak{g}_{p-1}\}$ is formally close to $\mathscr{C}$. The components of the points are the same, except that the numbers are inverted modulo p. The classes of the representatives of the inverses are taken from $\{1, \ldots, p-1\}$ and by convention the inverse of an integer divisible by p, which does not exist, is always replaced by 0. Let us remark that the influence of just a single component in the first point, while the others are obtained by circular rotations as in $\mathscr{G}$, has small and even negligible influence as $p \to \infty$, on the mutual distances between the points in $\mathscr{G}$. This is why we could keep, for balance, in $\mathscr{G}$ the components zero, even if zero has no inverse modulo p.

The main motivation for choosing $\mathscr{G}$ is the random spread of the inverses. Various ways to measure the randomness of inverses, triggered by [12, Problem F12], have been studied in [6–8]. In [23] the focus is on the values of polynomials and rational functions mod p and the results there might be also used to build other polytopes with similar characteristics.

Lemma 2.1. *If p is prime and $1 \le h \le p-1$, we have*

$$\sum_{x \in \mathbb{F}_p \setminus \{0, p-h\}} \left|(x+h)^{-1} - x^{-1}\right|^2 = \frac{p^3}{6} + O\left(p^{5/2}\log^2 p\right). \tag{2.8}$$

Here the inverses are calculated in $\mathbb{F}_p$ and the absolute value calculates the distance between two natural numbers in $\{0,1,\dots,p-1\}$, the corresponding representatives of the residue classes mod p of the inverses.

For $h=1$, the estimate (2.8) is a particular case of [23, Corollary 1.3], which is proved using Weil's bounds [22] for exponential sums with rational functions. For $h\geq 2$ the estimate (2.8) can be proved in a similar manner.

Taking the square root and dividing by $p^{3/2}$, we find by Lemma 2.1, that the normalized distance between any two points $\mathfrak{g}',\mathfrak{g}''\in\mathscr{G}$ is

$$\mathfrak{d}_d(\mathfrak{g}',\mathfrak{g}'')=\frac{1}{\sqrt{6}}\left(1+O\left(\frac{\log^2 p}{\sqrt{p}}\right)\right), \tag{2.9}$$

which proves part 3 of Theorem 4.

2.2.3. *Distances between $\mathfrak{c}\in\mathscr{C}$ and $\mathfrak{g}\in\mathscr{G}$.* It suffices to find the distance between $\mathfrak{c}$ and $\mathfrak{g}$ in (2.2). For this we have to estimate the sums

$$\mathfrak{d}^2(\mathfrak{g},\mathfrak{c})=\sum_{n=0}^{p-h-1}\left|(n+h)-n^{-1}\right|^2+\sum_{n=p-h}^{p-1}\left|(n+h-p)-n^{-1}\right|^2=\Sigma'_h+\Sigma''_h, \tag{2.10}$$

where Σ'_h and Σ''_h are the first and the second sum in (2.10), respectively. To calculate Σ'_h, let $L>1$ be fixed, denote $u=(p-h)/L$, $v=p/L$ and split the rectangle $[0,p-h]\times[0,p]$ into L^2 rectangles $T_{j,k}:=\big[ju,(j+1)u\big)\times\big[kv,(k+1v)\big)$. Then

$$\Sigma'_h=\sum_{j=0}^{L-1}\sum_{k=0}^{L-1}\sum_{(n,n^{-1})\in T_{j,k}}\left(n+h-n^{-1}\right)^2.$$

Here the size of the summand can be kept under control, so that we can replace it by its value on the lower left corner of $T_{j,k}$. Thus, on using (2.4) with $I\times J=T_{j,k}$, we have

$$\Sigma'_h=\sum_{j=0}^{L-1}\sum_{k=0}^{L-1}\left(ju+h-kv+O(p/L)\right)^2\left(\frac{uv}{p}+O\left(p^{1/2}\log^2 p\right)\right).$$

Now we factor the terms that do not depend on j,k and expand the square

$$\Sigma'_h=\left(\frac{uv}{p}+O\left(p^{1/2}\log^2 p\right)\right)\sum_{j=0}^{L-1}\sum_{k=0}^{L-1}g(j,k,u,v,p,L) \tag{2.11}$$

where

$$\begin{aligned}g(j,k,u,v,p,L)=&j^2u^2+h^2+k^2v^2+2juh-2jkuv-2hkv\\&+O(p^2/L^2)+O\Big(p(ju+h+kv)/L\Big).\end{aligned}$$

Adding together the terms separately over k and j, the sum of powers being denoted by $S_r(M)=1^r+\cdots+M^r$, and then collecting together the error terms, the double sum

from (2.11) becomes

$$\begin{aligned}\sum_{j=0}^{L-1}\sum_{k=0}^{L-1} g(j,k,u,v,p,L) =& Lu^2S_2(L-1)+L^2h^2+Lv^2S_2(L-1)\\ &+2LuhS_1(L-1)-2uvS_1^2(L-1)-2LhvS_1(L-1)\\ &+O(p^2)+O\Big(p\big(LuS_1(L)+hL^2+LvS_1(L)\big)/L\Big)\\ =&\frac{L^4u^2}{3}+L^2h^2+\frac{L^4v^2}{3}+L^3uh-\frac{L^4uv}{2}-L^3vh+O\left(p^2L\right).\end{aligned}$$

Next we insert this into (2.11), replace u,v by their definition and reduce the terms:

$$\begin{aligned}\Sigma'_h &= \Big(\frac{uv}{p}+O\big(\sqrt{p}\log^2 p\big)\Big)\left(\frac{L^4(u^2+v^2)}{3}+L^2h^2-\frac{L^4uv}{2}-L^3h(v-u)+O\left(p^2L\right)\right)\\ &= \Big(p-h+O\big(L^2\sqrt{p}\log^2 p\big)\Big)\left(\frac{p^2+(p-h)^2}{3}-\frac{p(p-h)}{2}+O\left(p^2/L\right)\right)\\ &= \frac{(p-h)(p^2+2h^2-hp)}{6}+O(p^3/L)+O\left(L^2p^{5/2}\log^2 p\right).\end{aligned}\tag{2.12}$$

To estimate Σ''_h, we make the change of variables $m=p-n$. Note that the representative in the interval $[0,p-1]$ of the inverse of m is $m^{-1}=p-n^{-1}$. Then

$$\Sigma''_h=\sum_{n=p-h}^{p-1}\left|(n+h-p)-n^{-1}\right)|^2=\sum_{m=1}^{h}\left|(m+p-h)-m^{-1}\right|^2.$$

Apart from the end limits of summation, this is exactly Σ'_{p-h}. Therefore we have

$$\Sigma''_h=\Sigma'_{p-h}+O(p^2).\tag{2.13}$$

On combining the estimate (2.12) with (2.13) and inserting the results into (2.10), we obtain

$$\begin{aligned}\mathfrak{d}^2(\mathfrak{g},\mathfrak{c}) &= \frac{(p-h)(p^2+2h^2-hp)+h(2p^2+2h^2-3hp)}{6}+O\left(p^3/L+L^2p^{5/2}\log^2 p\right)\\ &= \frac{p^3}{6}+O\left(p^3/L+L^2p^{5/2}\log^2 p\right).\end{aligned}\tag{2.14}$$

Balancing the error terms, we find the optimal L that we have fixed at the beginning, namely $L=\left\lfloor p^{1/6}\log^{-2/3}p\right\rfloor$. Thus we have proved the following result.

Proposition 2.2.2. *Let $d\geq 2$ and let p be prime such that $d=p-1$. Let $\mathscr{C}=\{\mathfrak{c}_0,\mathfrak{c}_1,\dots,\mathfrak{c}_d\}$ be the set of points whose coordinates are the rows of matrix $M_{\mathscr{C}}$ from (1.6) and let $\mathscr{G}=\{\mathfrak{g}_0,\mathfrak{g}_1,\dots,\mathfrak{g}_d\}$ be the set of points whose coordinates are the rows of matrix $M_{\mathscr{G}}$ from (1.7). Then, as p tends to infinity, the limit set of the normalized distances between any two points $\mathfrak{c}\in\mathscr{C}$ and $\mathfrak{g}\in\mathscr{G}$ is the single point $\{1/\sqrt{6}\}$ and*

$$\mathfrak{d}_d(\mathfrak{g},\mathfrak{c})=\frac{1}{\sqrt{6}}+O\left(p^{-1/6}\log^{2/3}p\right).\tag{2.15}$$

Proposition 2.2.2 proves part 4 of Theorem 4.

