An Improved Construction of Progression-Free Sets

Michael Elkin *

Abstract

The problem of constructing dense subsets S of $\{1, 2, \ldots, n\}$ that contain no three-term arithmetic progression was introduced by Erdős and Turán in 1936. They have presented a construction with $|S| = \Omega(n^{\log_3 2})$ elements. Their construction was improved by Salem and Spencer, and further improved by Behrend in 1946. The lower bound of Behrend is

$$|S| \;=\; \Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log_2 n}} \cdot \log^{1/4} n}\right) \;.$$

Since then the problem became one of the most central, most fundamental, and most intensively studied problems in additive number theory. Nevertheless, no improvement of the lower bound of Behrend has been reported since 1946.

In this paper we present a construction that improves the result of Behrend by a factor of $\Theta(\sqrt{\log n})$, and shows that

$$|S| \;=\; \Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log_2 n}}} \cdot \log^{1/4} n\right) \;.$$

In particular, our result implies that the construction of Behrend is not optimal.

Our construction and proof are elementary and self-contained. Also, the construction can be implemented by an efficient algorithm.

Behrend's construction has numerous applications in Theoretical Computer Science. In particular, it is used for fast matrix multiplication, for property testing, and in the area of communication complexity. Plugging in our construction instead of Behrend's construction in the matrix multiplication algorithm of Coppersmith and Winograd improves the state-of-the-art upper bound on the complexity of the matrix multiplication by a factor of $\log^{\nu} n$, for some fixed constant $\nu > 0$. We also present an application of our technique in Computational Geometry.

*Department of Computer Science, Ben-Gurion University of the Negev, Beer-Sheva, Israel, elkinm@cs.bgu.ac.il

This research has been supported by the Israeli Academy of Science, grant 483/06, by the BSF grant No. 2008390, and by the Lynn and William Center for Computer Science.

1 Introduction

A subset $S \subseteq \{1, 2, \ldots, n\}$ is called *progression-free* if it contains no three distinct elements $i, j, \ell \in S$ such that i is the arithmetic average of j and ℓ, i.e., $i = \frac{j+\ell}{2}$. For a positive integer n, let $\nu(n)$ denote the largest size of a progression-free subset S of $\{1, 2, \ldots, n\}$.

Providing asymptotic estimates on $\nu(n)$ is a central, fundamental, and very well-studied inverse problem in additive number theory. This problem was introduced[1] by Erdős and Turan [20], and they showed that $\nu(n) = \Omega(n^{\log_3 2})$. This estimate was improved by Salem and Spencer [36], and further improved by Behrend [10] in 1946. Behrend has shown that

$$\nu(n) \;=\; \Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log_2 n}} \cdot \log^{1/4} n}\right) \;, \tag{1.1}$$

and this bound has remained state-of-the-art for more than sixty years. A slightly weaker lower bound that does not rely on the Pigeonhole Principle was shown by Moser [30]. (Moser [30] cites [10] for the lower bound $\nu(n) \geq \frac{n}{2^{(2\sqrt{2}+\epsilon)\sqrt{\log_2 n}}}$, for every $\epsilon > 0$. This lower bound is slightly weaker than (1.1). The lower bound (1.1) can, however, also be derived by the argument of Behrend.) We refer the reader to [21] for a thorough general discussion of inverse problems in Additive Number Theory.

The first non-trivial upper bound $\nu(n) = O(\frac{n}{\log \log n})$ was proved in a seminal paper by Roth [34]. This bound was improved by Bourgain [12, 13], and the current state-of-the-art upper bound is $\nu(n) = O(n \cdot \frac{(\log \log n)^2}{\log^{2/3} n})$ [13]. The problem is also closely related to Szemerédi's theorem [39], which in particular, implies that $\nu(n) = o(n)$. It is also related to the problem of finding arbitrarily long arithmetic progressions of prime numbers (see, e.g., Green and Tao [24]), and to other central problems in the additive number theory.

In this paper we improve the lower bound of Behrend by a factor of $\Theta(\sqrt{\log n})$, and show that

$$\nu(n) \;=\; \Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log_2 n}}} \cdot \log^{1/4} n\right) \;.$$

[1] A closely related problem was studied by van der Waerden [40].

Though the improvement is not large, our result demonstrates that the bound of Behrend is not optimal. Also, it is the first lower bound that shows that $\nu(n)$ is asymptotically greater than $\frac{n}{2^{2\sqrt{2}\sqrt{\log_2 n}}}$.

Like the proof of Behrend, our proof relies on the Pigeonhole Principle. Consequently, the result of Moser [30] remains the best known lower bound achieved without relying on the Pigeonhole Principle. However, we hope that our argument can be made independent of the Pigeonhole Principle. (See also Section 8.)

Related work: The proof of Behrend was generalized by Rankin [33] to provide large subsets of $\{1, 2, \ldots, n\}$ that contain no arithmetic progression of length k, for any fixed k. (See also [29] the more recent variant of the proof of [33].) Ruzsa [35] and Shapira [37] extended the proofs of Behrend [10] and Rankin [33] further, and constructed large subsets of $\{1, 2, \ldots, n\}$ that exclude solutions of certain linear Diophantine equations. Abbott [1, 2, 3] and Bosznay [11] generalized the proof of Behrend [10] in another direction, and devised constructions of large non-averaging subsets of S of $\{1, 2, \ldots, n\}$. (A subset S is said to be *non-averaging* if no element $x \in S$ is an average of two or more other elements of S.) Gasarch et al. [23] studied the problem empirically, and constructed large progression-free subsets of $\{1, 2, \ldots, n\}$ for $n \leq 250$.

Consequent work: The preliminary version of this paper was published in the electronic archive [19] in January 2008. Since then several authors continued our line of research. Specifically, Green and Wolf [25] found a simpler proof of our result. They point out, however, that "the only advantage of our approach is brevity: it is based on ideas morally close to those of Elkin, and moreover, his argument is more constructive than ours." In an even more recent development O'Bryant [31] has combined our techniques with those of Rankin [33] and Green and Wolf [25], and improved Rankin's lower bound by a factor $\log^{\epsilon} n$, for some small positive $\epsilon = \epsilon(k)$. (In the preliminary version of our paper [19] we anticipated that our techniques could be useful to improve Rankin's bound.) We believe that similar ideas may help improving some of the results of [35, 37, 1, 2, 3, 11].

Overview of the proof: Our proof is elementary, and self-contained. The proof of Behrend is based on the observation that a sphere in any dimension is convexly independent, and thus cannot contain three vectors such that one of them is the arithmetic average of the two other. We replace the sphere by a thin annulus. Intuitively, we are able to produce larger progression-free sets because an annulus of non-zero width contains more integer lattice points than a sphere of the same radius does. However, unlike in a sphere, the set of integer lattice points in an annulus is not necessarily convexly independent. To counter this difficulty we show that as long as the annulus $\mathcal{A}$ is sufficiently thin, the set U of its integer lattice points contains a convexly independent subset $W \subseteq U$ whose size is at least a constant fraction of the size of U, i.e., $|W| = \Omega(|U|)$. The subset W is, in fact, the exterior set $Ext(U)$ of the set U. (By "exterior set of U" we mean the boundary of the convex hull of U.)

In our analysis we actually have to consider the intersection $\mathcal{S}$ of the annulus $\mathcal{A}$ with a certain hypercube C, and to show that $|Ext(U)| = \Omega(|U|)$ for the corresponding set U of integer lattice points of $\mathcal{S}$. To prove a lower bound on the cardinality of $Ext(U)$, we consider a set $\mathcal{F}$ of hyperplanes, and demonstrate that each point $x \in U \setminus Ext(U)$ belongs to one of the hyperplanes $\mathcal{H}_x$ from $\mathcal{F}$. We then argue that $\mathcal{F}$ contains only a small number of hyperplanes, and that each of these hyperplanes $\mathcal{H}$ contains only a relatively small number of points of U.

Our analysis of the number of integer lattice points in $\mathcal{H} \cap U$ boils down to estimating the $(k-1)$-dimensional volume of the corresponding high-dimensional body $\tilde{U} = \mathcal{H} \cap \mathcal{S} = \mathcal{H} \cap \mathcal{A} \cap C$, and showing that the discrepancy between the volume of $\tilde{U}$ and the number of integer lattice points in it is quite small. A naive upper bound on the volume of $\tilde{U}$ is the volume of the $(k-1)$-dimensional annulus $\mathcal{H} \cap \mathcal{A}$, where k is the dimension of the annulus $\mathcal{A}$. The latter volume is much easier to compute, but unfortunately, this upper bound turns out to be far too crude. Instead we show that after an appropriate rotation of the space, the body $\tilde{U} = \mathcal{H} \cap \mathcal{A} \cap C$ becomes contained in the intersection of the annulus $\mathcal{H} \cap \mathcal{A}$ with a relatively small number of octants, and use the volume of this intersection as our upper bound for the volume of $\tilde{U}$.

In addition, estimating the discrepancy between the volume and the number of integer lattice points of $\tilde{U}$ is not easy either. One technical difficulty is that the dimension k of this body is not fixed, but rather tends to infinity logarithmically with the radius of the annulus $\mathcal{A}$. On the other hand, most estimates for the discrepancy between the volume and the number of integer lattice points of high-dimensional bodies assume that the dimension is fixed, and consequently, these estimates are inapplicable for our purposes. To overcome this technical difficulty we explicated the dependency on the dimension in the relevant estimates. Another technical difficulty is that the annulus $\mathcal{A}$ is very thin. Intuitively, thin bodies may have a large volume but contain no (or a very small number of)

integer lattice points. From the technical perspective, this makes the analysis more elaborate.

However, even though the part of our proof that shows that $Ext(U)$ has large cardinality is technically challenging, we believe that our main contribution is in devising a *new scheme* for producing large progression-free sets. This scheme builds upon Behrend's construction, but it employs a *different strategy for constructing a convexly independent set* of integer lattice points. While Behrend's construction uses a set of integer lattice points that lie on a sphere, our scheme constructs a large convexly independent subset of the set of integer lattice points of an annulus. As was mentioned above, this annulus has to be sufficiently thin so that $Ext(U)$ will be of size which is at least a constant fraction of $|U|$. In our proof we set the width of the annulus to be the maximum (up to a constant factor) value for which this condition holds. (Note that the size of the resulting progression-free set is proportional to the width of the annulus that we use.)

Applications of our technique: Given a large positive real R, and an integer $k \geq 2$, let $C_k(R)$ denote the maximum size of a convexly independent set (henceforth, CIS) of k-dimensional integer vectors with norm at most R. Equivalently, $C_k(R)$ is the number of extreme points of the convex hull of the set of extreme points of the k-dimensional ball $B(R)$ of radius R centered at the origin. In 1925 Jarnik [27] proved that $C_2(R) = \Theta(R^{2/3})$. His proof is constructive, and it gives rise to an algorithm with running time $O(R^{2/3} \log R)$ for computing a CIS with an optimal (up to constant factors) size. More precise estimates on $C_2(R)$ were derived by Arnold [7] and Balog and Barany [8]. In 1963 Andrews [6] published a simple argument that extends the upper bound of Jarnik to larger dimensions, and shows that $C_k(R) = O(R^{k-2+\frac{2}{k+1}})$ for all $k \geq 3$. Establishing the corresponding lower bound turned out to be more difficult, and only in the end of nineties Barany and Larman [9] proved that $C_k(R) = \Omega(R^{k-2+\frac{2}{k+1}})$ for all $k \geq 3$. The proof of [9] is quite elaborate, and in particular, it relies on the Flatness Theorem of Khintchine [28]. In addition, the result of [9] is not constructive in the sense that, to the best of our knowledge, it does not give rise to an efficient algorithm for constructing a CIS of size $C_k(R)$. The fastest currently known algorithm for computing such a CIS invokes one of the existing algorithms for computing a convex hull of an arbitrary set of vectors on the ball $B(R)$. However, this approach is doomed to have running time $\Omega(|B(R)|) = \Omega(R^k)$. On the other hand, since the input and the output sizes for this problem are $O(\log R)$ and $O(R^{k-2+\frac{2}{k+1}})$, respectively, it is reasonable to expect that by other means one can do significantly better.

The fastest currently known algorithm for computing a k-dimensional convex hull of an arbitrary set of n points for $k = 3$ is due to Preparata and Hong [32], and it requires $O(n \log n)$ time. For $k \geq 4$ the best known output-sensitive algorithm is due to Chan [14], and it requires $O(n \cdot \log q + (nq)^{1-\frac{1}{\lfloor k/2 \rfloor + 1}} \cdot \log^{O(1)} n)$ time, where q is the size of the output. Using these algorithms for our problem ($n = R^k$, $q = R^{k-2+\frac{2}{k+1}}$) one obtains running time of $O(n \log n) = O(R^3 \log R)$ for computing the CIS of optimal size in 3 dimensions. For even (respectively, odd) $k \geq 4$ the running time is $O(n^{2-\frac{6}{k+2}+\frac{2}{(k+1)(k+2)}} \cdot \log^{O(1)} n)$ (resp., $O(n^{2-2\cdot\frac{3k+1}{(k+1)^2}} \cdot \log^{O(1)} n)$).

By applying the technique that we developed for constructing large progression-free sets we provide an alternative (to the one of [9]), simple and *constructive* proof of the lower bound $C_k(R) = \Omega(R^{k-2+\frac{2}{k+1}})$ for $k \geq 5$. For $k = 4$ and $k = 3$ our proof provides slightly suboptimal estimates. Specifically, for $k = 4$ it yields $C_4(R) = \Omega\left(\frac{R^{12/5}}{(\log \log R)^{2/5}}\right)$, which is suboptimal by a factor of $O((\log \log R)^{2/5})$. For $k = 3$ we show $C_3(R) = \Omega(R^{3/2-\epsilon})$, for an arbitrarily small $\epsilon > 0$, which is suboptimal by a factor of R^{ϵ}. Our proof relies on standard estimates for the discrepancy between the volume and the number of integer points in large k-dimensional balls, and is otherwise self-contained.

More importantly, our proof gives rise to a very efficient algorithm for computing a CIS of nearly optimal size. Specifically, the running time of our algorithm is $O(R^{k-1+\frac{1}{k+1}}) = O(n^{1-\frac{1}{k+1}})$ for all $k \geq 3$. In other words, it is *sublinear* in n, while any algorithm that computes the convex hull of $B(R)$ requires time at least *linear* in n. (In fact, with existing algorithms the running time is *superlinear* in n.) Moreover, the improvement in running time becomes even more significant as k grows to infinity; indeed, the exponent of n in the running time of our algorithm tends to 1, while the exponent of the algorithm that is based on computing the convex hull of $B(R)$ tends to 2. In addition, already for small k the running time of our algorithm is significantly smaller than that of the previously best-known one. See Table 1 for a concise comparison between the exponents of n in the running times of our and previous algorithms for this problem. In the bottom line of the table we provide the exponent of the output size, which serves as a lower bound for the running time of any algorithm for this problem. Note, however, that this improved efficiency comes with a price. For $k \geq 5$ our algorithm may provide a CIS which is of size smaller at most by a constant factor than the optimal one. For $k = 4$ the approximation factor becomes

k	3	4	5	6	7	8	9	10	∞
Previous	1	31/30	10/9	9/7	21/16	64/45	36/25	41/27	$\rightarrow 2$
Our $(1-\frac{1}{k+1})$	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	$\rightarrow 1$
Lower Bound $(1-\frac{2}{k+1})$	1/2	3/5	2/3	5/7	3/4	7/9	4/5	9/11	$\rightarrow 1$

Table 1: The exponents of previous upper bounds on the running time, our upper bounds, and lower bounds, for k in the range $3 \leq k \leq 10$ are summarized. Polylogarithmic factors are suppressed.

$O((\log\log R)^{2/5}) = O((\log\log n)^{2/5})$, and for $k = 3$ it is $O(R^{\epsilon}) = O(n^{\epsilon/3})$, for an arbitrarily small $\epsilon > 0$. On the other hand, we believe that in many applications one would be willing to use our far more efficient algorithm while compromising slightly on the output size.