2.3. **Proof of Theorem 4 – the spectra.**

From the origin, which we denote by $\mathbf{0}$, the normalized distances toward points in $\mathscr{C}$ or $\mathscr{G}$ are equal

$$\mathfrak{d}_d(\mathbf{0},\mathfrak{d}_h)=\mathfrak{d}_d(\mathbf{0},\mathfrak{c}_h)=\left(\frac{d(d+1)(2d+1)}{6}\right)^{1/2}d^{-3/2}=\frac{1}{\sqrt{3}}+O(1/d), \tag{2.16}$$

for $0\le h\le d$, if $p=d+1$. The distances between points in $\mathscr{C}$ are given by Proposition 2.2.1. Since the set of distances between points in $\mathscr{C}$ are in the limit, as $d\to\infty$, dense in the interval $[0,1/2]$, the limit of the spectrum $\sigma(\mathscr{C})$ is also a continuous interval. The end points of the limit $\sigma(\mathscr{C})$ come from the limit angles of the isosceles triangles with vertices in $\mathbf{0}$ having two edges equal to $1/\sqrt{3}+O(1/d)$, according to (2.16), while the third edge is the shortest and the longest distance between points in $\sigma(\mathscr{C})$, respectively. By (2.7), the acutest triangle has its third edge equal to $O(1/d)$, while the triangle with the largest angle at $\mathbf{0}$ has its third edge equal to $1/2+O(1/d)$. Then a straightforward calculation gives the end points of the limit spectrum 0 and $\sqrt{39}/8$.

If $d=p-1$, the mutual distances between points either from $\mathscr{G}$ or from $\mathscr{C}\cup\mathscr{G}$ are all equal to $1/\sqrt{6}+O\left(p^{-1/6}\log^{2/3}p\right)$, by relation (2.9) and Proposition 2.2.2, respectively. Then, the limit spectra $\sigma(\mathscr{G})$ and $\sigma(\mathscr{C}\cup\mathscr{G}))$ are equal and discrete, containing exactly one point. According to (2.16), the limit point is the sine of the acutest angle of the isosceles triangle with two edges equal to $1/\sqrt{3}$ and the third equal to $1/\sqrt{6}$. Since this is equal to $\sqrt{7}/4$, these concludes the proof of the remaining parts of Theorem 4.

2.4. **The average distance between points in $\mathscr{C}\cup\mathscr{G}$.**

Suppose $d=p-1$. The cardinality of $\mathscr{C}\cup\mathscr{G}$ is $4p^2$ and the average of the squares of distances between its points is $A(\mathscr{C}\cup\mathscr{G})=T/(4p^2)$, where T is the following sum

$$T=\sum_{(\mathfrak{c}',\mathfrak{c}'')\in\mathscr{C}\times\mathscr{C}}\mathfrak{d}^2(\mathfrak{c}',\mathfrak{c}'')+\sum_{(\mathfrak{g}',\mathfrak{g}'')\in\mathscr{G}\times\mathscr{G}}\mathfrak{d}^2(\mathfrak{g}',\mathfrak{g}'')+2\sum_{(\mathfrak{c},\mathfrak{g})\in\mathscr{C}\times\mathscr{G}}\mathfrak{d}^2(\mathfrak{c},\mathfrak{g}). \tag{2.17}$$

We denote by T_1,T_2,T_3 the sums on the right side of (2.17). By (2.6), the first sum is

$$T_1=2\sum_{h=1}^{p-1}(p-h)\cdot\big(ph(p-h)\big)=2p\big(pS_2(p-1)+S_3(p-1)\big)=\frac{p^5}{6}+O\left(p^4\right). \tag{2.18}$$

By (2.9) and (2.15) the last two sums in (2.17) together are

$$T_2+T_3=6\sum_{h=1}^{p-1}h\left(\frac{p^3}{6}+O\left(p^{17/6}\log^{2/3}p\right)\right)=\frac{p^5}{2}+O\left(p^{29/6}\log^{2/3}p\right). \tag{2.19}$$

Now we can find the normalized distances between points in $\mathscr{C}\cup\mathscr{G}$ which is defined by

$$A_p(\mathscr{C}\cup\mathscr{G}):=\sqrt{A(\mathscr{C}\cup\mathscr{G})}/p^{3/2}. \tag{2.20}$$

Thus, on inserting (2.19) and (2.18) into (2.17), we find that (2.20) becomes

$$\begin{aligned}A_p(\mathscr{C}\cup\mathscr{G})&=\left(\frac{1}{4p^2}\left(\frac{p^5}{6}+\frac{p^5}{2}+O\left(p^{29/6}\log^{2/3}p\right)\right)\right)^{1/2}p^{-3/2}\\&=\frac{1}{\sqrt{6}}+O\left(p^{-1/6}\log^{2/3}p\right).\end{aligned} \tag{2.21}$$

2.5. **A polytope stretched out along the longest diagonal of $\mathcal{W}$.**

On the same theme, we construct a polytope $\mathscr{B} = \{\mathfrak{b}_0, \dots, \mathfrak{b}_d\}$ that is a cousin of $\mathscr{C}$ and $\mathscr{G}$. The polytope $\mathscr{B}$ streches along the diagonal $\big[(0, \dots, 0);\ (d, \dots, d)\big]$, it has all vertices visible from each other and the normalized distances between them are dense in $[0, 1]$, as the dimension d tends to infinity. All components of $\mathfrak{b}_h$ are set to be equal to h, except the $(h-1)$-th and $(h+1)$-th, which are equal to $h-1$ and $h+1$, respectively. Thus

$$\mathfrak{b}_h = (\overset{0\text{-th}}{h},\ h,\ \dots,\ h,\ h-1,\ \overset{h\text{-th}}{h},\ h+1,\ h,\ \dots,\ \overset{d\text{-th}}{h}), \tag{2.22}$$

for $h = 0, 1, \dots, d$, with the convention that $-1 = d$ and $d+1 = 0$.

If $d = 3$ or $d \geq 5$ the points of $\mathscr{B}$ are visible from one another. For points that are not too close this follows because most components of any point are equal, while the neighbors of just one component are the neighbor integers of the rank of the point. For points that are near each other it can also be checked one by one that they are visible from each other.

The distances between the points of $\mathscr{B}$ are:

$$\mathfrak{d}_d(\mathfrak{b}_j, \mathfrak{b}_k) = d^{-3/2}\sqrt{d|k-j|^2 + O(1)} = \frac{|k-j|}{d} + O(1/d), \tag{2.23}$$

for $1 \leq j, k, \leq d$. As a consequence, (2.23) implies that the closure of the set of distances between points in $\mathscr{B}$ equals the full interval $[0, 1]$, as d tends to infinity.

Notice that all points in $\mathscr{B}$ are visible from the origin, because (2.22) assures that condition (1.1) is verified.

The distances from the origin to points in $\mathscr{B}$ are

$$\mathfrak{d}_d(\mathbf{0}, \mathfrak{b}_h) = \frac{h}{d} + O(1/d), \ \text{ for } 1 \leq h \leq d. \tag{2.24}$$

Since the limit set of the mutual distances between points in $\mathscr{B}$ is the full interval $[0, 1]$, on combining (2.23) and (2.24), we see that there is a limit of the spectrum $\sigma(\mathscr{B})$, which is also a closed interval. Its smallest end point comes from the triangle with vertices $\mathbf{0}$, $\mathfrak{b}_{d-2}$ and $\mathfrak{b}_{d-1}$, and its largest from the triangle with vertices $\mathbf{0}$, $\mathfrak{b}_1$ and $\mathfrak{b}_{d-1}$. Then the angle at $\mathbf{0}$ of the first triangle tends to zero, and that of the second triangle tends to $\pi/2$, as $d \to \infty$. As a consequence, the limit of the spectrum $\sigma(\mathscr{B})$ is the interval $[0, 1]$, as d tends to infinity.