Applications of our result: Throughout the years the construction of Behrend has found numerous applications in Theoretical Computer Science. In particular, it is a central ingredient in the celebrated matrix multiplication algorithm of Coppersmith and Winograd [18]. It was also used for property testing [5], and for devising multi-party protocols for basic communication-theoretic problems [15]. Our construction can be implemented by a deterministic algorithm with running time $n/2^{\Omega(\sqrt{\log n})}$, and thus it may be used instead of Behrend construction in all applications. (Our running time is sublinear in n, but superlinear in the size of the output.)

However, naturally, the resulting improvements are very small. Specifically, the current estimate for the running time of the algorithm of [18] (which is the current state-of-the-art algorithm for multiplying two square $n \times n$ matrices) is $n^{\omega} \cdot \zeta(n)$, with $\omega < 2.376$ being a universal constant, and $\zeta(n)$ being a function such that $\zeta(n) = n^{o(1)}$. Plugging in our construction instead of that of Behrend into the algorithm of [18] improves this estimate to $O(n^{\omega} \cdot \frac{\zeta(n)}{\log^{\nu} n})$, for some universal small constant $\nu > 0$. Even though the improvement is only by a factor of $\log^{\nu} n$, it may serve as an indication of the relevance and potential applicability of our result to the area of Algorithmics.

To our knowledge, our result provides no direct bearings to property testing. To improve the results of Alon and Shapira [5] one needs large subsets that exclude arithmetic progressions of length larger than 3, and subsets that exclude some more general patterns. However, it is plausible that our construction will be later extended to provide larger subsets of these kinds, and this, in turn, may result in improved estimates for the problems considered in [5].

Finally, Chandra *et al.*[15] used large progression-free subsets to devise efficient multi-party protocols for some basic problems in Communication Complexity. However, plugging in our construction instead of Behrend's construction in the analysis of [15] improves only lower-order terms in their estimates.

The Structure of the Paper: In Section 2 we provide definitions and notation that are used throughout the paper. In Section 3 we overview Behrend's construction [10]. Section 4 contains our construction and its analysis. The analysis uses an estimate (inequality (4.20)) of the discrepancy between the number of integer lattice points and the volume of a certain high-dimensional body. This estimate is closely related to known ones. For the sake of completeness, we provide a self-contained proof of this estimate in Section 6. In Section 5 we provide an alternative proof for the result of Barany and Larman [9]. In Section 7 we discuss the application of our result to fast matrix multiplication. In Section 8 we provide a short summary and discuss some directions for future research.

2 Preliminaries

For a pair a, b of real numbers, $a \leq b$, we denote by $[a, b]$ (respectively, (a, b)) the closed (resp., open) segment containing all numbers x, $a \leq x \leq b$ (resp., $a < x < b$). We also use the notation $(a, b]$ (respectively, $[a, b)$) for denoting the segment containing all numbers x, $a < x \leq b$ (resp., $a \leq x < b$). For integer numbers n and m, $n \leq m$, we denote by $[\{n, m\}]$ the set of integer numbers $\{n, n+1, \ldots, m\}$. If $n = 1$ then we use the notation $[\{m\}]$ as a shortcut for $[\{1, m\}]$. For a real number x, we denote by $\lfloor x \rfloor$ (respectively, $\lceil x \rceil$) the largest (resp., smallest) integer that is no greater (resp., no smaller) than x. For a k-vector δ, we denote by $gcd(\delta)$ the greatest common divisor of non-zero coordinates of δ.

A triple i, j, ℓ of distinct integers is called an *arithmetic triple* if one of these numbers is the average of two other numbers, i.e., $i = \frac{j+\ell}{2}$. A set S of integer numbers is called *progression-free* if it contains no arithmetic triple. For a positive integer number n, let $\nu(n)$ denote the largest size of a progression-free subset S of $[\{n\}]$.

Unless specified explicitly, log (respectively, ln) stands for the logarithm on *base 2* (resp., e).

For a positive integer k and a vector $v =$

$(v_1, v_2, \ldots, v_k)$, let $||v|| = \sqrt{\sum_{i=1}^{k} v_i^2}$ denote the *norm* of the vector v. The expression $||v||^2 = \sum_{i=1}^{k} v_i^2$ will be referred to as the *squared norm* of the vector v.

For a set $v^{(1)}, v^{(2)}, \ldots, v^{(t)} \in \mathbb{R}^k$ of vectors, a vector v is a *convex combination* of $v^{(1)}, v^{(2)}, \ldots, v^{(t)}$ if there exist non-negative numbers $p_1, p_2, \ldots, p_t$ that sum up to 1 (i.e., $\sum_{i=1}^{t} p_i = 1$), and $v = \sum_{i=1}^{t} p_i v^{(i)}$. A convex combination is called *trivial* if for some $i \in [\{t\}]$, $p_i = 1$. Otherwise, it is called *non-trivial*. For a set $U \subseteq \mathbb{R}^k$ of vectors, we say that U is a *convexly independent set* if for every vector $u \in U$, there is no non-trivial convex combination of vectors from U that is equal to u. For a set $X \subseteq \mathbb{R}^k$ of vectors, the *exterior set* of X, denoted $Ext(X)$, is the subset of X that contains all vectors $v \in X$ such that v cannot be expressed as a non-trivial convex combination of vectors from X. (This is the set of the extreme points of the convex hull of X.)

For a positive integer ℓ, let β_ℓ denote the volume of an ℓ-dimensional ball of unit radius. It is well-known (see, e.g, [22], p.3) that

$$\beta_\ell \;=\; \frac{\pi^{\ell/2}}{\Gamma(\frac{\ell}{2}+1)} \;, \tag{2.2}$$

where $\Gamma(\cdot)$ is the (Euler) Gamma-function. We use the Gamma-function either with a positive integer parameter n or with a parameter $n + \frac{1}{2}$ for a positive integer n. In these cases the Gamma-function is given by $\Gamma(n+1) = n!$ and

$$\Gamma\left(n+\frac{1}{2}\right) \;=\; \frac{(2n)!\sqrt{\pi}}{2^{2n}n!} \;. \tag{2.3}$$

(See [22], p.178.) Observe also that

$$\begin{aligned}\Gamma\left(n+\frac{1}{2}\right) \;&=\; \left(n-\frac{1}{2}\right)\left(n-\frac{3}{2}\right)\cdot\ldots\cdot\frac{1}{2}\cdot\sqrt{\pi} \\ &\geq\; (n-1)!\frac{\sqrt{\pi}}{2} \;.\end{aligned} \tag{2.4}$$

By definition, it is easy to verify that for an integer ℓ, $\ell \geq 2$, $\beta_\ell = \Theta(\frac{\beta_{\ell-1}}{\sqrt{\ell}})$.

3 Behrend's Proof

The state-of-the-art lower bound for $\nu(n)$ due to Behrend [10] states that for every positive integer n,

$$\nu(n) \;=\; \Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log n}}\cdot\log^{1/4} n}\right) \;. \tag{3.5}$$

In this paper we improve this bound by a factor of $\Theta(\sqrt{\log n})$, and show that for every positive integer n,

$$\nu(n) \;=\; \Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log n}}}\cdot\log^{1/4} n\right) \;. \tag{3.6}$$

Note that it is sufficient to prove this bound only for all sufficiently large values of n. The result for small values of n follows by using a sufficiently small universal constant c in the definition of Ω-notation.

We start with a short overview of the original construction of Behrend [10].

Fix a sufficiently large positive integer n. The construction involves a positive integer parameter k that will be determined later. Set $y = n^{1/k}/2$. In what follows we assume that y is an integer. The case that y is not an integer is analyzed later in this section.

Consider independent identically distributed random variables $Y_1, Y_2, \ldots, Y_k$, with each Y_i distributed uniformly over the set $[\{0, y-1\}]$, for all $i \in [\{k\}]$. Set $Z_i = Y_i^2$, for all $i \in [\{k\}]$, and $Z = \sum_{i=1}^{k} Z_i$. It follows that for all $i \in [\{k\}]$,

$$\mathbb{E}(Z_i) \;=\; \sum_{j=0}^{y-1}\frac{1}{y}\cdot j^2 \;=\; \frac{y^2}{3}+\Theta(y) \;.$$

Let $\mu_Z = \mathbb{E}(Z)$ denote the expectation of the random variable Z. It follows that

$$\mu_Z \;=\; \frac{k}{3}y^2+\Theta(k\cdot y) \;. \tag{3.7}$$

Also, for all $i \in [\{k\}]$, $Var(Z_i) = \mathbb{E}(Z_i^2) - \mathbb{E}(Z_i)^2 = \mathbb{E}(Y_i^4) - \frac{1}{9}y^4 + \Theta(y^3)$. Hence

$$Var(Z_i) \;=\; \frac{y^4}{5}+\Theta(y^3)-\frac{y^4}{9}+\Theta(y^3) \;=\; \frac{4}{45}\cdot y^4+O(y^3) \;.$$

Hence

$$Var(Z) \;=\; k\cdot y^4\cdot\frac{4}{45}+O(ky^3) \;=\; k\cdot y^4\cdot\frac{4}{45}\cdot(1+O(1/y)) \;,$$

and the standard deviation of Z, σ_Z, satisfies

$$\sigma_Z \;=\; \sqrt{k}\cdot y^2\cdot\frac{2}{3\cdot\sqrt{5}}\cdot(1+O(1/y)) \;. \tag{3.8}$$

By Chebyshev inequality, for any $a > 0$,

$$\mathbb{P}(|Z-\mu_Z| > a\cdot\sigma_Z) \leq \frac{1}{a^2} \;.$$

Hence, for a fixed value of a, $a > 0$, at least $(1-\frac{1}{a^2})$-fraction of all vectors v from the set $[\{0, y-1\}]^k$ have squared norm $||v||^2$ that satisfies

$$\mu_Z - a\cdot\sigma_Z \leq ||v||^2 \leq \mu_Z + a\cdot\sigma_Z \;.$$

These vectors are now going to be used as numbers in the $(2y)$-ary representation. Since all their coordinates are at most $y-1$, no base-$2y$ carries are needed to add two such numbers.

Note that each vector $v \in [\{0, y-1\}]^k$ has an integer squared norm. By the Pigeonhole Principle, there exists a value T such that $\mu_Z - a \cdot \sigma_Z \leq T \leq \mu_Z + a \cdot \sigma_Z$ such that at least $(1 - \frac{1}{a^2}) \cdot \frac{1}{2a \cdot \sigma_Z + 1} \cdot y^k$ vectors from the discrete cube $C = [\{0, y-1\}]^k$ have squared norm T. Let $\tilde{\mathcal{S}}$ denote the set of these vectors. See Figure 1 for an illustration. By (3.8),

$$\begin{aligned} |\tilde{\mathcal{S}}| &\geq (1 - \frac{1}{a^2}) \cdot \frac{1}{2a\sqrt{k} \cdot y^2 + 1} \cdot \frac{3\sqrt{5}}{2} \cdot (1 - O(\frac{1}{y})) \cdot y^k \\ &\geq \frac{y^{k-2}}{\sqrt{k}} \cdot c \ , \end{aligned}$$

for a fixed positive constant $c = c(a)$. Set $a = 2$. Now $c = c(2)$ is a universal constant, and consequently, $|\tilde{\mathcal{S}}| = \Omega\left(\frac{n^{1-2/k}}{2^k\sqrt{k}}\right)$. To maximize the right-hand-side, we set $k = \lceil \sqrt{2 \cdot \log n} \rceil$. It follows that

$$|\tilde{\mathcal{S}}| \ = \ \Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log n}} \cdot \log^{1/4} n}\right) \ .$$

Observe that all vectors in $\tilde{\mathcal{S}}$ have the same norm $\sqrt{T}$, and thus, for every three vectors $v, u, w \in \tilde{\mathcal{S}}$, $v \neq \frac{u+w}{2}$. To obtain a progression-free set $S \subseteq [\{n\}]$ we consider coordinates of vectors from $\tilde{\mathcal{S}}$ as digits of $(2y)$-ary representation. Specifically, for every vector $v = (v_1, v_2, \ldots, v_k) \in \tilde{\mathcal{S}}$, let $\hat{v} = \sum_{i=0}^{k-1} v_{i+1} \cdot (2y)^i$. The set S is now given by $S = \{\hat{v} \mid v \in \tilde{\mathcal{S}}\}$. Let $f(\cdot) : \tilde{\mathcal{S}} \to S$ denote this mapping.

Note that for every $v \in \tilde{\mathcal{S}}$,

$$0 \ < \ \hat{v} \leq (2y)^k - 1 \ = \ n - 1 \ .$$

Observe also that since $\tilde{\mathcal{S}} \subseteq [\{0, y-1\}]^k$, the mapping f is one-to-one, i.e., if $v \neq u$, for $v, u \in \tilde{\mathcal{S}}$, then $\hat{v} \neq \hat{u}$. Consequently,

$$|S| \ = \ |\tilde{\mathcal{S}}| \ = \ \Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log n}} \cdot \log^{1/4} n}\right) \ .$$

Finally, we argue that S is a progression-free set. Suppose for contradiction that for three distinct numbers $\hat{v}, \hat{u}, \hat{w} \in S$, $\hat{v} = \frac{\hat{u}+\hat{w}}{2}$. Let u, v, w be the corresponding vectors in $\tilde{\mathcal{S}}$, $v = (v_1, v_2, \ldots, v_k)$, $u = (u_1, u_2, \ldots, u_k)$, $w = (w_1, w_2, \ldots, w_k)$. Then

$$\hat{v} \ = \ \sum_{i=0}^{k-1} \frac{u_{i+1} + w_{i+1}}{2} \cdot (2y)^i \ = \ \sum_{i=0}^{k-1} v_{i+1} \cdot (2y)^i \ .$$

However, since all the coordinates $v_1, v_2, \ldots, v_k, u_1, u_2, \ldots, u_k, w_1, w_2, \ldots, w_k$ are in $[\{0, y-1\}]$, it follows that $v_i = \frac{u_i+w_i}{2}$, for every index $i \in [\{k\}]$. Consequently, $v = \frac{u+w}{2}$, a contradiction to the assumption that $||v|| = ||u|| = ||w||$. Hence S is a progression-free set of size $\Omega(\frac{n}{2^{2\sqrt{2}\sqrt{\log n}} \cdot \log^{1/4} n})$.

Consider now the case that $y = \frac{n^{1/k}}{2}$ is not an integer number. In this case the same construction is built with $\lfloor y \rfloor$ instead of y. Set $n' = (2\lfloor y \rfloor)^k$. By the previous argument, we obtain a progression-free set S that satisfies

$$\begin{aligned} |S| &= \Omega\left(\frac{n'}{2^{2\sqrt{2}\sqrt{\log n'}} \cdot \log^{1/4} n'}\right) \\ &= \Omega\left(\frac{n'}{2^{2\sqrt{2}\sqrt{\log n}} \cdot \log^{1/4} n}\right) \ . \end{aligned}$$

Observe that $\frac{n}{n'} \leq \left(\frac{y}{y-1}\right)^k = 1 + \Theta(\frac{k}{y}) = 1 + \Theta\left(\frac{\sqrt{\log n}}{2^{(1/\sqrt{2}) \cdot \sqrt{\log n}}}\right)$.

Hence $|S| = \Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log n}} \cdot \log^{1/4} n}\right)$, and we are done.

4 Our Proof

In this section we present our construction of progression-free sets $S \subseteq [\{n\}]$ with at least $\Omega\left(\frac{n}{2^{2\sqrt{2}\sqrt{\log n}}} \cdot \log^{1/4} n\right)$ elements. Fix $k = \lceil \sqrt{2 \log n} \rceil$, and $y = n^{1/k}/2$. Observe that

$$\begin{aligned} \frac{2^{k/2}}{4} &\leq \frac{1}{2\sqrt{2}} \cdot 2^{\frac{\sqrt{\log n}}{\sqrt{2}}} \\ &\leq y \ \leq \ \frac{1}{2} \cdot 2^{\frac{\sqrt{\log n}}{\sqrt{2}}} \ \leq \ \frac{2^{k/2}}{2} \ . \end{aligned} \tag{4.9}$$

For convenience we assume that y is an integer. If this is not the case, the same analysis applies with minor adjustments. (Specifically, we set $y = \lfloor n^{1/k}/2 \rfloor$. By the same argument as we used in Section 3, the resulting lower bound will be at most by a constant factor smaller than in the case when $n^{1/k}/2$ is an integer.)