3. The number of pairs in Ω

Let $(\boldsymbol{v}, \boldsymbol{w}) \in \Omega$ and suppose $\boldsymbol{v} = (v_1, \dots, v_d)$ and $\boldsymbol{w} = (w_1, \dots, w_d)$. Then $\boldsymbol{v}$ and $\boldsymbol{w}$ are visible from each other. This means that $\gcd(v_1 - w_1, \dots, v_d - w_d) = 1$. We can rewrite this condition by bringing Möbius summation into play. Thus, we have

$$\sum_{\substack{1 \leq b \leq N \\ b | v_1 - w_1 \\ \cdots \\ b | v_d - w_d}} \mu(b) = \begin{cases} 1 & \text{if } (\boldsymbol{v}, \boldsymbol{w}) \in \Omega, \\ 0 & \text{if } (\boldsymbol{v}, \boldsymbol{w}) \notin \Omega. \end{cases} \tag{3.1}$$

We start by finding an estimate for the cardinality of Ω, which is the object of the following lemma.

Lemma 3.1. *There exists an absolute constant $C_0 > 0$, such that for all $d \geq 2$ and all $N \geq 3d$, we have*

$$\left| \#\Omega - \frac{N^{2d}}{\zeta(d)} \right| \leq \begin{cases} C_0 N^3 \log N & \text{if } d = 2, \\ C_0 d N^{2d-1} & \text{if } d \geq 3. \end{cases} \tag{3.2}$$

Proof. By the definition and the counting formula (3.1), by changing the order of summation, we have

$$\#\Omega = \sum_{\boldsymbol{v}\in\mathcal{W}} \sum_{\substack{\boldsymbol{w}\in\mathcal{W}\\(\boldsymbol{v},\boldsymbol{w})\in\Omega}} 1 = \sum_{\boldsymbol{v}\in\mathcal{W}} \sum_{\boldsymbol{w}\in\mathcal{W}} \sum_{\substack{b|(v_1-w_1)\\ \cdots \\ b|(v_d-w_d)}} \mu(d) = \sum_{b=1}^{N} \mu(b) \sum_{\substack{0\le v_1,w_1\le N\\ b|v_1-w_1}} \cdots \sum_{\substack{0\le v_d,w_d\le N\\ b|v_d-w_d}} 1.$$

Since the variables run independently, the summation over $\boldsymbol{v}$ and $\boldsymbol{w}$ can be grouped as a product as follows

$$\#\Omega = \sum_{b=1}^{N} \mu(b) \Bigg(\sum_{\substack{0\le v_1,w_1\le N\\ b|v_1-w_1}} 1 \Bigg) \cdots \Bigg(\sum_{\substack{0\le v_d,w_d\le N\\ b|v_d-w_d}} 1 \Bigg). \tag{3.3}$$

Next, we estimate the inner sums that are equal to each other for all $j \in \{1,\dots,d\}$. Dropping the subscripts, we see that each of them is equal to

$$\begin{aligned} H := \sum_{\substack{0\le v,w\le N\\ b|v-w}} 1 &= \sum_{0\le r\le b-1} \#\{(v,w) : 0\le v,w\le N, v\equiv r \,(\mathrm{mod}\, b), w\equiv r \,(\mathrm{mod}\, b)\} \\ &= \sum_{0\le r\le b-1} \big(\#\{0\le a\le N \;:\; a\equiv r \,(\mathrm{mod}\, b)\}\big)^2. \end{aligned}$$

The cardinality of the inner set is equal to $\lfloor \frac{N-r}{b} \rfloor + 1 = \frac{N}{b} + \theta_1(b,r,N)$, with $|\theta(b,r,N)| \le 1$. Then

$$H = \sum_{0\le r\le b-1} \left(\left\lfloor \frac{N-r}{b} \right\rfloor + 1 \right)^2 = \sum_{0\le r\le b-1} \left(\frac{N^2}{b^2} + \theta_2(b,r,N)\cdot \frac{N}{b} \right) = \frac{N^2}{b} + \theta(b,N)N,$$

for some real numbers for which $|\theta_2(b,r,N)| \le 3$ and $|\theta(b,N)| \le 3$. On inserting this estimate in (3.3), it yields

$$\#\Omega = \sum_{b=1}^{N} \mu(b) H^d = \sum_{b=1}^{N} \mu(b) \left(\frac{N^{2d}}{b^d} + O\left(\sum_{k=1}^{d} \binom{d}{k} \left(\frac{N^2}{b} \right)^{d-k} (3N)^k \right) \right). \tag{3.4}$$

Here, the main term is

$$\sum_{b=1}^{N} \mu(b) \frac{N^{2d}}{b^d} = N^{2d} \sum_{b=1}^{\infty} \mu(b) b^{-d} + O\left(N^{2d} \int_N^{\infty} x^{-d}\, \mathrm{d}\, x \right) = \frac{N^{2d}}{\zeta(d)} + O\left(\frac{N^{d+1}}{d} \right). \tag{3.5}$$

Denoting the error term in (3.4) by E_1 and changing the order of summation we find that

$$|E_1| = O\left(\sum_{b=1}^{N} \sum_{k=1}^{d} \binom{d}{k} \frac{N^{2d-2k}}{b^{d-k}} 3^k N^k \right) = O\left(\sum_{k=1}^{d} 3^k N^{2d-k} \binom{d}{k} T_{d-k}(N) \right), \tag{3.6}$$

where $T_r(N) := 1^{-r} + \cdots + N^{-r}$. Then $T_{d-k}(N) = N$, if $k = d$, $T_{d-k}(N) = \log N$, if $k = d-1$ and $T_{d-k}(N) = O(1)$ if $1 \le k \le d-2$. Then

$$|E_1| = O\left(3^d N^{d+1} \right) + O\left(3^d N^{d+1} d \log N \right) + O\left(N^{2d} \sum_{k=1}^{d-2} 3^k N^{-k} \binom{d}{k} \right), \tag{3.7}$$

The lemma follows by inserting the estimates (3.5) and (3.7) into (3.4).

□

4. The average distance between points visible from each other

The average of the square of distances between points visible from each other is

$$A_{vis}(d,N) := \frac{1}{\#\Omega} \sum_{(\boldsymbol{v},\boldsymbol{w})\in\Omega} \mathfrak{d}^2(\boldsymbol{v},\boldsymbol{w}). \tag{4.1}$$

By (3.1), changing the order summation this is

$$A_{vis}(d,N) = \frac{1}{\#\Omega} \sum_{b=1}^{N} \mu(b) \sum_{\boldsymbol{v}\in\mathcal{W}} \sum_{\substack{\mathbf{w}\in\mathcal{W} \\ b|(v_1-w_1) \\ \cdots \\ b|(v_d-w_d)}} \mathfrak{d}^2(\boldsymbol{v},\boldsymbol{w}). \tag{4.2}$$

We rewrite (4.2) as

$$A_{vis}(d,N) = \frac{1}{\#\Omega} \sum_{b=1}^{N} \mu(b) \sum_{\boldsymbol{v}\in\mathcal{W}} \sum_{\substack{\mathbf{w}\in\mathcal{W} \\ b|(v_1-w_1) \\ \cdots \\ b|(v_d-w_d)}} \big((w_1-v_1)^2+\cdots+(w_d-v_d)^2\big) = \sum_{j=1}^{d} H_j, \tag{4.3}$$

where

$$H_j := \frac{1}{\#\Omega} \sum_{b=1}^{N} \mu(b) \sum_{\boldsymbol{v}\in\mathcal{W}} \sum_{\substack{\mathbf{w}\in\mathcal{W} \\ b|(v_1-w_1) \\ \cdots \\ b|(v_d-w_d)}} (w_j-v_j)^2. \tag{4.4}$$

By changing the order of summation to isolate the part that does not depend on the j variables, H_j can be rewritten as

$$\begin{aligned} H_j &= \frac{1}{\#\Omega} \sum_{b=1}^{N} \mu(b) \Bigg(\sum_{\substack{0\le v_j,w_j\le N \\ b|(v_j-w_j)}} (w_j-v_j)^2 \Bigg) \prod_{\substack{k=1 \\ k\neq j}}^{d} \Bigg(\sum_{\substack{0\le v_k,w_k\le N \\ b|(v_k-w_k)}} 1 \Bigg) \\ &= \frac{1}{\#\Omega} \sum_{b=1}^{N} \mu(b) \left(\frac{N^2}{b} + O(N) \right)^{d-1} \sum_{\substack{0\le v_j,w_j\le N \\ b|(v_j-w_j)}} (w_j-v_j)^2. \end{aligned} \tag{4.5}$$