Consider the k-dimensional ball centered at the origin that has radius R' given by

$$R'^2 \ = \ \mu_Z \ = \ \frac{k}{3}y^2 + \Theta(ky) \ . \tag{4.10}$$

(See (3.7).) By Chebyshev inequality, the annulus $\hat{\mathcal{S}}$ of all vectors with squared norm in $[R'^2 - 2 \cdot \sigma_Z, R'^2 + 2 \cdot \sigma_Z]$ contains at least $\frac{3}{4} \cdot y^k$ integer lattice points of the discrete cube $C = [\{0, y-1\}]^k$.

The annulus $\hat{\mathcal{S}}$ is far too thick for our needs, and next we "slice" it into many very thin annuli. One of these annuli will be later used to construct the

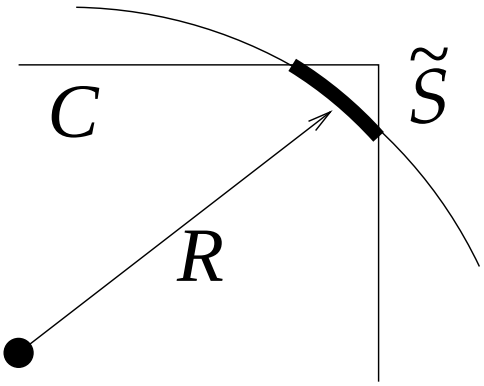

Figure 1: The intersection $\tilde{\mathcal{S}}$ of the discrete cube C with the sphere of radius R is depicted by the bold line.

convexly independent set W that was mentioned in the introduction.

Fix a parameter $g = \epsilon \cdot k$, for a universal constant $\epsilon > 0$ that will be determined later. Partition the (thick) annulus $\hat{\mathcal{S}}$ into $\lceil \frac{4\sigma_Z}{g} \rceil = \ell$ annuli $\hat{\mathcal{S}}_1, \hat{\mathcal{S}}_2, \ldots, \hat{\mathcal{S}}_\ell$, with the annulus $\hat{\mathcal{S}}_i$ containing all vectors with squared norms in the range $[R'^2 - 2\sigma_Z + (i-1) \cdot g, R'^2 - 2\sigma_Z + i \cdot g)$, for $i \in [\{\ell - 1\}]$, and $[R'^2 - 2\sigma_Z + (\ell - 1)\sigma_Z, R'^2 + 2\sigma_Z]$ for $i = \ell$. See Figure 2 for an illustration.

Observe that for distinct indices $i, j \in [\{\ell\}]$, the sets of integer lattice points in $\hat{\mathcal{S}}_i$ and $\hat{\mathcal{S}}_j$ are disjoint. Thus, by the Pigeonhole Principle, there exists an index $i \in [\{\ell\}]$ such that the annulus $\hat{\mathcal{S}}_i$ contains at least

$$\frac{3}{4\ell} \cdot y^k \;=\; \Omega(g \cdot \frac{y^{k-2}}{\sqrt{k}}) \;=\; \Omega(\epsilon\sqrt{k} \cdot y^{k-2}) \tag{4.11}$$

integer lattice points of $C \cap \hat{\mathcal{S}}$. (The first inequality is by (3.8).) In other words, there exists a radius $R = \sqrt{R'^2 - 2\sigma_Z + i \cdot g}$ for some i, that satisfies $R^2 \in [R'^2 - 2\sigma_Z, R'^2 + 2\sigma_Z]$, and such that the annulus $\mathcal{S}$ containing all vectors with squared norm in the range $[R^2 - g, R^2]$ contains at least $\Omega(\sqrt{k} \cdot y^{k-2})$ integer lattice points of $C \cap \hat{\mathcal{S}}$. (The annulus $\mathcal{S}$ is defined by $\mathcal{S} = \{b \in \mathbb{R}^k : ||b||^2 \in [R^2 - g, R^2]\}$.)

By (3.7), (3.8), and (4.10),

$$\begin{aligned} R^2 &\leq R'^2 + 2\sigma_Z \\ &\leq \frac{k}{3} \cdot y^2 + O(k \cdot y) + O(\sqrt{k} \cdot y^2) \\ &\leq \frac{k}{3} \cdot y^2 \left(1 + O\left(\frac{1}{\sqrt{k}}\right)\right) \ . \end{aligned} \tag{4.12}$$

Let $\tilde{\mathcal{S}}$ be the set of integer lattice points of $C \cap \mathcal{S}$. We will show that that $\tilde{\mathcal{S}}$ contains a convexly independent subset $\check{\mathcal{S}}$ with at least $|\check{\mathcal{S}}| \geq \frac{|\tilde{\mathcal{S}}|}{2}$ integer lattice points. Consequently,

$$|\check{\mathcal{S}}| \;\geq\; \frac{|\tilde{\mathcal{S}}|}{2} \;=\; \Omega(\sqrt{k} \cdot y^{k-2}) \;=\; \Omega\left(\log^{1/4} n \cdot \frac{n}{2^{2\sqrt{2}\sqrt{\log n}}}\right) . \tag{4.13}$$

Consider the set $\check{S} = f(\check{\mathcal{S}})$ constructed from $\check{\mathcal{S}}$ by the mapping f described in Section 3. Since $\mathcal{S}$ is a convexly independent set, by the same argument as in Section 3, $|\check{S}| = |\check{\mathcal{S}}|$, and moreover, $\check{S}$ is a progression-free set. Hence $|\check{S}| = \Omega\left(\log^{1/4} n \cdot \frac{n}{2^{2\sqrt{2}\sqrt{\log n}}}\right)$, and our result follows.

The following lemma is useful for showing an upper bound on the number of integer lattice points in $\mathcal{S}$ that do not belong to the exterior set of $\tilde{\mathcal{S}}$, $Ext(\tilde{\mathcal{S}})$. This lemma is due to Coppersmith [16]. Intuitively, this lemma states that for every non-exterior integer lattice point b in our annulus, there necessarily exists a small non-zero integer vector δ which is almost orthogonal to b. See Figure 3 for an illustration.

Let $B_k(R^2)$ denote the k-dimensional ball of radius R centered at the origin, and B denote the set of integer points contained in this ball. Denote $T = R^2$.

LEMMA 4.1. *[16] Let $b \in B \setminus Ext(B)$ be an integer lattice point that satisfies $T - g \leq ||b||^2 \leq T$. Then there exists a non-zero integer vector δ that satisfies $0 \leq \langle b, \delta \rangle \leq g$ and $0 < ||\delta||^2 \leq g$.*

Proof: Since $b \in B \setminus Ext(B)$, there exist integer lattice points $a_1, a_2, \ldots, a_\ell \in B$, for some positive integer $\ell \geq 2$, and constants $p_1, p_2, \ldots, p_l$, $0 < p_1, p_2, \ldots, p_\ell < 1$, such that $\sum_{i=1}^{\ell} p_i = 1$ and $b = \sum_{i=1}^{\ell} p_i \cdot a_i$. Since $a_1, a_2, \ldots, a_\ell \in B$, it follows that $||a_1||^2, ||a_2||^2, \ldots, ||a_\ell||^2 \leq T$. Observe that there exists an index $i \in [\ell]$ such that $\langle a_i, b \rangle$ is greater than or equal to $||b||^2$. (Otherwise, $||b||^2 = \langle \sum_{i=1}^{\ell} p_i \cdot a_i, b \rangle = \sum_{i=1}^{\ell} p_i \cdot \langle a_i, b \rangle < ||b||^2$, contradiction.)

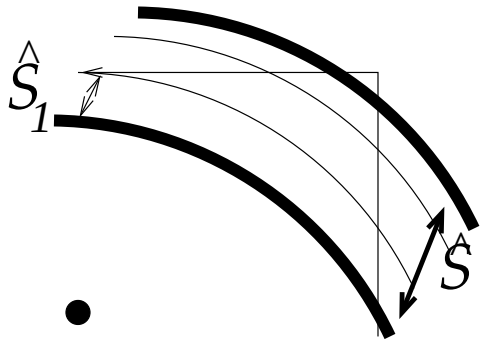

Figure 2: The annulus $\hat{\mathcal{S}}$ is sliced into thin annuli $\hat{\mathcal{S}}_1, \hat{\mathcal{S}}_2, \ldots$.

Suppose without loss of generality that $\langle a_1, b\rangle \geq ||b||^2$. Then $\langle a_1 - b, b\rangle \geq 0$. Set $\delta = a_1 - b$. Since $a_1, b \in B$ are integer lattice points, it follows that δ is an integer lattice point as well. Moreover, since $0 < p_1, p_2, \ldots, p_\ell < 1$, we have $\delta \neq 0$. Moreover,

$$T \;\geq\; ||a_1||^2 \;=\; ||b+\delta||^2 \;=\; ||b||^2 + 2\langle b, \delta\rangle + ||\delta||^2 \;.$$

Recall that $||b||^2 \geq T - g$. Hence $2\langle b, \delta\rangle + ||\delta||^2 \leq g$. As $\langle b, \delta\rangle = \langle a_1 - b, b\rangle \geq 0$, it follows that $\langle b, \delta\rangle, ||\delta||^2 \leq g$, as required. ∎

Since $||\delta||^2 \leq g = \epsilon \cdot k$, and since δ is an integer vector, we conclude that δ may may contain at most $\epsilon \cdot k$ non-zero entries. This property will be helpful for our argument.

Denote the number of integer vectors δ that have squared norm at most g by $\hat{D}(g)$. The next lemma provides an upper bound on $\hat{D}(g)$.

LEMMA 4.2. *For any $\epsilon > 0$ and g as above, there exists $\eta = \eta(\epsilon) > 0$ such that $\lim_{\epsilon\to 0} \eta(\epsilon) = 0$, and $\hat{D}(g) = O(2^{\eta \cdot k})$.*

Proof: Fix an integer h, $1 \leq h \leq g$. First, we count the number $N(h)$ of k-tuples $(q_1, q_2, \ldots, q_k)$ of non-negative integer numbers that sum up to h.

Consider permutations of $(k-1+h)$ elements of two types, with h elements of the first type and $k-1$ elements of the second type. Elements of the first type are called "balls", and elements of the second type are called "boundaries". Two permutations σ and σ' are said to be *equivalent* if they can be obtained one from another by permuting balls among themselves, and permuting boundaries among themselves.

Let Π be the induced equivalence relation. Observe that there is a one-to-one mapping between k-tuples $(q_1, q_2, \ldots, q_k)$ of non-negative integer numbers that sum up to h and the equivalence classes of the relation Π. Hence $N(h)$ is equal to the number of equivalence classes of Π, i.e.,

$$N(h) \;=\; \frac{(k-1+h)!}{(k-1)! \cdot h!} \;=\; \binom{k-1+h}{h} \,.$$

In a k-tuple $(\delta_1, \delta_2, \ldots, \delta_k)$ of integer numbers such that $\sum_{i=1}^k \delta_i^2 = h$, there can be at most h non-zero entries. Hence, for a fixed k-tuple of integers $(q_1, q_2, \ldots, q_k)$ such that $\sum_{i=1}^k q_i = h$, there may be at most 2^h k-tuples $(\delta_1, \delta_2, \ldots, \delta_k)$ of integers such that $\delta_i^2 = q_i$ for every index $i \in [\{k\}]$. Thus, the overall number $D(h)$ of integer k-tuples $(\delta_1, \delta_2, \ldots, \delta_k)$ such that $\sum_{i=1}^k \delta_i^2 = h$ satisfies

$$D(h) \;\leq\; 2^h \cdot N(h) = 2^h \binom{k-1+h}{h} \,.$$

Note that $\binom{k-1+h}{h} \leq \binom{k-1+g}{g}$, for every integer h, $1 \leq h \leq g$. Hence the number $\hat{D}(g)$ of integer k-tuples $(\delta_1, \delta_2, \ldots, \delta_k)$ with $1 \leq \sum_{i=1}^k \delta_i^2 \leq g$ satisfies

$$\begin{aligned}
\hat{D}(g) \;&=\; \sum_{h=1}^{g} D(h) \;\leq\; \sum_{h=1}^{g} 2^h \cdot N(h) \\
&\leq\; N(g) \cdot 2^{g+1} \;\leq\; 2^{g+1} \cdot \binom{k+g}{g} \\
&\leq\; 2^{g+1} \left(\frac{e(k+g)}{g}\right)^g \;=\; 2 \cdot (2e)^g \left(1 + \frac{1}{\epsilon}\right)^{\epsilon \cdot k} \\
&=\; 2 \cdot 2^{(\log 2e + \log(1+\frac{1}{\epsilon}))\epsilon \cdot k} \,.
\end{aligned}$$

Denote $\eta = \eta(\epsilon) = \epsilon(\log 2e + \log(1 + \frac{1}{\epsilon}))$. Then $\hat{D}(g) \leq 2 \cdot 2^{\eta(\epsilon)\cdot k}$. Finally,

$$\lim_{\epsilon\to 0} \eta(\epsilon) \;=\; \lim_{\epsilon\to 0} \frac{\log(1+\frac{1}{\epsilon})}{\frac{1}{\epsilon}} \;=\; \frac{1}{\ln 2} \cdot \lim_{y\to\infty} \frac{\ln(1+y)}{y} \;=\; 0 \,,$$

completing the proof. ∎

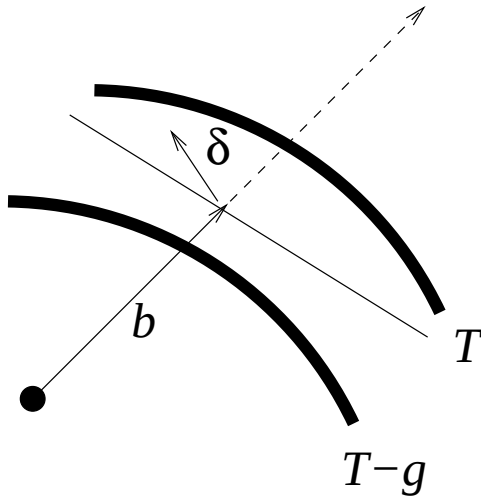

Figure 3: The annulus is depicted by bold curves. The integer vectors b and δ are almost orthogonal, and δ has a very small norm.

Consider again the annulus $\mathcal{S} = \{b \in \mathbb{R}^k : T - g \leq ||b||^2 \leq T\}$, $T = R^2$, and the set $\tilde{\mathcal{S}}$ of integer lattice points of $\mathcal{S}$. For an integer vector δ that satisfies $0 < ||\delta||^2 \leq g$, let $\hat{Z}(\delta)$ denote the set of integer lattice points $b \in \tilde{\mathcal{S}}$ that satisfy $0 \leq \langle b, \delta \rangle \leq g$. Let $\hat{W}(\delta) = \hat{Z}(\delta) \cap C$ denote the intersection of $\hat{Z}(\delta)$ with the discrete cube $C = [\{0, y-1\}]^k$, and let $W(\delta) = |\hat{W}(\delta)|$. Also, let $\hat{W} = \bigcup\{\hat{W}(\delta) : 0 < ||\delta||^2 \leq g\}$, and $W = |\hat{W}|$.

Let $\hat{N}$ denote the set of integer lattice points of $C \cap \mathcal{S}$ that do not belong to $Ext(B)$, and $N = |\hat{N}|$. By Lemma 4.1, $\hat{N} \subseteq \hat{W}$. Consequently,

$$N \ \leq \ W \ \leq \ \sum_{0<||\delta||^2 \leq g} W(\delta) \ . \tag{4.14}$$

Fix a vector δ, $0 < ||\delta||^2 \leq g$. In the sequel we provide an upper bound for $W(\delta)$.