In the interior sum from (4.5) we group the terms with v_j and w_j in the same residue classes mod b as follows:

$$\begin{aligned} \sum_{\substack{0\le v_j,w_j\le N \\ b|(v_j-w_j)}} (w_j-v_j)^2 &= \sum_{r=0}^{b-1} \sum_{\substack{0\le v_j,w_j\le N \\ v_j\equiv w_j\equiv r \pmod b}} (w_j-v_j)^2 \\ &= \sum_{r=0}^{b-1} \sum_{m=0}^{\lfloor \frac{N-r}{b} \rfloor} \sum_{n=0}^{\lfloor \frac{N-r}{b} \rfloor} \big((r+mb)-(r+nb)\big)^2 \\ &= b^2 \sum_{r=0}^{b-1} \sum_{m=0}^{\lfloor \frac{N-r}{b} \rfloor} \sum_{n=0}^{\lfloor \frac{N-r}{b} \rfloor} (m-n)^2. \end{aligned} \tag{4.6}$$

With $M = \lfloor \frac{N-r}{b} \rfloor$, the sums over m and n are equal to

$$\begin{aligned}\sum_{m=0}^{M}\sum_{n=0}^{M}(m-n)^2 &= 2\sum_{m=0}^{M}\sum_{n=0}^{M}m^2 - 2\sum_{m=0}^{M}\sum_{n=0}^{M}mn \\ &= 2\left(\frac{1}{3}M^4 + O(M^3)\right) - 2\left(\frac{1}{4}M^4 + O(M^3)\right) \\ &= \left(\frac{1}{6}M^4 + O(M^3)\right).\end{aligned} \tag{4.7}$$

Combining (4.7) into (4.6) we find that

$$\sum_{\substack{0\le v_j, w_j \le N \\ b|(v_j - w_j)}} (w_j - v_j)^2 = \frac{N^4}{6b} + O(N^3). \tag{4.8}$$

On inserting this estimate in (4.5), we obtain

$$\begin{aligned}H_j &= \frac{1}{\#\Omega}\sum_{b=1}^{N}\mu(b)\left(\frac{N^2}{b} + O(N)\right)^{d-1}\left(\frac{N^4}{6b} + O(N^3)\right) \\ &= \frac{N^{2d+2}}{6\cdot\#\Omega}\sum_{b=1}^{N}\frac{\mu(b)}{b^d}\left(1 + O\left(\frac{b}{N}\right)\right)^d \\ &= \frac{N^{2d+2}}{6\cdot\#\Omega}\left(\frac{1}{\zeta(d)} + O\left(\frac{1}{dN^{d-1}}\right) + \sum_{b=1}^{N}\frac{1}{b^d}\sum_{k=1}^{d}\binom{d}{k}\left(\frac{b}{N}\right)^k\right).\end{aligned} \tag{4.9}$$

Here the interior sums are

$$\sum_{b=1}^{N}\frac{1}{b^d}\sum_{k=1}^{d}\binom{d}{k}\left(\frac{b}{N}\right)^k = \begin{cases} O\left(\frac{\log N}{N}\right) & \text{if } d = 2, \\ O\left(\frac{d}{N}\right) & \text{if } d \ge 3. \end{cases}$$

Introducing this estimate in (4.9) and the result in (4.3) we summarize in the next lemma the estimate obtained for $A_{vis}(d, N)$.

Lemma 4.1. *We have*

$$A_{vis}(d,N) = \begin{cases} \frac{dN^{2d+2}}{6\#\Omega\zeta(d)}\left(1 + O\left(\frac{\log N}{N}\right)\right) & \text{if } d = 2, \\ \frac{dN^{2d+2}}{6\#\Omega\zeta(d)}\left(1 + O\left(\frac{d}{N}\right)\right) & \text{if } d \ge 3. \end{cases} \tag{4.10}$$

Taking into account the size of the cardinality of Ω evaluated in Lemma 3.1 into (4.10), it yields the following simple estimate for $A_{vis}(d, N)$.

Lemma 4.2. *There exists an absolute constant $C_1 > 0$, such that for all $d \ge 2$ and all $N \ge 3d$, we have*

$$\left|A_{vis}(d,N) - \frac{dN^2}{6}\right| \le \begin{cases} C_1 N\log N & \text{if } d = 2, \\ C_1 d^2 N & \text{if } d \ge 3. \end{cases} \tag{4.11}$$

5. The second moment about the mean

The second moment about the mean $A_{vis}(d,N)$ is the average of the squares of the differences between the expected and the true distance between the pairs of points from $\mathcal{W}$ that are visible from each other, that is,

$$\mathfrak{M}_{2,vis}(d,N) := \frac{1}{\#\Omega}\sum_{\boldsymbol{v}\in\mathcal{W}}\sum_{\substack{\mathbf{w}\in\mathcal{W}\\(\boldsymbol{v},\boldsymbol{w})\in\Omega}}\left|\mathfrak{d}^2(\boldsymbol{v},\boldsymbol{w})-A_{vis}(d,N)\right|^2. \tag{5.1}$$

Replacing the coprimality condition by means of the characteristic function (3.1) and changing the order of summation, we have

$$\mathfrak{M}_{2,vis}(d,N) = \frac{1}{\#\Omega}\sum_{b=1}^{N}\mu(b)\sum_{\boldsymbol{v}\in\mathcal{W}}\sum_{\substack{\mathbf{w}\in\mathcal{W}\\ b|(v_1-w_1)\\ \cdots\\ b|(v_d-w_d)}}\left|\mathfrak{d}^2(\boldsymbol{v},\boldsymbol{w})-A_{vis}(d,N)\right|^2.$$

Next, by expanding the square it yields

$$\begin{aligned}\mathfrak{M}_{2,vis}(d,N) =& \frac{1}{\#\Omega}\sum_{b=1}^{N}\mu(b)\sum_{\boldsymbol{v}\in\mathcal{W}}\sum_{\substack{\mathbf{w}\in\mathcal{W}\\ b|(v_1-w_1)\\ \cdots\\ b|(v_d-w_d)}}\mathfrak{d}^4(\boldsymbol{v},\boldsymbol{w})\\ &-\frac{2A_{vis}(d,N)}{\#\Omega}\sum_{b=1}^{N}\mu(b)\sum_{\boldsymbol{v}\in\mathcal{W}}\sum_{\substack{\mathbf{w}\in\mathcal{W}\\ b|(v_1-w_1)\\ \cdots\\ b|(v_d-w_d)}}\mathfrak{d}^2(\boldsymbol{v},\boldsymbol{w})\\ &+\frac{A_{vis}^2(d,N)}{\#\Omega}\sum_{b=1}^{N}\mu(b)\sum_{\boldsymbol{v}\in\mathcal{W}}\sum_{\substack{\mathbf{w}\in\mathcal{W}\\ b|(v_1-w_1)\\ \cdots\\ b|(v_d-w_d)}}1\\ =&\frac{1}{\#\Omega}\cdot\Sigma_{vis}-A_{vis}^2(d,N).\end{aligned} \tag{5.2}$$

Here we have denoted by Σ_{vis} the multiple sum over $\boldsymbol{v}$ and $\boldsymbol{w}$ from the first row of relation (5.2) and have taken into account the fact that the term from the second row is equal to $-2A_{vis}^2(d,N)$, while the term from the third row is equal to $A_{vis}^2(d,N)$. Next, changing the order of summation, we split Σ_{vis} into d^2 similar sums $H_{j,k}$

$$\Sigma_{vis}=\sum_{b=1}^{N}\mu(b)\sum_{\boldsymbol{v}\in\mathcal{W}}\sum_{\substack{\mathbf{w}\in\mathcal{W}\\ b|(v_1-w_1)\\ \cdots\\ b|(v_d-w_d)}}\sum_{j=1}^{d}\sum_{k=1}^{d}(w_j-v_j)^2(w_k-v_k)^2=\sum_{j=1}^{d}\sum_{k=1}^{d}H_{j,k}, \tag{5.3}$$

where

$$H_{j,k}=\sum_{b=1}^{N}\mu(b)\sum_{\boldsymbol{v}\in\mathcal{W}}\sum_{\substack{\mathbf{w}\in\mathcal{W}\\ b|(v_1-w_1)\\ \cdots\\ b|(v_d-w_d)}}(w_j-v_j)^2(w_k-v_k)^2. \tag{5.4}$$

Now fix $j \neq k$. In each $H_{j,k}$ the summand depends only on four of the $2d$ variables $v_1, \dots, v_d, w_1, \dots, w_d$, so that

$$H_{j,k} = \sum_{b=1}^{N} \mu(b) \Bigg(\sum_{\substack{0 \le v_j, w_j \le N \\ b \mid (v_j - w_j)}} \sum_{\substack{0 \le v_k, w_k \le N \\ b \mid (v_k - w_k)}} (w_j - v_j)^2 (w_k - v_k)^2 \Bigg) \prod_{\substack{s=1 \\ s \neq j \\ s \neq k}}^{d} \Bigg(\sum_{\substack{0 \le v_s, w_s \le N \\ b \mid (v_s - w_s)}} 1 \Bigg). \tag{5.5}$$

Since the products count the number of terms in some arithmetic progressions and are equal, we derive that

$$H_{j,k} = \sum_{b=1}^{N} \mu(b) \left(\frac{N^2}{b} + O(N) \right)^{d-2} \sum_{\substack{0 \le v_j, w_j \le N \\ b \mid (v_j - w_j)}} \sum_{\substack{0 \le v_k, w_k \le N \\ b \mid (v_k - w_k)}} (w_j - v_j)^2 (w_k - v_k)^2 \tag{5.6}$$

By relation (4.8), we find that the interior sums are

$$\sum_{\substack{0 \le v_j, w_j \le N \\ b \mid (v_j - w_j)}} \sum_{\substack{0 \le v_k, w_k \le N \\ b \mid (v_k - w_k)}} (w_j - v_j)^2 (w_k - v_k)^2 = \frac{N^8}{36 b^2} + O\left(N^7 / b\right). \tag{5.7}$$

On combining (5.7) and (5.6), it follows that

$$\begin{aligned} H_{j,k} &= \sum_{b=1}^{N} \mu(b) \left(\frac{N^2}{b} + O(N) \right)^{d-2} \left(\frac{N^8}{36 b^2} + O\left(N^7 / b\right) \right) \\ &= \frac{N^{2d+4}}{36} \sum_{b=1}^{N} \frac{\mu(b)}{b^d} \left(1 + O\left(\frac{b}{N} \right) \right)^d. \end{aligned} \tag{5.8}$$

Following the reasoning from (4.9) and the relation that follows, we obtain the following estimate

$$H_{j,k} = \begin{cases} \frac{N^{2d+4}}{36\zeta(d)} \left(1 + O\left(\frac{\log N}{N} \right) \right) & \text{if } d = 2, \\ \frac{N^{2d+4}}{36\zeta(d)} \left(1 + O\left(\frac{d}{N} \right) \right) & \text{if } d \geq 3. \end{cases} \tag{5.9}$$

If $j = k$, adapting the same steps after relation (5.4) we obtain the upper bound

$$H_{j,j} = O(N^{2d+4}), \text{ for } 1 \leq j \leq d. \tag{5.10}$$

Then on inserting (5.10) and (5.9) into (5.3), yields

$$\Sigma_{vis} = \frac{d^2 N^{2d+4}}{36\zeta(d)} \left(1 + O\left(\frac{1}{d} + \frac{d}{N} \right) \right). \tag{5.11}$$

On combining (5.11), (5.2), Lemma 3.1 and Lemma 4.2, we obtain the following result.

Lemma 5.1. *There exists an absolute constant $C_2 > 0$, such that for all $d \geq 2$ and all $N \geq 3d$, we have*

$$\mathfrak{M}_{2,vis}(d, N) \leq C_2 (dN^4 + d^3 N^3). \tag{5.12}$$

6. Effective results and the proofs of Theorems 1, 3 and Corollary 1

We scale the bound for $\mathfrak{M}_{2,vis}(d,N)$ from Lemma 5.1 by d^2N^4, in order to have all spacings between points measured by the normalized distance situated in the interval $[0,1]$. Note first that for any $d \geq 2$ and any $N \geq 3d$, we have

$$\frac{\mathfrak{M}_{2,vis}(d,N)}{d^2N^4} \leq C_2\left(\frac{1}{d}+\frac{d}{N}\right).$$

Then, on combining the above inequalities with Lemma 4.2, there is an absolute constant $C_3>0$ such that

$$\frac{1}{\#\Omega}\sum_{(\boldsymbol{v},\boldsymbol{w})\in\Omega}\left(\mathfrak{d}_d^2(\boldsymbol{v},\boldsymbol{w})-\frac{1}{6}\right)^2 \leq C_3\left(\frac{1}{d}+\frac{d}{N}\right). \tag{6.1}$$

Now, for any parameters $a,T>0$, imposing supplementary conditions on the summation, we find the following lower bounds of the left-side term of the inequality (6.1):

$$\begin{aligned}\frac{1}{\#\Omega}\sum_{(\boldsymbol{v},\boldsymbol{w})\in\Omega}\left(\mathfrak{d}_d^2(\boldsymbol{v},\boldsymbol{w})-\frac{1}{6}\right)^2 &\geq \frac{1}{\#\Omega}\sum_{\substack{(\boldsymbol{v},\boldsymbol{w})\in\Omega\\ \left|\mathfrak{d}_d^2(\boldsymbol{v},\boldsymbol{w})-\frac{1}{6}\right|\geq\frac{1}{aT}}}\left(\mathfrak{d}_d^2(\boldsymbol{v},\boldsymbol{w})-\frac{1}{6}\right)^2\\ &\geq \frac{1}{\#\Omega}\sum_{\substack{(\boldsymbol{v},\boldsymbol{w})\in\Omega\\ \left|\mathfrak{d}_d^2(\boldsymbol{v},\boldsymbol{w})-\frac{1}{6}\right|\geq\frac{1}{aT}}}\frac{1}{a^2T^2}.\end{aligned} \tag{6.2}$$

Then, on combining (6.1) and (6.2), we find that

$$\frac{1}{\#\Omega}\#\left\{(\boldsymbol{v},\boldsymbol{w})\in\Omega\ :\ \left|\mathfrak{d}_d^2(\boldsymbol{v},\boldsymbol{w})-\frac{1}{6}\right|\geq\frac{1}{aT}\right\} \leq C_3a^2T^2\left(\frac{1}{d}+\frac{d}{N}\right). \tag{6.3}$$

Now, since

$$\left|\mathfrak{d}_d^2(\boldsymbol{v},\boldsymbol{w})-\frac{1}{6}\right| = \left|\mathfrak{d}_d(\boldsymbol{v},\boldsymbol{w})-\frac{1}{\sqrt{6}}\right|\left(\mathfrak{d}_d(\boldsymbol{v},\boldsymbol{w})+\frac{1}{\sqrt{6}}\right) \geq \frac{1}{\sqrt{6}}\left|\mathfrak{d}_d(\boldsymbol{v},\boldsymbol{w})-\frac{1}{\sqrt{6}}\right|,$$

by sharpening the restriction in the definition of the set on the left side of (6.3), the set remains with fewer elements, so that with $a=\sqrt{6}$, we derive that

$$\frac{1}{\#\Omega}\#\left\{(\boldsymbol{v},\boldsymbol{w})\in\Omega\ :\ \left|\mathfrak{d}_d(\boldsymbol{v},\boldsymbol{w})-\frac{1}{\sqrt{6}}\right|\geq\frac{1}{T}\right\} \leq 6C_3T^2\left(\frac{1}{d}+\frac{d}{N}\right). \tag{6.4}$$

In particular, this proves Theorem 1.

More generally, we consider the set Ω_K of K-polytopes $P=\{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K\}\subset\mathcal{W}$ the property that any of its two vertices are visible from each other. Then Corollary 1 follows from the following more general statements.