Observe that since $\hat{W}(\delta)$ is a set of integer lattice points, it follows that for every $b \in \hat{W}(\delta)$, $\langle b, \delta \rangle \in [\{0, g\}]$.

For an integer number $h \in [\{0, g\}]$, let $\hat{W}(\delta, h)$ denote the subset of $\hat{W}(\delta)$ of integer lattice points b that satisfy $\langle b, \delta \rangle = h$. Let $W(\delta, h) = |\hat{W}(\delta, h)|$. Observe that for distinct values $h \neq h'$, $h, h' \in [\{0, g\}]$, the sets $\hat{W}(\delta, h)$ and $\hat{W}(\delta, h')$ are disjoint. Consequently,

$$W(\delta) \ = \ \sum_{h=0}^{g} W(\delta, h) \ . \tag{4.15}$$

Next, we provide an upper bound for $W(\delta, h)$.

Consider the hyperplane $\mathcal{H} = \mathcal{H}(\delta, h) = \{\alpha \in \mathbb{R}^k \mid \langle \alpha, \delta \rangle = h\}$. Observe that $\hat{W}(\delta, h) = \mathcal{H} \cap \mathcal{S} \cap C$ is the intersection of the hyperplane $\mathcal{H}$ with the annulus $\mathcal{S}$ and with the discrete cube C.

Let S denote the k-dimensional sphere with squared radius T centered at the origin, i.e., $S = \{\alpha \in \mathbb{R}^k : ||\alpha||^2 = T\}$. Consider the intersection S' of S with the hyperplane $\mathcal{H}$.

LEMMA 4.3. *$S' \subseteq \mathcal{H}$ is a $(k-1)$-dimensional sphere with squared radius $(T - \frac{h^2}{||\delta||^2})$ centered at $\frac{h}{||\delta||^2} \cdot \delta$.*

Proof: For a vector $\alpha \in S \cap \mathcal{H}$,

$$||\alpha - \frac{h}{||\delta||^2} \cdot \delta||^2 \ = \ \sum_{i=1}^{k} (\alpha_i - \frac{h}{||\delta||^2} \cdot \delta_i)^2$$

$$= \ ||\alpha||^2 + \frac{h^2}{||\delta||^2} - 2\frac{h}{||\delta||^2}\langle \alpha, \delta \rangle \ = \ ||\alpha||^2 - \frac{h^2}{||\delta||^2} \ .$$

(For the last equality, note that since $\alpha \in \mathcal{H}$, we have $\langle \alpha, \delta \rangle = h$.) ∎

Recall that for a vector $\alpha \in \mathcal{S}$, $T - g \leq ||\alpha||^2 \leq T$. Hence the intersection of the hyperplane $\mathcal{H}$ with the annulus $\mathcal{S}$ is the $(k-1)$-dimensional annulus $\mathcal{S}' \subseteq \mathcal{H}$, centered at $\frac{h}{||\delta||^2} \cdot \delta$, containing vectors α such that

$$T - g - \frac{h^2}{||\delta||^2} \ \leq \ ||\alpha - \frac{h}{||\delta||^2} \cdot \delta||^2 \ \leq \ T - \frac{h^2}{||\delta||^2} \ .$$

Let $T' = T - \frac{h^2}{||\delta||^2}$. Then $\mathcal{S}'$ is given by

$$\mathcal{S}' \ = \ \{\alpha \in \mathcal{H} : \ T' - g \leq ||\alpha - \frac{h}{||\delta||^2} \cdot \delta||^2 \leq T'\} \ . \tag{4.16}$$

Note that since $h \geq 0$, it follows that $T' \leq T$ for all h and δ. (By definition, it also holds that $\mathcal{S}' = H \cap \mathcal{S}$.)

Recall that our goal at this stage is to provide an upper bound for the number $W(\delta,h)$ of integer lattice points in $\hat{W}(\delta,h) = \mathcal{H} \cap \mathcal{S} \cap C = \mathcal{S}' \cap C$. Let $\mathcal{C} = [0, y-1]^k$ be the (continuous) cube. (The discrete cube $C = [\{0, y-1\}]^k$ is the set of integer lattice points of $\mathcal{C}$.) Let $\tilde{W} = \mathcal{S}' \cap \mathcal{C}$ be the continuous version of $\hat{W}(\delta,h)$. Since $\hat{W}(\delta,h)$ is the set of integer lattice points in $\tilde{W}$, we are interested in providing an upper bound for the number of integer lattice points in $\tilde{W}$. Our strategy is to show an upper bound for the $(k-1)$-dimensional volume $Vol(\tilde{W})$ of $\tilde{W}$, and to estimate the discrepancy between $Vol(\tilde{W})$ and the number of integer lattice points in $\tilde{W}$.

Note that since $\tilde{W} \subseteq \mathcal{S}'$, it follows that $Vol(\tilde{W}) \leq Vol(\mathcal{S}')$. However, this upper bound is too crude. (In particular, it is greater than our lower bound (4.13) on the volume and on the number of integer lattice points in $\check{\mathcal{S}}$.) Instead we will show that if the axes are rotated appropriately, then $\tilde{W}$ becomes contained in the intersection of the annulus $\mathcal{S}'$ with a relatively small number q of octants. Therefore, its volume is at most $Vol(\mathcal{S}') \cdot \frac{q}{2^{k-1}}$. (Because 2^{k-1} is the overall number of octants.) Our estimate for q is $q \leq 2^g = 2^{\epsilon k}$, and thus this upper bound is smaller than the trivial one by a factor of $\frac{2^{\epsilon k}}{2^{k-1}} = 2^{-(1-\epsilon)k+1}$. Since $y = \Theta(2^{k/2})$, this is very significant.

Let $\mathcal{H}' = \{\alpha \in \mathbb{R}^k \mid \langle \alpha, \delta \rangle = 0\}$ be the parallel hyperplane to $\mathcal{H}$ that passes through the origin. Next, we will construct below an orthonormal basis $\Upsilon = \{\gamma_1, \gamma_2, \ldots, \gamma_{k-1}\}$ for $\mathcal{H}'$. The axes will be rotated so that the vectors of Υ will become the new unit vectors of $\mathcal{H}'$.

Recall that δ satisfies $0 < ||\delta||^2 \leq g = \epsilon \cdot k$, and it is an integer vector. Consequently, $\delta = (\delta_1, \delta_2, \ldots, \delta_k)$ contains at most $g = \epsilon \cdot k$ non-zero entries. Let $I \subseteq [\{k\}]$ be the subset of indices such that $\delta_i \neq 0$. Let $m = |I|$. It follows that $m \leq g = \epsilon \cdot k$.

For every vector $\alpha = (a_1, a_2, \ldots, a_k) \in \mathcal{H}'$, it holds that

$$\sum_{i \in I} a_i \delta_i = 0 \ . \tag{4.17}$$

Let $\gamma^{(1)}, \gamma^{(2)}, \ldots, \gamma^{(m-1)}$ be an arbitrary orthonormal basis for the solution space of the equation (4.17). These vectors are in $\mathbb{R}^m$. For each index $j \in [\{m-1\}]$, we view the vector $\gamma^{(j)}$ as $\gamma^{(j)} = (\gamma_i^{(j)} \mid i \in I)$. (In other words, $\gamma_i^{(j)}$ is the ith coordinate of the vector $\gamma^{(j)}$.)

We form orthonormal vectors $\hat{\gamma}^{(1)}, \hat{\gamma}^{(2)}, \ldots, \hat{\gamma}^{(m-1)} \in \mathbb{R}^k$ in the following way. For each index $j \in [\{m-1\}]$, and each index $i \in I$, the ith entry $\hat{\gamma}_i^{(j)}$ of $\hat{\gamma}^{(j)}$ is set as $\gamma_i^{(j)}$, and for each index $i \in [\{k\}] \setminus I$, the entry $\hat{\gamma}_i^{(j)}$ is set as zero. (The vectors $\hat{\gamma}^{(1)}, \hat{\gamma}^{(2)}, \ldots, \hat{\gamma}^{(m-1)}$ agree with vectors $\gamma^{(1)}, \gamma^{(2)}, \ldots, \gamma^{(m-1)}$ on all entries with indices from I, and have zeros in all other entries.) In addition, for each index $j \in [\{k\}] \setminus I$, we insert the vector $\xi_j = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$, $\xi_j \in \mathbb{R}^k$, with 1 at the jth entry and zeros in all other entries into the basis Υ. Observe that $\xi_j \in \mathcal{H}'$, i.e., $\langle \xi_j, \delta \rangle = 0$. The resulting basis Υ is $\{\hat{\gamma}^{(1)}, \hat{\gamma}^{(2)}, \ldots, \hat{\gamma}^{(m-1)}\} \cup \{\xi_j : j \in [\{k\}] \setminus I\}$. It is easy to verify that Υ is an orthonormal basis for $\mathcal{H}'$.

Order the vectors of Υ so that $\hat{\gamma}^{(j)} = \gamma_j$ for all $j \in [\{0, m-1\}]$, and so that the vectors $\{\xi_j : j \in [\{k\}] \setminus I\}$ appear in an arbitrary order among $\gamma_m, \gamma_{m+1}, \ldots, \gamma_{k-1}$.

Move the origin to the center $\frac{h}{||\delta||^2} \cdot \delta$ of the annulus $\mathcal{S}'$, and rotate the annulus so that new axes become the colinear with vectors $\gamma_1, \gamma_2, \ldots, \gamma_{k-1}$ of the orthonormal basis Υ. Obviously, this mapping is volume-preserving.

For a vector $\zeta \in \mathcal{H}'$, let $\zeta_1[\Upsilon], \zeta_2[\Upsilon], \ldots, \zeta_{k-1}[\Upsilon]$ denote the coordinates of ζ with respect to the basis Υ, i.e., $\zeta_i[\Upsilon] = \langle \zeta - \frac{h}{||\delta||^2} \cdot \delta, \gamma_i \rangle$. Observe that since $\langle \delta, \gamma_i \rangle = 0$ for all $i \in [\{k-1\}]$, it follows that $\zeta_i[\Upsilon] = \langle \zeta, \gamma_i \rangle$, for all $i \in [\{k-1\}]$.

LEMMA 4.4. *For a vector $\zeta \in \tilde{W} = \mathcal{S}' \cap \mathcal{C}$, and an index $i \in [\{m, \ldots, k-1\}]$, we have $\zeta_i[\Upsilon] \geq 0$. In particular, ζ has at least $(1-\epsilon) \cdot k$ non-negative coordinates with respect to the basis Υ.*

Proof: Note that for every index $i \in [\{m, k-1\}]$, all entries of γ_i are non-negative. (Because these are the vectors ξ_j, $j \in [\{k\}] \setminus I$ of the standard Kronecker basis.) Since $\zeta \in \mathcal{C} = [0, y-1]^k$, it follows that for all indices $i \in [\{m, k-1\}]$, the ith coordinate of ζ with respect to the basis Υ is non-negative, that is, $\zeta_i[\Upsilon] = \langle \zeta, \gamma_i \rangle \geq 0$. Hence ζ has at least $(k-1) - (m-1) \geq (1-\epsilon) \cdot k$ non-negative coordinates with respect to the basis Υ. ∎

Recall that $\tilde{W} \subseteq \mathcal{S}'$, and the annulus $\mathcal{S}'$ is given by (with respect to the basis Υ)

$$\mathcal{S}' \ = \ \{\alpha \in \mathbb{R}^{k-1} : T' - g \leq ||\alpha||^2 \leq T'\} \ .$$

Let

$$\tilde{Q} \ = \ \{\alpha \in \mathbb{R}^{k-1} : \\ T' - g \leq ||\alpha||^2 \leq T', \ \forall i \in [\{m, k-1\}], \alpha_i[\Upsilon] \geq 0\}$$

be the intersection of $\mathcal{S}'$ with the $(k-m)$ half-spaces $\alpha_i[\Upsilon] \geq 0$, for all $i \in [\{m, k-1\}]$. Let $\mathcal{S}''$ be the intersection of the annulus $\mathcal{S}'$ with the positive octant (with respect to Υ), i.e.,

$$\mathcal{S}'' = \{\alpha \in (\mathbb{R}^+)^{k-1} : T' - g \leq ||\alpha||^2 \leq T'\} \ .$$

It follows that $\tilde{W} \subseteq \tilde{Q}$. Hence $Vol(\tilde{W}) \leq Vol(\tilde{Q}) = 2^{m-1} \cdot Vol(\mathcal{S}'') \leq 2^{\epsilon \cdot k - 1} \cdot Vol(\mathcal{S}'')$.

Next, we provide an upper bound for $Vol(\mathcal{S}'')$.

LEMMA 4.5. *For a sufficiently large integer k,*

$$Vol(\mathcal{S}'') \;\le\; g\cdot\left(\frac{\pi e}{6}\right)^{k/2}\cdot y^{k-3}\cdot 2^{O(\sqrt{k})}\ .$$

Proof: Let $R'=\sqrt{T'}$. Observe that $R'\le R=\sqrt{T}$. Let β_{k-1} be the volume of the $(k-1)$-dimensional ball of unit radius. Then $Vol(\mathcal{S}'')=\frac{1}{2^{k-1}}\beta_{k-1}((T')^{\frac{k-1}{2}}-(T'-g)^{\frac{k-1}{2}})$. Note that

$$\begin{aligned}(R'^2-g)^{\frac{k-1}{2}} &= \left(1-\frac{g}{R'^2}\right)^{\frac{k-1}{2}}\cdot R'^{k-1}\\ &\ge R'^{k-1}\left(1-\frac{g(k-1)}{2R'^2}\right)\\ &\ge R'^{k-1}-R'^{k-3}g\cdot k\ .\end{aligned}$$

Hence by (2.2),

$$\begin{aligned}(4.18)\,Vol(\mathcal{S}'') &\le \frac{1}{2^{k-1}}\cdot k\cdot g\cdot\beta_{k-1}\cdot R'^{k-3}\\ &\le \frac{1}{2^{k-1}}\cdot k\cdot g\cdot\frac{\pi^{\frac{k-1}{2}}}{\Gamma\left(\frac{k+1}{2}\right)}\cdot R^{k-3}\ .\end{aligned}$$

By (4.12), $T=R^2\le\frac{k}{3}\cdot y^2\left(1+O\left(\frac{1}{\sqrt{k}}\right)\right)$, and so $R\le\sqrt{\frac{k}{3}}\cdot y\left(1+O\left(\frac{1}{\sqrt{k}}\right)\right)$. Hence

$$Vol(\mathcal{S}'')\;\le\;\frac{1}{2^{k-1}}\cdot k\cdot g\cdot\frac{\pi^{\frac{k-1}{2}}}{\Gamma\left(\frac{k+1}{2}\right)}\cdot\left(\frac{k}{3}\right)^{\frac{k-3}{2}}\cdot y^{k-3}\cdot 2^{O(\sqrt{k})}\ .$$

By Stirling formula, if $k+1$ is even then for a sufficiently large k,

$$\begin{aligned}\Gamma\left(\frac{k+1}{2}\right) &= \left(\frac{k-1}{2}\right)!\;\ge\;\sqrt{\frac{k-1}{2}}\cdot\frac{\left(\frac{k-1}{2}\right)^{\frac{k-1}{2}}}{e^{\frac{k-1}{2}}}\\ &= e^{1/2}\frac{\left(\frac{k}{2}\right)^{\frac{k}{2}}\left(1-\frac{1}{k}\right)^{\frac{k}{2}}}{e^{k/2}}\;\ge\;\frac{1}{2}\cdot\frac{k^{\frac{k}{2}}}{(2e)^{\frac{k}{2}}}\ .\end{aligned}$$