Theorem 5. *There exists an effectively computable absolute constant $C_4>0$ such that for any integers $d\geq 2$, $N\geq 3d$, $K\geq 2$ and any real $T>0$, we have*

$$\frac{1}{\#\Omega_K}\cdot\#\left\{\{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K\}\subset\Omega_K\ :\ \max_{1\leq m\neq n\leq K}\left|\mathfrak{d}_d(\boldsymbol{w}_m,\boldsymbol{w}_n)-\frac{1}{\sqrt{6}}\right|\geq\frac{1}{T}\right\} \leq C_4T^2K^2\left(\frac{1}{d}+\frac{d}{N}\right).$$

Corollary 2. *Let $\eta \in (0, 1/4)$ be fixed. Then, there exists an effectively computable absolute constant $C_5 > 0$ such that for any integers $d \geq 2$, $N \geq d^2$, $2 \leq K \leq d^{1/2-2\eta}$, we have*

$$\frac{1}{\#\Omega_K} \cdot \# \left\{ \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_K\} \subset \Omega_K \ : \ \begin{array}{c} \vartheta_d(\boldsymbol{w}_m, \boldsymbol{w}_n) \in \left[\frac{1}{\sqrt{6}} - \frac{1}{d^\eta}, \frac{1}{\sqrt{6}} + \frac{1}{d^\eta} \right] \\ \text{for all } 1 \leq m \neq n \leq K \end{array} \right\} \geq 1 - \frac{C_5}{d^{2\eta}}.$$

For the proof of Theorem 3 one can follow the path from Sections 4 and 5 with one component $\boldsymbol{v} = \boldsymbol{0}$ fixed in the involved summations. One finds that almost all normalized distances between the origin and the components of points in Ω are close to $1/\sqrt{3}$.

On the other hand, we know that, according to Theorem 1, almost all normalized distances between $\boldsymbol{v}$ and $\boldsymbol{w}$ with $(\boldsymbol{v}, \boldsymbol{w}) \in \Omega$ are almost always almost equal to $1/\sqrt{6}$. Therefore, almost all triangles with vertices $\boldsymbol{0}, \boldsymbol{v}, \boldsymbol{w}$ with $(\boldsymbol{v}, \boldsymbol{w}) \in \Omega$ are almost isosceles having the normalized edges almost equal to $1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{6}$ and Theorem 3 follows immediately.

7. The probability that a K-polytope is Self-Visible

Let $K \geq 2$ be a fixed integer. The set of self-visible K-polytopes with vertices in the lattice $\mathcal{W}$ is

$$\Omega_K = \left\{ P \in \mathcal{W}^K \ : \ \boldsymbol{v}', \boldsymbol{v}'' \text{ visible from each other, for all } \boldsymbol{v}', \boldsymbol{v}'' \in P \right\}. \tag{7.1}$$

Our object here is to see if there is a tendency of the probabilities that a K-polytope is self-visible as N gets large. We show that if d and K are kept fixed, the limit of the ratios

$$Prob(d, N, K) = \lim_{N \to \infty} \frac{\#\Omega_K}{\#\mathcal{W}^K}$$

does exist.

If $K = 2$, then Ω_K coincides with Ω, but the Möbius summation method used in the proof of Lemma 3.1 to estimate $\#\Omega_2$ is not suitable for larger K, because of the size of the multitude of new terms introduced. We need to have a better control on the large divisors, so we will proceed accordingly.

Denote a generic polytope by $P = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_K\}$ and the coordinates of its vertices by $\boldsymbol{v}_j = (v_{j,1}, \dots, v_{j,d})$ for $1 \leq j \leq K$. Note that, for each positive integer m, we have the following inequality

$$\# \left\{ (\boldsymbol{v}_j, \boldsymbol{v}_k) \in \mathcal{W}^2 \ : \ \boldsymbol{v}_j \neq \boldsymbol{v}_k, \ m \mid \gcd(v_{j,1} - v_{k,1}, \dots, v_{j,d} - v_{k,d}) \right\} \leq \frac{\#\mathcal{W}^2}{m^d}, \tag{7.2}$$

because, say, $v_{j,1}, \dots, v_{j,d}$ are free and then each of $v_{k,1}, \dots, v_{k,d}$ belongs to the corresponding shifted arithmetic progression of ratio m. Also, if $m \geq N$, the left side of (7.2) equals zero, since there are no pairs to count.

Fix $M > 0$, a parameter to be chosen later, and sum the inequalities (7.2) for all $m > M$. Then the size of the resulted sum is

$$\leq \sum_{m > M} \frac{\#\mathcal{W}^2}{m^d} = O\left(\frac{\#\mathcal{W}^2}{M^{d-1}} \right).$$

As a consequence, any such subsum is also $\ll \#\mathcal{W}^2 / M^{d-1}$. In particular, the sum over all positive integers m that have at least one prime factor larger than M. This holds for each pair $(\boldsymbol{v}_j, \boldsymbol{v}_k)$, and there are $K(K-1)/2$ such pairs with $1 \leq j < k \leq K$. As a consequence,

it follows that

$$\#\left\{(\boldsymbol{v}_1,\dots,\boldsymbol{v}_K)\in\mathcal{W}^K \;:\; \begin{array}{c} q\mid \gcd(v_{j,1}-v_{k,1},\dots,v_{j,d}-v_{k,d}) \\ \text{for some prime } q>M, \\ \text{for some } 1\le j<k\le K \end{array}\right\} \ll \frac{K^2\#\mathcal{W}^K}{M^{d-1}}. \tag{7.3}$$

In other words, with the exception of at most $O\left(\frac{K^2\#\mathcal{W}^K}{M^{d-1}}\right)$ K-tuples $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_K)$, for all the other polytopes $P=(\boldsymbol{v}_1,\dots,\boldsymbol{v}_K)\in\mathcal{W}^K$, the condition $P\in\Omega_K$ is equivalent to the condition that $P\in\Omega_K(M)$, where

$$\Omega_K(M):=\left\{(\boldsymbol{v}_1,\dots,\boldsymbol{v}_K)\in\mathcal{W}^K \;:\; \begin{array}{c} \gcd(B,v_{j,1}-v_{k,1},\dots,v_{j,d}-v_{k,d})=1 \\ \text{for all } 1\le j<k\le K \end{array}\right\}, \tag{7.4}$$

where B is the primorial number

$$B:=\prod_{\substack{p \text{ prime}\\ p\le M}} p.$$

Therefore, the probability that a polytope $P\in\mathcal{W}^K$ has all vertices visible from each other is

$$\frac{\#\Omega_K}{\#\mathcal{W}^K}=\frac{\#\Omega_K(M)}{\#\mathcal{W}^K}+O\left(\frac{K^2\#\mathcal{W}^K}{M^{d-1}}\right). \tag{7.5}$$

By the Prime Number Theorem, we know that $B=e^{(1+o(1))M}$, so that we will eventually choose M of size $\log N$ to assure that $B<N$.

Next, we split the interval $[0,N]$ in subintervals of size B. Accordingly, the cube $[0,N]^d$ is split in boxes of side length B. The number of these boxes is

$$\text{The number of boxes}=\left(\frac{N}{B}+O(1)\right)^d=\frac{N^d}{B^d}+O\left(\frac{dN^{d-1}}{B^{d-1}}\right). \tag{7.6}$$

Observe, by the definition, that $\Omega_K(M)$ has the same number of elements in each such box. Denote this number by $H(B)$, that is,

$$H(B):=\#\left\{(\boldsymbol{v}_1,\dots,\boldsymbol{v}_K)\in\Omega_K(M)\;:\;0\le v_{j,l}<B \text{ for all } 1\le j\le K,\,1\le l\le d\right\}. \tag{7.7}$$

Then, by (7.6) and (7.7), as each $\boldsymbol{v}_1,\dots,\boldsymbol{v}_K$ runs over each box, it follows that

$$\#\Omega_K(M)=H(B)\left(\frac{N^d}{B^d}+O\left(\frac{dN^{d-1}}{B^{d-1}}\right)\right)^K=\frac{N^{dK}H(B)}{B^{dK}}\left(1+O\left(\frac{dKB}{N}\right)\right). \tag{7.8}$$

Since $H(B)\le B^{dK}$ and since $\#\mathcal{W}^K=(N+1)^{dK}=N^{dK}(1+O(dK/N))$, it follows that

$$\frac{\#\Omega_K(M)}{\#\mathcal{W}^K}=\frac{H(B)}{B^{dK}}\left(1+O\left(\frac{dKB}{N}\right)\right)=\frac{H(B)}{B^{dK}}+O\left(\frac{dKB}{N}\right). \tag{7.9}$$