By (2.4), if $k+1$ is odd then for a sufficiently large k,

$$\begin{aligned}\Gamma\left(\frac{k+1}{2}\right) &= \Gamma\left(\frac{k}{2}+\frac{1}{2}\right)\\ &\ge \frac{\sqrt{\pi}}{2}\left(\frac{k}{2}-1\right)!\\ &\ge \frac{\pi}{\sqrt{2}}\cdot\sqrt{\frac{k}{2}-1}\cdot\frac{\left(\frac{k}{2}-1\right)^{\frac{k}{2}-1}}{e^{\frac{k}{2}-1}}\\ &= \frac{\pi e}{\sqrt{2}}\cdot\frac{1}{\sqrt{\frac{k}{2}-1}}\cdot\frac{(\frac{k}{2}-1)^{k/2}}{e^{\frac{k}{2}}}\\ &\ge \frac{\pi e}{\sqrt{k}}\cdot\frac{\left(\frac{k}{2}\right)^{\frac{k}{2}}\cdot\left(1-\frac{2}{k}\right)^{\frac{k}{2}}}{e^{\frac{k}{2}}}\\ &\ge \frac{1}{\sqrt{k}}\cdot\frac{k^{k/2}}{(2e)^{k/2}}\ .\end{aligned}$$

Hence in both cases, for a sufficiently large k,

$$\Gamma\left(\frac{k+1}{2}\right)\;\ge\;\frac{1}{\sqrt{k}}\cdot\frac{k^{\frac{k}{2}}}{(2e)^{\frac{k}{2}}}\ .$$

Consequently,

$$\begin{aligned}&Vol(\mathcal{S}'')\;\le\\ &\le\;(k\cdot g)\frac{1}{2^{k-1}}\frac{\pi^{\frac{k}{2}}\sqrt{k}(2e)^{\frac{k}{2}}}{\sqrt{\pi}k^{\frac{k}{2}}}\cdot\frac{k^{\frac{k-3}{2}}}{3^{\frac{k-3}{2}}}y^{k-3}\cdot 2^{O(\sqrt{k})}\\ &=\;O(1)\cdot(k\cdot g)\cdot k^{\frac{-3}{2}}\sqrt{k}\cdot\left(\frac{\pi e}{6}\right)^{\frac{k}{2}}\cdot y^{k-3}\cdot 2^{O(\sqrt{k})}\\ &\le\;g\cdot\left(\frac{\pi e}{6}\right)^{\frac{k}{2}}\cdot y^{k-3}\cdot 2^{O(\sqrt{k})}\ .\qquad\blacksquare\end{aligned}$$

We conclude that

$$(4.19)\quad Vol(\tilde{W})\;\le\;Vol(\tilde{Q})\;\le\;2^{\epsilon k-1}\cdot Vol(\mathcal{S}'')\;\le\;\le\;\frac{1}{2}\cdot g\cdot 2^{\epsilon k}\cdot\left(\frac{\pi e}{6}\right)^{\frac{k}{2}}\cdot y^{k-3}\cdot 2^{O(\sqrt{k})}\ .$$

Since $\tilde{W}\subseteq\tilde{Q}$, the number $W(\delta,h)$ of integer lattice points in $\tilde{W}$ is at most the number Q of integer lattice points in $\tilde{Q}$. In Section 6 we will show that Q is not much larger than $Vol(\tilde{Q})$. Specifically,

$$(4.20)\;\;Q\;\le\;k^{O(1)}\cdot 2^{\epsilon k}\cdot\left(\frac{\pi e}{6}\right)^{\frac{k}{2}}\cdot y^{k-3}\cdot 2^{O(\sqrt{k})}\ .$$

We remark that this estimate is quite crude, as it says that the number Q of integer lattice points in $\tilde{Q}$ cannot be larger than by a factor of $k^{O(1)}$ than $Vol(\tilde{Q})$. However, it is sufficient for our argument.

Next, we put all parts together and complete the proof. By (4.20),

$$(4.21)\qquad W(\delta,h)\;\le\;Q\;\le\;\le\;k^{O(1)}\cdot 2^{\epsilon k}\cdot\left(\frac{\pi e}{6}\right)^{\frac{k}{2}}\cdot y^{k-3}\cdot 2^{O(\sqrt{k})}\ .$$

By (4.15),

$$\begin{aligned}W(\delta) &= \sum_{h=0}^{g}W(\delta,h)\;\le\\ &\le\;(g+1)\cdot k^{O(1)}\cdot 2^{\epsilon k}\cdot\left(\frac{\pi e}{6}\right)^{\frac{k}{2}}\cdot y^{k-3}\cdot 2^{O(\sqrt{k})}\ .\end{aligned}$$

Hence by (4.14), the overall number N of integer lattice points in $C\cap\mathcal{S}$ that do not belong to $Ext(B)$ (and thus, do not belong to $Ext(C\cap\mathcal{S})$, because $\mathcal{S}\subseteq B$) satisfies

$$\begin{aligned}N &\le \sum_{0<||\delta||^2\le g}W(\delta)\;\le\\ &\le\;k^{O(1)}\cdot 2^{\epsilon k}\cdot\left(\frac{\pi e}{6}\right)^{\frac{k}{2}}\cdot y^{k-3}\cdot 2^{O(\sqrt{k})}\cdot\hat{D}(g)\ .\end{aligned}$$

Recall that $g \leq k$. By Lemma 4.2, and since for a sufficiently large k, $k^{O(1)} \leq 2^{O(\sqrt{k})}$, it follows that

$$\begin{aligned} N &\leq k^{O(1)} \cdot 2^{\epsilon k} \cdot \left(\frac{\pi e}{6}\right)^{\frac{k}{2}} \cdot y^{k-3} \cdot 2^{O(\sqrt{k})} \cdot O(2^{\eta \cdot k}) = \\ &= 2^{((\epsilon+\eta(\epsilon)+O(1/\sqrt{k}))+\frac{1}{2}\log\frac{\pi e}{6})\cdot k} \cdot y^{k-3} . \end{aligned}$$

By (4.11), the set $\tilde{\mathcal{S}}$ of integer lattice points of $C \cap \mathcal{S}$ contains

$$|\tilde{\mathcal{S}}| = \Omega(\epsilon\sqrt{k} \cdot y^{k-2})$$

integer lattice points. Let $c' > 0$ be a universal constant such that $|\tilde{\mathcal{S}}| \geq c' \cdot \epsilon\sqrt{k} \cdot y^{k-2}$. By (4.9), $y \geq \frac{2^{k/2}}{4}$. Hence the inequality

$$y \geq \frac{2^{k/2}}{4} > 2 \cdot \frac{1}{c' \cdot \epsilon \cdot \sqrt{k}} \cdot 2^{((\epsilon+\eta(\epsilon)+O(1/\sqrt{k}))+\frac{1}{2}\log\frac{\pi e}{6})\cdot k} \quad (4.22)$$

holds whenever $\epsilon > 0$, k is sufficiently large, and ϵ and k satisfy

$$1 > \log\frac{\pi e}{6} + 2 \cdot (\epsilon + \eta(\epsilon)) + O\left(\frac{1}{\sqrt{k}}\right) .$$

By Lemma 4.2, $\lim_{\epsilon \to 0} \eta(\epsilon) = 0$. Thus, for a sufficiently small universal constant $\epsilon > 0$, and sufficiently large k, the inequality (4.22) holds, and thus $|\tilde{\mathcal{S}}| \geq 2N$. (More specifically, one needs to set ϵ so that $0 < 2 \cdot (\epsilon + \eta(\epsilon)) < 1 - \log\frac{\pi e}{6}$.) Hence the set $\tilde{\mathcal{S}}$ contains a subset $\check{\mathcal{S}}$ of integer lattice points that belong to $Ext(B)$, and moreover, by (4.13),

$$\begin{aligned} |\check{\mathcal{S}}| &\geq |\tilde{\mathcal{S}}| - N \geq \frac{1}{2}|\tilde{\mathcal{S}}| = \\ &= \Omega(\epsilon \cdot \sqrt{k} \cdot y^{k-2}) = \\ &= \Omega(\log^{1/4} n \cdot \frac{n}{2^{2\sqrt{2}\sqrt{\log n}}}) . \end{aligned} \quad (4.23)$$

Finally, we argue that our construction can be implemented by a deterministic algorithm that requires time $\frac{n}{2^{\sqrt{2}(1-\eta')\sqrt{\log_2 n}}}$, for an arbitrarily small $\eta' > 0$. The algorithm starts with computing the values $R'^2 - 2 \cdot \sigma_Z + i \cdot g$, for $i \in [\{0, \ell - 1\}]$, and $R'^2 + 2 \cdot \sigma_Z$. (See the beginning of Section 4.) These values determine the minimum and maximum values of vector norms in the ℓ annuli $\hat{\mathcal{S}}_1, \hat{\mathcal{S}}_2, \ldots, \hat{\mathcal{S}}_\ell$. This computation requires $O(\ell) = O(\sigma_Z/g) = O(y) = O(n^{1/k}) = 2^{O(\sqrt{\log n})}$ time. Next, for every point $b \in C = [\{0, y-1\}]^k$, the algorithm tests whether this point belongs to one of the annuli. If it does, it adds the point b into the set of elements of the respective annulus, and increments the size counter that corresponds to this annulus. In the end of this computation (which requires $O(y^k \cdot k)$ time), for each annulus $\hat{\mathcal{S}}_i$ the algorithm knows its size s_i. The algorithm selects the annulus $\hat{\mathcal{S}}_i$ with the greatest size. Then for each point $b \in \hat{\mathcal{S}}_i$, and for every vector δ such that $0 < ||\delta||^2 \leq g$, the algorithm tests whether $0 \leq \langle b, \delta \rangle \leq g$. If this is the case, the algorithm removes the point b from $\hat{\mathcal{S}}_i$. (By Lemma 4.1, the resulting set is convexly independent.) By Lemma 4.2, the number of vectors δ as above is $2^{\eta \cdot k}$, for $\eta > 0$ being an arbitrarily small constant. Thus this computation requires at most $O(y^k \cdot 2^{\eta \cdot k} \cdot k)$ time. Finally, the algorithm outputs the numbers that are associated with the vectors that are left in $\hat{\mathcal{S}}_i$. By (4.23), there are $\Omega(\log^{1/4} n \cdot \frac{n}{2^{2\sqrt{2}\sqrt{\log n}}})$ numbers in this set. Since $n = (2y)^k$, and $k = \lceil\sqrt{2 \cdot \log_2 n}\rceil$, it follows that the overall running time is at most

$$\begin{aligned} O(y^k \cdot 2^{\eta \cdot k}) &\leq O(y^k \cdot 2^{\eta' \cdot k}) = O(\frac{(2y)^k}{2^k} \cdot 2^{\eta' \cdot k}) \\ &= O(\frac{n}{2^{k(1-\eta')}}) = O(\frac{n}{2^{\sqrt{2}(1-\eta')\sqrt{\log_2 n}}}) , \end{aligned}$$

for an arbitrarily small constant $\eta' > \eta > 0$. Note that this running time is sublinear in n, but superlinear in the size of the output.

5 A Convex Hull of the Lattice Points of a Large Ball

Consider a k-dimensional ball $B = B(T)$ of squared radius $T = R^2$ centered at the origin. Suppose that $k \geq 5$ is a fixed constant, and R tends to infinity. (We will later extend the proof to cases $k = 4$ and $k = 3$. For the case $k = 2$ there is a classic construction of Jarnik [27] that can be efficiently implemented.)

In 1963 Andrews [6] has shown that $Ext(B) = O(R^{k-2+\frac{2}{k+1}})$. Much more recently Barany and Larman [9] proved the tight lower bound $Ext(B) = \Omega(R^{k-2+\frac{2}{k+1}})$. For small values of k the problem was studied in [7, 8]. The proof of Barany and Larman [9] is quite elaborate. In particular, it employs the Flatness Theorem of Khintchine [28]. In addition, the proof of [9] does not give rise to an efficient algorithm for constructing a set of $\Omega(R^{k-2+\frac{2}{k+1}})$ convexly independent vectors with norm at most R each. Next, we provide a simple and (almost) self-contained proof that $|Ext(B)| = \Omega(R^{k-2+\frac{2}{k+1}})$. (The proof uses standard estimates for the discrepancy between the volume of a large ball and the number of integer lattice points in it.) Later in the sequel we convert our proof into an efficient algorithm for constructing a large set of convexly independent k-vectors of bounded norm.

Let $g = g(R) > 0$ be a large number that tends to infinity as R grows. The precise dependence of g on R will be determined later. Let $\mathcal{A} = \mathcal{A}(T, g) = B(T) \setminus B(T-g)$ be an annulus with squared radius T and

squared width g centered at the origin. As $\beta_k = \Theta(1)$, the volume of $\mathcal{A}$, $Vol(\mathcal{A})$, satisfies

$$(5.24)\quad Vol(\mathcal{A}) \;=\; Vol(B(T)) - Vol(B(T-g)) \;= \\ =\; \beta_k \cdot (T^{k/2} - (T-g)^{k/2}) \;=\; \Omega(T^{\frac{k}{2}-1} \cdot g)\ .$$

The number $A(\mathcal{A})$ of integer lattice points in $\mathcal{A}$ satisfies $A(\mathcal{A}) \geq A(B(T)) - A(B(T-(g-1)))$. By standard estimates (see, e.g., [4]),

$$(5.25)\quad |A(B(T)) - Vol(B(T))| \;\leq\; O(T^{\frac{k}{2}-1})\ ,$$

and thus,

$$(5.26)\qquad A(\mathcal{A}) \;\geq\; Vol(B(T)) - \\ - Vol(B(T-(g-1))) - O(T^{\frac{k}{2}-1}) \;= \\ =\; \Omega(T^{\frac{k}{2}-1} \cdot g)\ .$$

Next, we show that if g is not too large, then at most a constant fraction of integer lattice points of $\mathcal{A}$ do not belong to $Ext(B)$. This will imply that $Ext(B) = \Omega(A(\mathcal{A})) = \Omega(T^{\frac{k}{2}-1} \cdot g)$.

By Lemma 4.1, it is sufficient to provide an upper bound on the number of integer lattice points in $\bigcup\{\mathcal{A} \cap \mathcal{H}(\delta, h) : 0 < ||\delta||^2 \leq g, 0 \leq h \leq g\}$ (with h and δ being an integer number and vector, respectively). Observe that we need only to count the integer lattice points in $\mathcal{H}(\delta, h)$ for δ with $gcd(\delta) = 1$. This is because for a vector $\delta = c \cdot \delta'$, for an integer $c > 0$, each integer lattice point $\alpha \in \mathcal{H}(\delta, h)$ belongs to the hyperplane $\mathcal{H}(\delta, h/c)$ as well, and $h/c = \langle \delta', \alpha \rangle$ is an integer.

Fix some integer value j, $0 < j \leq g$, and consider a vector δ with $||\delta||^2 = j$. The number $r_k(j)$ of integer vectors δ on the surface of the k-dimensional sphere of squared radius j centered at the origin satisfies $r_k(j) = O(j^{\frac{k}{2}-1})$. (See, e.g., [4].) For each fixed δ and h, the body $(\mathcal{A} \cap \mathcal{H}(\delta, h))$ is a $(k-1)$-dimensional annulus of squared radius at most T and squared width at most g. (See Lemma 4.3 and equation (4.16).) Hence its volume is $O(T^{\frac{k-1}{2}-1} \cdot g)$. Moreover, this annulus is contained in the hyperplane $\mathcal{H} = \mathcal{H}(\delta, h)$, and the lattice of integer points in $\mathcal{H}$ has determinant $||\delta||$. It follows that the number of integer lattice points in $\mathcal{A} \cap \mathcal{H}$ is $\frac{1}{||\delta||} \cdot O(T^{\frac{k}{2}-\frac{3}{2}} \cdot g) = \frac{1}{\sqrt{j}} \cdot O(T^{\frac{k}{2}-\frac{3}{2}} \cdot g)$. Summing up over all possible integer values of h, $0 \leq h \leq g$, we obtain at most

$$|\bigcup_{0 \leq h \leq g} (\mathcal{A} \cap \mathcal{H}(\delta, h))| \;=\; \frac{1}{\sqrt{j}} \cdot O(T^{\frac{k}{2}-\frac{3}{2}} \cdot g^2)$$

integer lattice points.