On combining (7.8) and (7.5), it yields

$$\frac{\#\Omega_K}{\#\mathcal{W}^K}=\frac{H(B)}{B^{dK}}+O\left(\frac{dKB}{N}\right)+O\left(\frac{K^2}{M^{d-1}}\right). \tag{7.10}$$

Now, for each prime $p\mid B$, consider the analogue of the set $\Omega_K(M)$ defined by (7.4). Its cardinality is analogous to $H(B)$ and is given by

$$H(p):=\#\left\{(\boldsymbol{v}_1,\dots,\boldsymbol{v}_K)\in\mathcal{W}^K \;:\; \begin{array}{c} \gcd(p,v_{j,1}-v_{k,1},\dots,v_{j,d}-v_{k,d})=1 \\ \text{for all } 1\le j<k\le K, \\ 0\le v_{j,l}\le p-1 \text{ for all } 1\le j\le K,\,1\le l\le d \end{array}\right\}. \tag{7.11}$$

Note that each $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_K)\in\mathcal{W}^K$ that contributes to $H(B)$ produces, via reduction modulo p, a K-tuple that contributes to $H(p)$, and this holds for each prime divisor p of B. Conversely, by the Chinese Remainder Theorem, each collection of K-tuples, with one K-tuple for each prime divisor of B, produces a unique K-tuple that is counted in $H(B)$. In conclusion,

$$H(B)=\prod_{\substack{p \text{ prime}\\ p|B}} H(p)=\prod_{\substack{p \text{ prime}\\ p\le M}} H(p),$$

which combined with (7.10) implies

$$\frac{\#\Omega_K}{\#\mathcal{W}^K}=\prod_{\substack{p \text{ prime}\\ p\le M}}\frac{H(p)}{p^{dK}}+O\left(\frac{dKB}{N}\right)+O\left(\frac{K^2}{M^{d-1}}\right). \tag{7.12}$$

Next, let us observe that since each of the coordinates $v_{j,1},\dots,v_{j,d}$ and $v_{k,1},\dots,v_{k,d}$ belongs to $\{0,1,\dots,p-1\}$, the difference $v_{j,1}-v_{k,1}$ cannot be divisible by p unless $v_{j,1}=v_{k,1}$, and similarly for all differences $v_{j,2}-v_{k,2},\dots,v_{j,d}-v_{k,d}$. As a consequence, the condition $\gcd(p,v_{j,1}-v_{k,1},\dots,v_{j,d}-v_{k,d})=1$ from the definition of $H(p)$ given by (7.11) is equivalent to the condition that the d-tuples $(v_{j,1},\dots,v_{j,d})$ and $(v_{k,1},\dots,v_{k,d})$ are distinct. In other words

$$H(p)=\#\left\{(\boldsymbol{v}_1,\dots,\boldsymbol{v}_K)\in\mathcal{W}^K\ :\ \begin{array}{l}\boldsymbol{v}_j\neq\boldsymbol{v}_k \text{ for } 1\le j\neq k\le K,\\ 0\le v_{j,l}\le p-1 \text{ for all } 1\le j\le K \text{ and } 1\le l\le d\end{array}\right\}.$$

Here, there are exactly p^d choices for $\boldsymbol{v}_1$. Then, for each fixed $\boldsymbol{v}_1$, the only restriction on $\boldsymbol{v}_2$ is to not coincide with $\boldsymbol{v}_1$, so that there are p^d-1 choices for $\boldsymbol{v}_2$. With $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ fixed, the only restrictions on $\boldsymbol{v}_3$ are $\boldsymbol{v}_3\neq\boldsymbol{v}_1$ and $\boldsymbol{v}_3\neq\boldsymbol{v}_2$, so that there are p^d-2 choices for $\boldsymbol{v}_3$. And so on, up to $\boldsymbol{v}_K$, for which there are $p^d-(K-1)$ choices. In conclusion

$$H(p)=p^d\left(p^d-1\right)\cdots\left(p^d-(K-1)\right).$$

On combining this with (7.12) we see that

$$\frac{\#\Omega_K}{\#\mathcal{W}^K}=\prod_{\substack{p \text{ prime}\\ p\le M}}\left(1-\frac{1}{p^d}\right)\cdots\left(1-\frac{K-1}{p^d}\right)+O\left(\frac{dKB}{N}\right)+O\left(\frac{K^2}{M^{d-1}}\right). \tag{7.13}$$

The finite product over primes in (7.13) can be replaced with the completed product over all primes, with a change in the error term that is swallowed inside the last error term. Indeed, if we denote

$$\Lambda_{d,K}(M):=\prod_{\substack{p \text{ prime}\\ p>M}}\left(1-\frac{1}{p^d}\right)\cdots\left(1-\frac{K-1}{p^d}\right),$$

an infinite product that converges if $d\ge 2$, then

$$\log\Lambda_{d,K}(M)=\sum_{\substack{p \text{ prime}\\ p>M}}\sum_{1\le k\le K-1}\log\left(1-\frac{k}{p^d}\right)=\sum_{\substack{p \text{ prime}\\ p>M}}\sum_{1\le k\le K-1}O\left(\frac{k}{p^d}\right).$$

This implies

$$|\log\Lambda_{d,K}(M)|=O\Big(\sum_{\substack{p \text{ prime}\\ p>M}}\sum_{1\le k\le K-1}\frac{k}{p^d}\Big)=O\Big(\sum_{\substack{p \text{ prime}\\ p>M}}\frac{K^2}{p^d}\Big)=O\Big(\sum_{m>M}\frac{K^2}{m^d}\Big)=O\left(\frac{K^2}{M^{d-1}}\right).$$

It follows that

$$\Lambda_{d,K}(M) = \exp\left(O\left(\frac{K^2}{M^{d-1}}\right)\right) = 1 + O\left(\frac{K^2}{M^{d-1}}\right).$$

Therefore, if we denote by $\Lambda_{d,K}$ the complete infinite product,

$$\Lambda_{d,K} := \prod_{p \text{ prime}} \left(1 - \frac{1}{p^d}\right) \cdots \left(1 - \frac{K-1}{p^d}\right), \tag{7.14}$$

which is constant for any fixed d and K, we have

$$\prod_{\substack{p \text{ prime} \\ p \le M}} \left(1 - \frac{1}{p^d}\right) \cdots \left(1 - \frac{K-1}{p^d}\right) = \frac{\Lambda_{d,K}}{\Lambda_{d,K}(M)} = \Lambda_{d,K} + O\left(\frac{K^2}{M^{d-1}}\right), \tag{7.15}$$

where the implied constant in the big O estimate is absolute, because $\Lambda_{d,2} = \zeta(d)^{-1}$ for $d \ge 2$ and, for any fixed d, the sequence $\{\Lambda_{d,K}\}_{K \ge 2}$ is decreasing.

Then, inserting (7.15) and (7.14) in (7.13), we arrive at the following result

$$\frac{\#\Omega_K}{\#\mathcal{W}^K} = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^d}\right) \cdots \left(1 - \frac{K-1}{p^d}\right) + O\left(\frac{dKB}{N}\right) + O\left(\frac{K^2}{M^{d-1}}\right).$$

We now take B to be the largest primorial that is $\le \sqrt{N}$, which means that $M \sim (\log N)/2$. Then,

$$\frac{\#\Omega_K}{\#\mathcal{W}^K} = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^d}\right) \cdots \left(1 - \frac{K-1}{p^d}\right) + O\left(\frac{dK}{\sqrt{N}}\right) + O\left(\frac{2^d K^2}{\log^{d-1} N}\right), \tag{7.16}$$

and the constants implied in the big O terms are absolute. This concludes the proof of Theorem 2.