Since there are $r_k(j) = O(j^{\frac{k}{2}-1})$ possible integer vectors δ with $||\delta||^2 = j$, it follows that

$$|\bigcup\{\mathcal{A} \cap \mathcal{H}(\delta, h) : 0 \leq h \leq g, ||\delta||^2 = j\}| \;= \\ O(j^{\frac{k}{2}-\frac{3}{2}} \cdot T^{\frac{k-3}{2}} \cdot g^2)\ .$$

Finally, the value of j runs over all positive integers that are no greater than g. Hence the total number of integer lattice points in $\mathcal{A} = B(T) \setminus B(T-g)$ that do not belong to $Ext(B(T))$ is at most

$$(5.27)\quad W \;=\; \sum_{j=1}^{g} O(j^{\frac{k}{2}-\frac{3}{2}} \cdot T^{\frac{k-3}{2}} \cdot g^2) \;= \\ =\; O(T^{\frac{k-3}{2}} \cdot g^2) \cdot g^{\frac{k}{2}-\frac{1}{2}} \;=\; O(T^{\frac{k-3}{2}} \cdot g^{\frac{k+3}{2}})\ .$$

By (5.26), $A(\mathcal{A}) = \Omega(g \cdot T^{\frac{k}{2}-1})$. We set g to be the largest value so that the inequality $W \leq \frac{1}{2}A(\mathcal{A})$ holds. In other words, $g = c \cdot T^{\frac{1}{k+1}} = c \cdot R^{\frac{2}{k+1}}$, for a sufficiently small universal constant $c > 0$. (Observe that $g = g(R)$ tends to infinity as R grows.)

For this choice of g, it holds that

$$|Ext(B)| \;\geq\; \frac{1}{2}A(\mathcal{A}) \;=\; \Omega(T^{\frac{k}{2}-1+\frac{1}{k+1}}) \;=\; \Omega(R^{k-2+\frac{2}{k+1}})\ .$$

This proves the theorem of Barany and Larman [9] for $k \geq 5$. Next, we discuss the cases $k = 4$ and $k = 3$. We start with $k = 4$. The right-hand-side in the estimate (5.25) for the discrepancy between $A(B(T))$ and $V(B(T))$ becomes $O(T \cdot \log^{2/3} T)$ [4]. Hence to carry out the proof we will need to set $g = \omega(\log^{2/3} T)$. Also, for $k = 4$, the value of $r_k(j)$ is no longer bounded by $j^{\frac{k}{2}-1}$. However, $r_4(j) = O(\sigma(j))$, where $\sigma(j)$ is the sum of the divisors of j. The function $\sigma(j)$ can be bounded by $\sigma(j) = O(j \log\log j)$. (See, e.g., [26], p.266.)

Hence the inequality (5.27) becomes $W = O(T^{\frac{k-3}{2}} g^{\frac{k+3}{2}} \log\log g)$, and g has to be set as

$$g \;=\; \Theta\left(\frac{T^{\frac{1}{k+1}}}{(\log\log T)^{\frac{2}{k+1}}}\right) = \Theta\left(\frac{T^{1/5}}{(\log\log T)^{2/5}}\right)\ .$$

Hence $|Ext(B)| = \Omega\left(\frac{T^{\frac{k}{2}-1+\frac{1}{5}}}{(\log\log T)^{2/5}}\right) = \Omega\left(\frac{R^{12/5}}{(\log\log R)^{2/5}}\right)$. This estimate is weaker than the optimal estimate (due to Barany and Larman [9]) by a factor of $O((\log\log R)^{2/5})$.

Next, we turn to the case $k = 3$. Here the estimate (5.27) becomes $A(B(T)) - V(B(T)) = O(T^{2/3} \cdot \log^6 T)$ [4]. Hence we need to set g so that $T^{\frac{k}{2}-1} \cdot g = T^{\frac{1}{2}} \cdot g = \omega(T^{2/3} \cdot \log^6 T)$, i.e., $g = \omega(T^{\frac{1}{6}} \log^6 T)$. To provide an upper bound on $r_3(j)$ we use Siegel's Theorem [38].

THEOREM 5.1. *For any $\epsilon > 0$, $r_3(j) = O(j^{\frac{1}{2}+\epsilon})$, where the dependence on ϵ is hidden by the O-notation.*

Propagating this change through our analysis we obtain $W = O(T^{\frac{k-3}{2}} g^{\frac{k+3}{2}+\epsilon})$. Hence $g = \Theta(T^{\frac{1}{k+1+2\epsilon}}) = T^{1/4-O(\epsilon)}$, and $|Ext(B)| = R^{\frac{3}{2}-O(\epsilon)}$. This estimate is

weaker than the estimate of [9] by a factor of $R^{O(\epsilon)}$, where $\epsilon > 0$ is an arbitrarily small constant.

Next, we use our proof to devise an efficient algorithm for constructing large convexly independent sets (henceforth, CISs) of k-dimensional vectors with norm at most R. For $k = 2$ the construction of Jarnik [27] (see also [17]) yields directly an efficient algorithm for constructing such a CIS with an optimal number of $\Omega(R^{2/3})$ vectors. The running time of this algorithm is $O(R^{2/3} \cdot \log R)$, which does not leave much room for improvement. Hence we restrict our attention to the case $k \geq 3$.

The trivial approach to the problem is to invoke one of the algorithms for computing extreme points of a convex hull of an arbitrary set of n points on the set S of integer points of the ball of radius R centered at the origin. Observe that $n = |S| = \Theta(R^k)$. Moreover, the size of the output, that is, the number of vertices in the convex hull of S, is $q = \Theta(R^{k-2+\frac{2}{k+1}}) = \Theta(n^{1-\frac{2}{k}+\frac{2}{k(k+1)}})$. For $k = 3$ there is an algorithm of Preparata and Hong [32] that requires $O(n \cdot \log n) = O(R^3 \cdot \log R)$ time. For $k \geq 4$ the best-known output-sensitive algorithm for computing convex hulls has running time $O(n \log q + (nq)^{1-\frac{1}{\lfloor k/2 \rfloor + 1}} \cdot \log^{O(1)} n)$ [14]. Substituting $q = \Theta(n^{1-\frac{2}{k}+\frac{2}{k(k+1)}})$, we get the bound of $O(n^{2-\frac{6}{k+2}+\frac{2}{(k+1)(k+2)}} \cdot \log^{O(1)} n)$ for even k, and $O(n^{2-2\cdot\frac{3k+1}{(k+1)^2}} \cdot \log^{O(1)} n)$ for odd k.

Next, we describe the algorithmic version of our proof, and analyze its running time. We initialize a set S to contain all k-vectors b of squared norm between $T-g$ and T, i.e., $T-g \leq ||b||^2 \leq T$. For each $b \in S$ and for each integer k-vector δ of squared norm at most g, we compute the scalar product $\langle b, \delta \rangle$ and test whether $0 \leq \langle b, \delta \rangle \leq g$. If it is, we remove b from S. After testing all vectors b the algorithm returns the set S.

More specifically, we fix a vector δ, and for each integer j, $0 \leq j \leq g$, we list all points b that satisfy $T - g \leq ||b||^2 \leq T$ and $\langle b, \delta \rangle = j$. To do it we test for each possible $(k-2)$-tuples $(b_1, b_2, \ldots, b_{k-2})$, $-R \leq b_i \leq R$, for $1 \leq i \leq k-2$, whether there are possible values of b_{k-1} and b_k such that both $T - g \leq ||b||^2 \leq T$ and $\langle b, \delta \rangle = j$ hold for $b = (b_1, b_2, \ldots, b_k)$. For each $(k-2)$-tuple $(b_1, b_2, \ldots, b_{k-2})$, all possible pairs b_{k-1}, b_k as above can be computed in $O(1)$ time. Thus, this step requires $O(R^{k-2})$ time. Since we do it for each of the $g+1$ possible values of j, $0 \leq j \leq g$, for each fixed vector δ the running time is $O(R^{k-2} \cdot g)$. Since there are $O(g^{k/2})$ possible values of vector δ (since it is an integer vector with squared norm at most g), it follows that the total running time is $O(R^{k-2} \cdot g^{\frac{k}{2}+1}) = O(R^{k-2} \cdot R^{\frac{k+2}{k+1}}) = O(R^{k-1+\frac{1}{k+1}}) = O(n^{1-\frac{1}{k+1}})$. (Observe that the output size, which serves also a lower bound on the required time, is $\Omega(R^{k-2+\frac{2}{k+1}}) = \Omega(R^{k(1-\frac{2}{k+1})}) = \Omega(n^{1-\frac{2}{k+1}})$.) See Table 1 (in the introduction) for a concise comparison between the values of the exponent γ in the running time $O(n^\gamma)$ of our algorithm with the running time of the previous best known algorithm (that computes extreme points of the convex hull of an arbitrary body).

6 Discrepancy between Volume and Number of Integer Lattice Points

Consider the annulus $\mathcal{S}' = \{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{k-1}) \in \mathbb{R}^{k-1} \mid T' - g \leq ||\alpha||^2 \leq T'\}$, and its intersection $\tilde{Q}$ with the half-spaces $\alpha_i \geq 0$ for all $i \in [\{m, k-1\}]$. In this section we argue that the number Q (denoted also by $A(\tilde{Q})$) of integer lattice points in $\tilde{Q}$ is not much larger than $Vol(\tilde{Q})$. Specifically, we show that

$$Q \;=\; 2^{O(\sqrt{k})} \cdot 2^{\epsilon k} \cdot \left(\frac{\pi e}{6}\right)^{k/2} \cdot y^{k-3} \;. \tag{6.28}$$

This proves (4.20), and hence completes the proof of our result.

Consider the $(k-1)$-dimensional ball B of squared radius t centered at the origin, for some sufficiently large $t > 0$. Let $A(B)$ denote the number of integer lattice points in B. For a positive integer j, let $V_j(t)$ denote the volume of the j-dimensional ball of squared radius t centered at the origin. It is well-known (see, e.g., the survey of Adhikari [4]) that for a *constant* dimension $k-1$, $|A(B) - V(B)| = O(V_{k-3}(t))$. However, in our case the dimension $k-1$ *grows logarithmically* with t. Fortunately, the following analogous inequality holds in this case:

$$|A(B) - V(B)| \;=\; k^{O(1)} \cdot V_{k-3}(t) \;. \tag{6.29}$$

We prove (6.29) in the sequel. It is worth mentioning that $V_{k-3}(t)$ is almost as large as the volume of the annulus $\mathcal{S}'$, and consequently, one has to provide quite precise estimates for the discrepancy between $A(B)$ and $V(B)$. In particular, a crude estimate of $2^{O(k)} \cdot V_{k-3}(t)$ would not be sufficient for our argument, but rather a polynomial dependence in k is needed, i.e., $k^{O(1)} \cdot V_{k-3}(t)$. Providing such a precise estimate in a $(k-1)$-dimensional space with the dimension growing to infinity logarithmically in the radius of B is technically somewhat involved.

Another subtle point is that we have rotated the vector space to move from the standard Kronecker basis to the orthonormal basis Υ. (In fact, Υ is an orthonormal basis for the hyperplane $\mathcal{H}'$, but it can be completed to an orthonormal basis for $\mathbb{R}^k$ by inserting the unit vector $\frac{\delta}{||\delta||}$ into it.) Consequently, the integer lattice was rotated as well, and so in our context $A(B)$

and $A(\tilde{Q})$ are actually the numbers of points of the *rotated* integer lattice that are contained in B and in $\tilde{Q}$, respectively. It is easy to see that the set of points of the integer lattice that lie in B is in one-to-one correspondence to the set of points of the rotated integer lattice that lie in B. On the other hand, this is not necessarily the case for $\tilde{Q}$. However, we argue below that the estimate (6.29) and its analogue for $\tilde{Q}$ apply for the rotated integer lattice as well, for any rotation.

Recall that $m = |I|$. Let

$$\tilde{Q}_{ext} = \{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{k-1}) \in \mathbb{R}^{k-1} : ||\alpha||^2 \leq T',\ \alpha_i \geq 0 \text{ for all } i \geq m\} \quad (6.30)$$

$$\tilde{Q}_{int} = \{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{k-1}) \in \mathbb{R}^{k-1} : ||\alpha||^2 \leq T' - (g+1),\ \alpha_i \geq 0 \text{ for all } i \geq m\} \quad (6.31)$$

Observe that $\tilde{Q} \subseteq \tilde{Q}_{ext} \setminus \tilde{Q}_{int}$. Also, let $\tilde{Z}$ denote

$$\tilde{Z} = \{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{k-3}) \in \mathbb{R}^{k-3} : ||\alpha||^2 \leq T', \alpha_i \geq 0 \text{ for all } i \geq m\} . \quad (6.32)$$

The set $\tilde{Q}_{ext}$ (respectively, $\tilde{Q}_{int}$) is the intersection of the $(k-1)$-dimensional ball of squared radius T' (resp., $T' - (g+1)$) centered at the origin with the half-spaces $\alpha_i \geq 0$ for all $i \geq m$. The set $\tilde{Z}$ is the intersection of the $(k-3)$-dimensional ball of squared radius T' centered at the origin with the half-spaces $\alpha_i \geq 0$ for all $i \geq m$. The analogue of (6.29) that is required for our argument is

$$|A(\tilde{Q}_{ext}) - Vol(\tilde{Q}_{ext})| = k^{O(1)} \cdot Vol(\tilde{Z}) . \quad (6.33)$$

Given (6.33) we show (6.28) by the following argument.

LEMMA 6.1. $A(\tilde{Q}) = 2^{O(\sqrt{k})} \cdot 2^{\epsilon k} \cdot \left(\frac{\pi e}{6}\right)^{k/2} \cdot y^{k-3}$.

Proof: By (6.33),

$$\begin{aligned} A(\tilde{Q}) &\leq A(\tilde{Q}_{ext}) - A(\tilde{Q}_{int}) \\ &\leq Vol(\tilde{Q}_{ext}) + k^{O(1)} \cdot Vol(\tilde{Z}) - \\ &\quad - Vol(\tilde{Q}_{int}) + + k^{O(1)} \cdot Vol(\tilde{Z}) \\ &= (Vol(\tilde{Q}_{ext}) - Vol(\tilde{Q}_{int})) + k^{O(1)} \cdot Vol(\tilde{Z}). \end{aligned}$$

Observe that $Vol(\tilde{Z}) = \frac{2^{\epsilon k}}{2^{k-3}} \cdot \beta_{k-3} \cdot (T')^{\frac{k-3}{2}}$. Also, since T' is much greater than g,

$$\begin{aligned} &Vol(\tilde{Q}_{ext}) - Vol(\tilde{Q}_{int}) = \\ &= \frac{2^{\epsilon k}}{2^{k-1}} \cdot \beta_{k-1}((T')^{\frac{k-1}{2}} - (T' - (g+1))^{\frac{k-1}{2}}) \\ &\leq O(1) \cdot \frac{2^{\epsilon k}}{2^{k-1}} \cdot \beta_{k-1} \cdot k \cdot (g+1) \cdot (T')^{\frac{k-3}{2}} . \end{aligned}$$

Hence

$$A(\tilde{Q}) \leq O(1) \cdot \frac{2^{\epsilon k}}{2^{k-3}} \cdot (k^{O(1)} \cdot \beta_{k-3} + + k \cdot (g+1) \cdot \beta_{k-1}) \cdot (T')^{\frac{k-3}{2}} .$$

Since $\beta_{k-3} = \Theta(k \cdot \beta_{k-1})$ and $g \leq k$, it follows that

$$A(\tilde{Q}) \leq k^{O(1)} \cdot \frac{2^{\epsilon k}}{2^{k-1}} \cdot \beta_{k-1} \cdot (T')^{\frac{k-3}{2}} .$$

By (4.12), $T' \leq \frac{k}{3} \cdot y^2(1 + O(\frac{1}{\sqrt{k}}))$. Also, $\beta_{k-1} = \frac{\pi^{\frac{k-1}{2}}}{\Gamma(\frac{k+1}{2})}$. Hence

$$A(\tilde{Q}) = 2^{O(\sqrt{k})} \cdot 2^{\epsilon k} \cdot \left(\frac{\pi e}{6}\right)^{k/2} \cdot y^{k-3} . \qquad \blacksquare$$

Hence it remains to prove (6.33). Our proof is closely related to the argument in [22], pp. 94-97, and is provided for the sake of completeness. In addition, our argument is more general than the one in [22], as the latter argument applies only for balls, while our argument applies for intersections of balls with half-spaces.