8. Probabilistic Intuition

In this section, we show how to interpret the constant $1/\sqrt{6}$ in Theorem 1 via probabilistic intuition. Similar arguments can give intuition for some of the other particular constants we obtain in this paper. We recall

$$\mathfrak{d}_d(\boldsymbol{v}, \boldsymbol{w}) = \frac{1}{Nd^{1/2}} \left(\sum_{n=1}^{d} (w_n - v_n)^2\right)^{1/2} = \left(\frac{1}{d} \sum_{n=1}^{d} \left(\frac{w_n}{N} - \frac{v_n}{N}\right)^2\right)^{1/2},$$

As $N \to \infty$, if we select $\boldsymbol{v}$ uniformly at random from $\mathcal{W}$ (or, indeed, the subset of integer vectors visible from the origin in $\mathcal{W}$), the normalized vector $\boldsymbol{v}/N$ becomes equidistributed in the hypercube $[0,1]^d$. That is,

$$\sum_{\boldsymbol{v} \in \mathcal{W}} \delta_{\boldsymbol{v}/N} \to m,$$

where m is the standard Lebesgue measure on $[0,1]^d$, and the convergence is in the weak*-sense as $N \to \infty$. The same result is true with $\mathcal{W}$ replaced by the subset of primitive vectors in $\mathcal{W}$. Thus, to try and get intuition about the $d \to \infty$ behavior of our normalized distance $\mathfrak{d}_d(\boldsymbol{v}, \boldsymbol{w})$, we can consider the following probabilistic analogue. Let $\mathbf{X} = (X_1, \dots X_d), \mathbf{Y} = (Y_1, \dots, Y_d)$ be independent random vectors chosen according to

Lebesgue measure on the hypercube $[0,1]^d$. Thus $X_1, \ldots, X_d$ and $Y_1, \ldots Y_d$ are independent, identically distributed (i.i.d.) uniform $[0,1]$ random variables, and also indpendent from each other. We define

$$\mathfrak{d}_d(\mathbf{X}, \mathbf{Y}) = \left(\frac{1}{d} \sum_{n=1}^{d} (X_n - Y_n)^2 \right)^{1/2}.$$

Thus

$$\mathfrak{d}_d^2(\mathbf{X}, \mathbf{Y}) = \frac{1}{d} \sum_{n=1}^{d} (X_n - Y_n)^2$$

is the sample mean of d independent random variables of the form $(U-V)^2$, where U and V are independent uniform $[0,1]$ random variables. By the strong law of large numbers, as $d \to \infty$, this converges with probability 1 to the mean

$$E((U-V)^2) = \int_0^1 \int_0^1 (u-v)^2 du dv = 1/6.$$

That is, as $d \to \infty$, with probability 1,

$$\mathfrak{d}_d^2(\mathbf{X}, \mathbf{Y}) \to 1/6,$$

so

$$\mathfrak{d}_d(\mathbf{X}, \mathbf{Y}) \to 1/\sqrt{6}.$$

To be clear, this does not give a direct proof of Theorem 1, since there are tricky issues with the interchange of limits. Similar arguments can yield intuition for the other constants in our results.

Acknowledgement. We thank Sara Billey, Sam Fairchild, Alex Kontorovich, and Doug West for valuable discussions at a variety of times about the problem of the limiting density of Ω_K. J.S.A. was partially supported by NSF grant DMS 2003528, 'Curves, Counting, and Correlations'. J.S.A. also acknowledges the hospitality of the Mathematical Sciences Research Institute (MSRI) during the Spring 2022 program on Analysis and Geometry of Random Spaces.

References

[1] Charu C. Aggarwal, Alexander Hinneburg, Daniel A. Keim, *On the Surprising Behavior of Distance Metric in High-Dimensional Space*, Van den Bussche, Jan (ed.) et al., Database theory - ICDT 2001. 8th international conference, London, GB, January 4–6, 2001. Proceedings. Berlin: Springer (ISBN 3-540-41456-8). Lect. Notes Comput. Sci. 1973, 420–434 (2001). 1

[2] D. H. Baileya, J. M. Borwein, R. E. Crandall, *Box integrals*, J. Comput. Appl. Math. **206** (2007), no. 1, 196–208. 1

[3] Uwe Bäsel, *The moments of the distance between two random points in a regular polygon*, preprint January 11, 2021. https://arxiv.org/pdf/2101.03815.pdf 1

[4] Sebastien Bubeck, Mark Sellke, *A universal law of robustness via isoperimetry*, Advances in Neural Information Processing Systems 34 pre-proceedings (NeurIPS 2021), https://arxiv.org/pdf/2105.12806.pdf. 1, 4

[5] B. Burgstaller, F. Pillichshammer, *The average distance between two points*, Bull. Aust. Math. Soc., **80** (2009), no. 3 , 353–359. 1

[6] C. I. Cobeli, S. M. Gonek, A. Zaharescu, *The distribution of patterns of inverses modulo a prime* J. Number Theory **101** (2003), no. 2, 209–222 . 7

[7] Cristian Cobeli, Alexandru Zaharescu, *On the distribution of the $\mathbb{F}_p$-points on an affine curve in r dimensions*, Acta Arith. **99** (2001), no. 4, 321–329. 7

[8] Cristian Cobeli, Alexandru Zaharescu, *Generalization of a problem of Lehmer*, Manuscripta Math. **104** (2001), no. 3, 301–307. 6, 7

[9] Steven R. Dunbar, *The average distance between points in geometric figures*, Coll. Math. J., The College Mathematics Journal, **28** (1997), no. 3, 187–197. 1
[10] Willy Feller, Erhard Tornier, *Mengentheoretische Untersuchung von Eigenschaften der Zahlenreihe, Mathematische Annalen*, **107** (1932), 188–232. 3
[11] A. Gafni, A. Iosevich, E. Wyman, *Uniform distribution and geometric incidence theory*, preprint February 10, 2022. https://arxiv.org/pdf/2202.05359.pdf 1
[12] Richard K. Guy, *Unsolved problems in number theory*. Third edition. Problem Books in Mathematics. Springer-Verlag, New York, 2004. xviii+437 pp. 6, 7
[13] A. Iosevich, *On the approximate unit distance problem*, Aldroubi, Akram (ed.) et al., New trends in applied harmonic analysis. Volume 2. Harmonic analysis, geometric measure theory, and applications. Collected papers based on courses given at the 2017 CIMPA school, Buenos Aires, Argentina, July 31 - August 11, 2017. Cham: Birkhäuser. Appl. Numer. Harmon. Anal., pages 121–128 (2019). 1
[14] A. Iosevich, M. Mourgoglou, K. Taylor, *On the Mattila-Sjölin theorem for distance sets*, Ann. Acad. Sci. Fenn., Math. **37** (2012), no. 2, 557–562. 1
[15] A. Iosevich, M. Rudnev, I. Uriarte-Tuero, *Theory of dimension for large discrete sets and applications*, Math. Model. Nat. Phenom. **9** (2014), no. 5, 148–169. 1
[16] A. Iosevich and S. Senger, *Sharpness of Falconer's $\frac{d+1}{2}$ estimate*, Ann. Acad. Sci. Fenn. Math. **41** (2016), no. 2, 713–720. 1
[17] Hongjun Li, Xing Qiu, *Moments of distance from a vertex to a uniformly distributed random point within arbitrary triangles*, Math. Probl. Eng., **2016** (2016), Article ID 8371750, 10 p. 1
[18] A. M. Mathai, P. Moschopoulos, G. Pederzoli, *Random points associated with rectangles*, Rend. Circ. Mat. Palermo II. Ser. **48** (1999), no. 1, 163–190. 1
[19] D. Oberlin, R. Oberlin, *Unit distance problems,* Am. J. Math. **137** (2015), no. 1, 251–270. 1
[20] The On-line Encyclopedia of Integer Sequences, *Sequence A065493*, https://oeis.org/A065493. 3
[21] Sunil Srinivasa, Martin Haenggi, *Distance distributions in finite uniformly random networks: theory and application*, IEEE Transactions on vehicular technology, **59** (2010), no. 2, 940–949. 1
[22] André Weil, *On some exponential sums*, Proc. Nat. Acad. Sci. U.S.A. **34** (1948), 204–207. 8
[23] Alexandru Zaharescu, *The distribution of the values of a rational function modulo a big prime*, J. Théor. Nombres Bordx., **15** (2003), no. 3, 863–872. 7, 8

JA: Department of Mathematics, University of Washington, Padelford Hall, Seattle, WA 98195, USA
Email address: jathreya@uw.edu

CC: Simion Stoilow Institute of Mathematics of the Romanian Academy, P. O. Box 1-764, RO-014700 Bucharest, Romania
Email address: cristian.cobeli@gmail.com

AZ: Department of Mathematics, University of Illinois at Urbana-Champaign, Altgeld Hall, 1409 W. Green Street, Urbana, IL, 61801, USA and Simion Stoilow Institute of Mathematics of the Romanian Academy, P. O. Box 1-764, RO-014700 Bucharest, Romania
Email address: zaharesc@illinois.edu