Fix m to be a positive integer number. (In our application $m = |I|$.) For positive integer numbers k and t, let $Q_k(t)$ denote the intersection of the k-dimensional ball $B_k(t)$ centered at the origin with squared radius t with the half-spaces $\mathcal{H}^{(i)} = \{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k) \mid \alpha_i \geq 0\}$, for all $i \geq m$. Let $\bar{V}_k(t)$ denote the volume $Vol(Q_k(t))$, and $\bar{A}_k(t)$ denote the number of points of the rotated integer lattice in $Q_k(t)$. Note that $\bar{V}_k(t) = \frac{\beta_k}{2^{\max\{k-m+1,0\}}} \cdot t^{k/2}$. The next lemma provides an upper bound for the discrepancy between $\bar{V}_k(t)$ and $\bar{A}_k(t)$ in terms of $\bar{V}_{k-2}(t)$.

LEMMA 6.2. *For a sufficiently large real $t > 0$ and an integer $k \geq 5$,*

$$|\bar{A}_k(t) - \bar{V}_k(t)| = O(k^{3/2} \cdot \bar{V}_{k-2}(t)) .$$

Remark: This lemma applies even if $k = k(t)$ is a function of t.

Before proving Lemma 6.2, we first provide a number of auxiliary lemmas that will be useful for its proof. We start with Euler Sum-formula ([22], Satz 29.1, p.185).

LEMMA 6.3. *For a real-valued function $f(u)$ differentiable in a segment $[a, b]$,*

$$\sum_{a < \ell \leq b} f(\ell) = \int_a^b f(u)du + \psi(a) \cdot f(a) - \psi(b) \cdot f(b) + \int_a^b \psi(u) \cdot f'(u)du ,$$

where $\psi(u) = u - \lfloor u \rfloor - \frac{1}{2}$.

In addition, we will use the following property of the function $\psi(\cdot)$.

LEMMA 6.4. *For any two real numbers κ and λ, $\kappa \le \lambda$, $-1/8 \le \int_\kappa^\lambda \psi(u)du \le 1/8$.*

Proof: The function ψ is periodic with period 1. Its integral is 0 over each complete period. Hence the integral is maximized by setting $\kappa = i + 1/2$ and $\lambda = i + 1$, for some integer i. Hence the integral is at most $1/8$. Analogously, the integral is minimized by setting $\kappa = i$ and $\lambda = i + 1/2$, for some integer i. Hence it is at least $-1/8$. ∎

Next, we use Lemma 6.4 to derive another useful property of the function $\psi(\cdot)$.

LEMMA 6.5. *For a positive real number t and a positive integer $p \ge 2$,*

$$|\int_0^{\sqrt{t}} u \cdot \psi(u)(t-u^2)^{\frac{p}{2}-1}du| \le t^{\frac{p-1}{2}} .$$

Proof: Since $f(u) = u$ is a monotone increasing function, there exists $\xi \in [0, \sqrt{t}]$ such that

$$\int_0^{\sqrt{t}} u\cdot\psi(u)(t-u^2)^{\frac{p}{2}-1}du = \sqrt{t}\int_\xi^{\sqrt{t}} \psi(u)(t-u^2)^{\frac{p}{2}-1}du .$$

Since $g(u) = (t-u^2)^{\frac{p}{2}-1}$ is a monotone decreasing function in $[\xi, \sqrt{t}]$, there exists $\eta \in [\xi, \sqrt{t}]$ such that the right-hand-side is equal to $\sqrt{t}\cdot(t-\xi^2)^{\frac{p}{2}-1}\int_\xi^\eta \psi(u)du$. By Lemma 6.4,

$$|\sqrt{t}(t-\xi^2)^{\frac{p}{2}-1}\int_\xi^\eta \psi(u)du| \le$$
$$\le \sqrt{t}\cdot(t-\xi^2)^{\frac{p}{2}-1} \le \sqrt{t}\cdot t^{\frac{p}{2}-1} = t^{\frac{p-1}{2}} . \quad ∎$$

Note also that

$$|\int_{-\frac{1}{2}}^0 u\cdot\psi(u)(t-u^2)^{\frac{p}{2}-1}du| \le$$
$$\le |\int_{-\frac{1}{2}}^0 (t-u^2)^{\frac{p}{2}-1}du| \le \frac{1}{2}\cdot t^{\frac{p}{2}-1} .$$

Hence for any t and p as above,

$$|\int_{-\frac{1}{2}}^{\sqrt{t}} u\cdot\psi(u)(t-u^2)^{\frac{p}{2}-1}du| \le \quad (6.34)$$
$$\le t^{\frac{p-1}{2}} + \frac{1}{2}\cdot t^{\frac{p}{2}-1} \le$$
$$\le t^{\frac{p-1}{2}}\left(1+\frac{1}{2\sqrt{t}}\right) .$$

We are now ready to prove Lemma 6.2.

Proof of Lemma 6.2:

We prove by induction on k that there exists a universal constant $c > 0$ such that

$$|\bar{A}_k(t) - \bar{V}_k(t)| \le \quad (6.35)$$
$$\le \left(c\cdot\sum_{j=1}^{k-1}\sqrt{j}\right)\left(1+\frac{1}{2\sqrt{t}}\right)\cdot\bar{V}_{k-2}(t) .$$

The constant c will be determined later.

The induction base is $k = 5$. It is well-known (see, e.g., [4]) that $|\bar{A}_5(t) - \bar{V}_5(t)| = O(\bar{V}_3(t)) = O(t^{3/2})$.

Next, we prove the induction step.

In all summations below, ℓ is an integer index. The analysis splits into two cases. The first case is $k+1 < m$, and the second is $k+1 \ge m$. In the first case

$$\bar{A}_{k+1}(t) = \sum_{|\ell|\le\sqrt{t}} \bar{A}_k(t-\ell^2) =$$
$$= \sum_{|\ell|\le\sqrt{t}} \bar{V}_k(t-\ell^2) +$$
$$+ \sum_{|\ell|\le\sqrt{t}} (\bar{A}_k(t-\ell^2) - \bar{V}_k(t-\ell^2)) ,$$

and so

$$|\bar{A}_{k+1}(t) - \sum_{|\ell|\le\sqrt{t}} \bar{V}_k(t-\ell^2)| =$$
$$= |\sum_{|\ell|\le\sqrt{t}} (\bar{A}_k(t-\ell^2) - \bar{V}_k(t-\ell^2))| \le$$
$$\le \sum_{|\ell|\le\sqrt{t}} |\bar{A}_k(t-\ell^2) - \bar{V}_k(t-\ell^2)| .$$

In the second case the same inequalities apply, but the index ℓ runs in the range $0 \le \ell \le \sqrt{t}$ in all summations. It turns out to be more convenient to have the index ℓ vary in the range $-\frac{1}{2} \le \ell \le \sqrt{t}$ rather than $0 \le \ell \le \sqrt{t}$ in these summations.

By the induction hypothesis (that is, by (6.35)),

$$|\bar{A}_k(t-\ell^2) - \bar{V}_k(t-\ell^2)| \le$$
$$\le \left(c\cdot\sum_{j=1}^{k-1}\sqrt{j}\right)\left(1+\frac{1}{2\sqrt{t}}\right)\cdot\bar{V}_{k-2}(t-\ell^2) .$$

Hence

$$|\bar{A}_{k+1}(t) - \sum_{|\ell|\le\sqrt{t}} \bar{V}_k(t-\ell^2)| \le \quad (6.36)$$
$$\le \left(c\cdot\sum_{j=1}^{k-1}\sqrt{j}\right)\left(1+\frac{1}{2\sqrt{t}}\right)\cdot\sum_{|\ell|\le\sqrt{t}}\bar{V}_{k-2}(t-\ell^2) .$$

Next, we estimate $\sum_{|\ell|\le\sqrt{t}}\bar{V}_k(t-\ell^2)$ via Euler Sum-formula (Lemma 6.3). In the first case, we substitute

$a = -\sqrt{t}$, $b = \sqrt{t}$, and $f(u) = \bar{V}_k(t-u^2)$. Then $f(a) = f(b) = \bar{V}_k(0) = 0$, and

$$\frac{df}{du}(u) \;=\; \frac{d}{du}\bar{V}_k(t-u^2) \;= \\ =\; \beta_k \frac{d}{du}(t-u^2)^{\frac{k}{2}} \;=\; -\beta_k \cdot k \cdot (t-u^2)^{\frac{k}{2}-1} u \;.$$

By Lemma 6.3 it follows that

$$\sum_{|\ell|\le\sqrt{t}} \bar{V}_k(t-\ell^2) \;= \\ =\; \int_{\sqrt{t}}^{\sqrt{t}} \bar{V}_k(t-u^2)du - \\ -k\cdot\beta_k \int_{\sqrt{t}}^{\sqrt{t}} \psi(u)(t-u^2)^{\frac{k}{2}-1} u du \;. \tag{6.37}$$

In the second case ($k \ge m-1$), $a = -\frac{1}{2}$, $b = \sqrt{t}$, and again $f(a) = f(b) = 0$. Also,

$$\frac{df}{du}(u) \;=\; -\beta_k \cdot \frac{1}{2^{k-m+1}} \cdot k \cdot (t-u^2)^{\frac{k}{2}-1} u \;,$$

and thus,

$$\sum_{-\frac{1}{2}<\ell\le\sqrt{t}} \bar{V}_k(t-\ell^2) \;= \\ =\; \int_{-\frac{1}{2}}^{\sqrt{t}} \bar{V}_k(t-u^2)du - \\ -k\cdot\frac{\beta_k}{2^{k-m+1}} \int_{-\frac{1}{2}}^{\sqrt{t}} \psi(u)(t-u^2)^{\frac{k}{2}-1} u du \;. \tag{6.38}$$

In the first case, since $h(u) = u\psi(u)$ is an even function on $\mathbb{R}\setminus\mathbb{Z}$, the right-hand-side in (6.37) is equal to

$$\int_{\sqrt{t}}^{\sqrt{t}} \bar{V}_k(t-u^2)du - 2k\cdot\beta_k \int_0^{\sqrt{t}} \psi(u)(t-u^2)^{\frac{k}{2}-1} u du \;.$$

Let J denote $|\int_0^{\sqrt{t}} u\cdot\psi(u)(t-u^2)^{\frac{k}{2}-1}du|$. By Lemma 6.5, $J \le t^{\frac{k-1}{2}}$. Hence

$$\sum_{|\ell|\le\sqrt{t}} \bar{V}_k(t-\ell^2) \;= \\ =\; \int_{\sqrt{t}}^{\sqrt{t}} \bar{V}_k(t-u^2)du - 2k\beta_k\cdot J \;= \\ =\; \bar{V}_{k+1}(t) - 2k\beta_k\cdot J \;.$$

It follows that

$$|\bar{V}_{k+1}(t) - \sum_{|\ell|\le\sqrt{t}} \bar{V}_k(t-\ell^2)| \;= \\ =\; 2k\beta_k\cdot J \;\le\; 2k\beta_k\cdot t^{\frac{k-1}{2}} \;. \tag{6.39}$$

In the second case by (6.38) and since $\int_0^{\sqrt{t}} \bar{V}_k(t-u^2)du = \bar{V}_{k+1}(t)$, it follows that

$$\sum_{-\frac{1}{2}<\ell\le\sqrt{t}} \bar{V}_k(t-\ell^2) \;=\; \int_{-\frac{1}{2}}^{0} \bar{V}_k(t-u^2)du + \\ +\bar{V}_{k+1}(t) - k\cdot\frac{\beta_k}{2^{k-m+1}} \int_{-\frac{1}{2}}^{\sqrt{t}} \psi(u)u(t-u^2)^{\frac{k}{2}-1}du \;.$$

Let J' denote $|\int_{-\frac{1}{2}}^{\sqrt{t}} u\cdot\psi(u)(t-u^2)^{\frac{k}{2}-1}du|$. By (6.34),

$$J' \;\le\; t^{\frac{k-1}{2}} + \frac{1}{2}\cdot t^{\frac{k}{2}-1} \;\le\; t^{\frac{k-1}{2}}\left(1+\frac{1}{2\sqrt{t}}\right) \;. \tag{6.40}$$

Since $\bar{V}_k(t-u^2) \ge 0$ for all u, $-\frac{1}{2} \le u \le 0$, the integral $\int_{-\frac{1}{2}}^{0} \bar{V}_k(t-u^2)du$ is non-negative as well. Thus,

$$|\bar{V}_{k+1}(t) - \sum_{-\frac{1}{2}\le\ell\le\sqrt{t}} \bar{V}_k(t-\ell^2)| \;\le \\ \le\; k\cdot\frac{\beta_k}{2^{k-m+1}}\cdot J' \\ \le\; k\cdot\frac{\beta_k}{2^{k-m+1}}\cdot t^{\frac{k-1}{2}}\left(1+\frac{1}{2\sqrt{t}}\right) \;.$$

In the first case, by the triangle inequality, by (6.36) and (6.39),

$$|\bar{A}_{k+1}(t) - \bar{V}_{k+1}(t)| \;\le\; |\bar{A}_{k+1}(t) - \sum_{|\ell|\le\sqrt{t}} \bar{V}_k(t-\ell^2)| \tag{6.41}$$

$$+\;\; |\sum_{|\ell|\le\sqrt{t}} \bar{V}_k(t-\ell^2) - \bar{V}_{k+1}(t)| \\ \le\;\; \left(c\cdot\sum_{j=1}^{k-1}\sqrt{j}\right) \sum_{|\ell|\le\sqrt{t}} \bar{V}_{k-2}(t-\ell^2) + \\ +\;\; |\sum_{|\ell|\le\sqrt{t}} \bar{V}_k(t-\ell^2) - \bar{V}_{k+1}(t)| \;\le \\ \left(c\cdot\sum_{j=1}^{k-1}\sqrt{j}\right) \sum_{|\ell|\le\sqrt{t}} \bar{V}_{k-2}(t-\ell^2) + 2k\beta_k\cdot t^{\frac{k-1}{2}} \;. \tag{6.42}$$

Analogously, in the second case,

$$|\bar{A}_{k+1}(t) - \bar{V}_{k+1}(t)| \;\le \\ \left(c\cdot\sum_{j=1}^{k-1}\sqrt{j}\right)\left(1+\frac{1}{2\sqrt{t}}\right) \sum_{-\frac{1}{2}\le\ell\le\sqrt{t}} \bar{V}_{k-2}(t-\ell^2) + \\ +\;\; 2k\cdot\frac{\beta_k}{2^{k+2-m}}\cdot t^{\frac{k-1}{2}}\cdot\left(1+\frac{1}{2\sqrt{t}}\right) \;. \tag{6.43}$$

However, in the first case $\sum_{|\ell|\leq\sqrt{t}} \bar{V}_{k-2}(t-\ell^2) \leq \int_{-\sqrt{t}}^{\sqrt{t}} \bar{V}_{k-2}(u)du = \bar{V}_{k-1}(t)$. In the second case,

$$\begin{aligned}&\sum_{-\frac{1}{2}\leq\ell\leq\sqrt{t}} \bar{V}_{k-2}(t-\ell^2) = \\ &= \sum_{0\leq\ell\leq\sqrt{t}} \bar{V}_{k-2}(t-\ell^2) \leq \\ &\leq \int_0^{\sqrt{t}} \bar{V}_{k-2}(u)du = \bar{V}_{k-1}(t) \ .\end{aligned}$$

(In the first case the $(k-1)$st coordinate varies between $-\sqrt{t}$ and $\sqrt{t}$, while in the second case it is non-negative, and thus varies between 0 and $\sqrt{t}$.) Hence in both cases the first terms in (6.42) and in the right-hand-side of (6.43) are at most

$$\left(c\cdot\sum_{j=1}^{k-1}\sqrt{j}\right)\left(1+\frac{1}{2\sqrt{t}}\right)\bar{V}_{k-1}(t) \ .$$

Consequently, in both cases,

$$\begin{aligned}&|\bar{A}_{k+1}(t)-\bar{V}_{k+1}(t)| \leq \\ &\left(c\cdot\sum_{j=1}^{k-1}\sqrt{j}\right)\left(1+\frac{1}{2\sqrt{t}}\right)\cdot\bar{V}_{k-1}(t)+ \\ +\quad &2k\cdot\frac{\beta_k}{2^{\max\{k+2-m,0\}}}\cdot t^{\frac{k-1}{2}}\cdot\left(1+\frac{1}{2\sqrt{t}}\right) \ .\end{aligned}$$

By (2.2), $\beta_k = \Theta\left(\frac{\beta_{k-1}}{\sqrt{k}}\right)$. Set c to be a universal constant such that $c \geq \frac{\sqrt{k}\cdot\beta_k}{2\beta_{k-1}}$, for all integer $k \geq 2$. Then

$$\begin{aligned}&|\bar{A}_{k+1}(t)-\bar{V}_{k+1}(t)| \leq \\ &\leq \left(c\cdot\sum_{j=1}^{k-1}\sqrt{j}\right)\left(1+\frac{1}{2\sqrt{t}}\right)\cdot\bar{V}_{k-1}(t)+ \\ &c\cdot\sqrt{k}\cdot\left(1+\frac{1}{2\sqrt{t}}\right)\cdot\frac{\beta_{k-1}}{2^{\max\{(k-1)-m+1,0\}}}\cdot t^{\frac{k-1}{2}} \\ &= \left(c\cdot\sum_{j=1}^{k}\sqrt{j}\right)\left(1+\frac{1}{2\sqrt{t}}\right)\bar{V}_{k-1}(t) \ .\end{aligned}$$

Finally, $\sum_{j=1}^{k}\sqrt{j} \leq k^{3/2}$, completing the proof. ∎

7 Fast Matrix Multiplication

In this section we argue that our improved construction of progression-free subsets can be used to improve slightly the state-of-the-art matrix multiplication algorithm of Coppersmith and Winograd [18]. Specifically, the upper bound of [18] on the running time of their algorithm is $n^\omega\cdot\zeta(n)$, with $\omega < 2.376$ being a universal constant and $\zeta(n)$ being a function such that $\zeta(n) = n^{o(1)}$. In what follows we show that by plugging in our construction instead of the construction of Behrend in the algorithm of [18] one can improve their upper bound by a factor of $\log^\delta n$, for some small universal constant $\delta > 0$.

The following variant of the result of Behrend [10] (or, actually, of [36]) is used in [18].

THEOREM 7.1. *[36] Given $\epsilon > 0$ there exists an integer M_ϵ such that for all $M \geq M_\epsilon$ there exists a progression-free subset $B \subseteq [\{M\}]$ of size $M' > M^{1-\epsilon}$.*

There are three different variants of the same matrix multiplication algorithm described in [18]. The simplest of them achieves $\omega < 2.404$. The lightly more elaborate intermediate variant achieves $\omega < 2.388$, and, finally, the most elaborate variant achieves $\omega < 2.376$. Each of these variants can be slightly improved by using our progression-free subsets instead of those provided by Behrend construction. We will argue this for the two variants of the algorithm, specifically, for the intermediate and the most elaborate ones.

The inequality that governs the running time of the intermediate variant of the algorithm of [18] is

$$(q+2)^3 \geq \frac{27\cdot q^{(1-\beta)\omega}}{\beta^\beta(1+\beta)^{1+\beta}(2-2\beta)^{2-2\beta}}\cdot c'^\epsilon \ , \tag{7.44}$$

where q and β are parameters, w is the exponent of the resulting algorithm, $\epsilon > 0$ is an arbitrarily small constant from Theorem 7.1, and $0 < c' < 1$ is a universal constant. (The only unknown in this inequality is ω.) As shown in [18], if $\epsilon = 0$ then the minimum value ω^* of ω for which the inequality (7.44) holds is achieved by setting $q = 6$ and $\beta = 0.048$. This value is $\omega^* < 2.387$.

By a simple calculation it follows that for an arbitrary $\epsilon > 0$, the minimum value of ω for which this inequality holds is $\omega = \omega^* + \Theta(\epsilon)$. The running time of the intermediate variant of the algorithm of [18] is then $n^{\omega^*+\Theta(\epsilon)}\cdot\tau(n)$, where $\tau(n) = n^{o(1)}$ is a function from the Schonhage theorem (see [18], p.253). Let c denote the specific constant hidden by the Θ-notation above. Consequently, this running time can be presented as $n^{\omega^*}\cdot\tau(n)\cdot(n^\epsilon)^c$. We denote $\eta(n) = (n^\epsilon)^c$. Since this analysis assumes only that for any sufficiently large M there exists a progression-free subset of $[\{M\}]$ of size at least $M^{1-\epsilon}$, it follows from the result of Behrend [10] that one can set here ϵ such that $n^\epsilon = O(2^{\sqrt{8}\sqrt{\log_2 n}}\cdot\log^{1/4} n)$, and then $\eta(n)$ becomes $(2^{\sqrt{8}\sqrt{\log_2 n}}\cdot\log^{1/4} n)^c$. By using our result we get a

slightly smaller value ϵ' of ϵ, that is, the value that satisfies $n^{\epsilon'} = O(\frac{2^{\sqrt{8}\sqrt{\log_2 n}}}{\log^{1/4} n})$, and a slightly smaller function $\eta'(n)$ given by $\eta'(n) = (\frac{2^{\sqrt{8}\sqrt{\log_2 n}}}{\log^{1/4} n})^c$. Hence $\eta'(n) = \frac{\eta(n)}{\log^{c/2} n}$, and the running time of the intermediate variant of [18] becomes $n^{\omega^*} \cdot \tau(n) \cdot \eta'(n) = n^{\omega^*} \cdot \tau(n) \cdot \frac{\eta(n)}{\log^{c/2} n}$. Hence this running time is better by a factor of $\log^{c/2} n$ than the original one. The constant $\delta = c/2 > 0$ is a (small) universal constant.

Now, we turn to the most elaborate variant of the algorithm of [18]. On the bottom of p.268 it is stated that after pruning there are "approximately" $\binom{n}{A_0,A_1,A_2,A_3,A_4}$ triples of remaining blocks, where $A_0, \ldots, A_4$ are parameters of the algorithm. By inspection of the pruning process described on p.263 of [18] we obtain that the more precise form of this expression is $\binom{n}{A_0,A_1,A_2,A_3,A_4} \cdot c'^{\epsilon n}$, where c' and ϵ are as above. Hence the inequality that governs the running time of this variant of the algorithm (see p.269 of [18]) becomes

(7.45)

$$(q+2)^2 \;\geq\; \frac{(2q)^{2\omega \cdot \hat{b}} (q^2+2)^{\omega \cdot \bar{c}} [4q^\omega (q^\omega + 2)]^{\bar{d}}}{F \cdot G \cdot H \cdot L} \cdot c'^{\epsilon} \;,$$

where

$$\begin{aligned} F &= (2\bar{a} + 2\bar{b} + \bar{c})^{2\bar{a}+2\bar{b}+\bar{c}} \;, \\ G &= (2\bar{b} + 2\bar{d})^{2\bar{b}+2\bar{d}} \;, \\ H &= (2\bar{c} + \bar{d})^{2\bar{c}+\bar{d}} \;, \\ L &= (2\bar{b})^{2\bar{b}} \bar{a}^{\bar{a}} \;, \end{aligned}$$

and where $q, \bar{a}, \bar{b}, \bar{c}, \bar{d}$ are parameters that satisfy certain constraints specified in [18], ω is the exponent of the resulting algorithm, and $0 < c' < 1$ is a universal constant. Again, let ω^* be the minimum value of ω for which the inequality (7.45) holds when $\epsilon = 0$. The specific values of the parameters for which this minimum value is achieved are listed in [18]. This minimum value satisfies $\omega^* < 2.376$.

Similarly to the intermediate variant, for an arbitrary value of $\epsilon > 0$, the minimum value ω for which the inequality (7.45) holds is $w = w^* + c \cdot \epsilon$, for some universal positive constant c. Thus, the overall running time of the algorithm is $n^{\omega^* + c\epsilon} \cdot \tau(n)$. Using the construction of Behrend one can write here $n^\epsilon = O(2^{\sqrt{8}\sqrt{\log_2 n}} \cdot \log^{1/4} n)$, and using our construction we get $n^\epsilon = O(\frac{2^{\sqrt{8}\sqrt{\log_2 n}}}{\log^{1/4} n})$. Hence the overall running time of this variant of the algorithm of [18] when using our construction of progression-free sets is smaller by a factor of $\log^{c/2} n = \log^\delta n$ than when using the construction of Behrend, for $\delta = c/2$.

To summarize, we have shown that our improvement of the construction of Behrend implies an improvement by a factor of $\log^\delta n$, for some small universal constant $\delta > 0$, of the best known estimates on the running time of the state-of-the-art algorithms for matrix multiplication.

8 Conclusion

In this paper we improved the lower bound of Behrend by a factor of $\Theta(\sqrt{\log n})$. As was already mentioned, both Behrend's and our proof arguments rely on the Pigeonhole Principle. It is reasonable to believe that by choosing $T = R^2 = \mu_Z$ (see (4.10)) one can get an annulus with at least as many integer points as in the annulus $\mathcal{S}$ chosen via the Pigeonhole Principle. To prove that this is the case one should probably use normal approximation of the discrete random variable Z (see Sections 3 and 4), and employ probablistic estimates to argue that the probability that Z is between $(\mu_Z - \frac{\epsilon k}{2})$ and $(\mu_Z + \frac{\epsilon k}{2})$ is at least as large as the probability that it is between $(\mu_Z - 2\sigma_Z)$ and $(\mu_Z + 2\sigma_Z)$, divided by $\frac{\epsilon k}{4\sigma_Z}$. Although this appears to be quite clear intuitively, so far we were not able to find sufficiently precise probabilistic estimates to prove this statement formally. Once this intuition is formalized, our construction will become independent of the Pigeonhole Principle. This, in turn, would be a significant improvement of the lower bound of Moser [30].

Acknowledgements

The author is indebted to Don Coppersmith, who was offered a coauthorship on this paper. In particular, Lemma 4.1 is due to Don. In addition, fingerprints of Don can be found in numerous other places in this paper. The author is also grateful to Benny Sudakov for introducing him to the problem.

References

[1] H. Abbott. On a conjecture of Erdős and Straus on non-averaging sets of integers. In *Proc. of the 5th British Combinatorial Conference, Congress Numerartium XV*, pages 1–4, 1975.

[2] H. Abbott. Extremal problems on non-averaging and non-dividing sets. *Pacific J. Math*, 91, 1980.

[3] H. Abbott. On the Erdős-Straus non-averaging set problem. *Acta Math. Hungar.*, 47:117–119, 1986.

[4] S. D. Adhikari. Lattice points in spheres. *Bulletin of the Allahabad Mathematical Society*, 8-9:1–13, 1993-1994.

[5] N. Alon and A. Shapira. Linear equations, arithmetic progressions, and hypergraph property testing. In *Proc. of the 16th Annual Symp. on Discr. Algorithms*, pages 708–717, 2005.

[6] G. E. Andrews. A lower bound for the volumes of strictly convex bodies with many boundary points. *Trans. Amer. Math. Soc.*, 106:270–279, 1963.

[7] V. I. Arnold. Statistics of integer convex polytopes. *Funk. Anal. Pril. (in Russian)*, 14(1):1–3, 1980.

[8] A. Balog and I. Barany. On the convex hull of the integer points in a disc. *DIMACS Series on Discrete and Computational Geometry*, 6:39–44, 1991.

[9] I. Barany and D. Larman. The convex hull of the integer points in a large ball. *Mathematische Annalen*, 312:167 – 181, 1998.

[10] F. Behrend. On sets of integers which contain no three terms in arithmetic progression. *Proc. Nat. Acad. Sci.*, 32:331–332, 1946.

[11] A. Bosznay. On the lower estimation of non-averaging sets. *Acta Math. Hungar.*, 53:155–157, 1989.

[12] J. Bourgain. On triples in arithmetic progression. *GAFA*, 9:968–984, 1999.

[13] J. Bourgain. Roth's theorem in progressions revisited. *manuscript*, 2007.

[14] T. Chan. Output-sensitive results on convex hulls, extreme points, and related problems. *Discrete Computational Geometry*, 16:369–387, 1996.

[15] A. Chandra, M. Furst, and R. Lipton. Multi-party protocols. In *Proc. of the 24th Annual Symp. on Foundations of Computer Science, FOCS'83*, pages 94–99, 1983.

[16] D. Coppersmith. Personal communication, 2003.

[17] D. Coppersmith and M. Elkin. Sparse source-wise and pair-wise distance preservers. In *SODA: ACM-SIAM Symposium on Discrete Algorithms*, pages 660–669, 2005.

[18] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation*, 9:251–280, 1990.

[19] M. Elkin. An improved construction of progression-free sets. arXiv:0801.4310v1, 2008.

[20] P. Erdős and P. Turán. On some sequences of integers. *J. London Math. Society*, 11:261–264, 1936.

[21] G. A. Freiman. Inverse problems of additive number theory. *Izv. Akad. Nauk SSSR (Russian) Ser. Mat.*, 19:275–284, 1955.

[22] F. Fricker. *Einführung in die Gitterpunktlehre*. Birkhauser, 1982.

[23] W. Gasarch, J. Glenn, and C. Kruskal. Finding large 3-free sets I: The small n case. *Journal of Computer and System Sciences*, 74:628–655, 2008.

[24] B. Green and T. Tao. New bounds for Szemerédi's theorem, II: A new bound for $r_4(n)$. *manuscript*, 2008.

[25] B. Green and J. Wolf. A note on Elkin's improvement of Behrend's construction. arXiv:0810.0732v1, 2008.

[26] G. Hardy and E. Wright. *An Introduction to the Theory of Numbers*. Oxford Science Publications (5th Edition), 2004.

[27] V. Jarnik. Über Gitterpunkte und konvex Kurven. *Math. Z.*, 2:500–518, 1925.

[28] A. Khintchine. A qualitative formulation of Kronecker's theory of approximation. *Izv. Akad. Nauk. SSSR Ser. Mat. (in Russian)*, 12:113–122, 1948.

[29] I. Laba and M. T. Lacey. On sets of integers not containing long arithmetic progressions. *ArXiv Mathematics e-prints*, Aug. 2001.

[30] L. Moser. On non-averaging sets of integers. *Canadian J. Math.*, 5:245–253, 1953.

[31] K. O'Bryant. Sets of integers that do not contain long arythmetic progressions. arXiv:0811.3057v2, 2008.

[32] F. P. Preparata and S. J. Hong. Convex hulls of finite sets of points in two and three dimensions. *Communications of the ACM*, 20:87–93, 1977.

[33] R. Rankin. Sets not containing more than a given number of terms in arithmetic progression. *Proc. Roy. Soc. Edinburgh Section A*, 65:332–344, 1960.

[34] K. Roth. On certain sets of integers. *J. London Math. Society*, 28:245–252, 1953.

[35] I. Ruzsa. Solving a linear equation on a set of integers I. *Acta Arithmetica*, 65:259–282, 1993.

[36] R. Salem and D. Spencer. On sets of integers which contain no three in arithmetic progression. *Proc. Nat. Acad. Sci. (USA)*, 28:561 – 563, 1942.

[37] A. Shapira. Behrend-type constructions for sets of linear equations. *Acta Arithmetica*, 122:17–33, 2006.

[38] C. Siegel. Über die Klassenzahl quadratischer Zahlkörper. *Acta Arithmetica*, 1:83–86, 1935.

[39] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arithm.*, 27:299–345, 1975.

[40] B. L. van der Waerden. Beweis einer Baudetischen Vermutung. *Nieuw Arch. Wiskunde*, 2(15):212–216, 1927.