Rev. Mat. Iberoam. **33** (2017), no. 4, 1351–1367
DOI 10.4171/RMI/974

Various approaches for the study of the complexity of some families of pseudorandom subsets

Cécile Dartyge and Domingo Gómez-Pérez

Abstract. Studying randomness in different structures is important from the development of applications and theory. Dartyge, Mosaki and Sárközy (among others) have studied measures of randomness for families of subsets of integers. In this article, we improve results on the complexity of some families defined by polynomials, introducing new techniques from areas such as combinatorial geometry, geometry of numbers and additive combinatorics.

1. Introduction

Randomness is required by a large number of applications in areas like watermarking, wireless communications and simulation. Therefore, it is necessary to evaluate and compare randomness associated to different structures.

Sequences are the most studied structures by far. In 1996, Mauduit and Sárközy [14] introduced two measures for binary sequences which formalize the "pseudorandomness" of a sequence. These measures were adapted by Dartyge and Sárközy [7] to the context of subsets of $\{1,\dots,N\}$ or of $\mathbb{Z}/n\mathbb{Z}$. We include here the definitions for completeness of the paper.

Given $\mathcal{R} \subset \{1,\dots,N\}$, we associate to this set the corresponding sequence $\{e_n\}_{1\le n\le N}$ defined by:

$$e_n = \begin{cases} 1-|\mathcal{R}|/N & \text{for } n \in \mathcal{R}, \\ -|\mathcal{R}|/N & \text{for } n \notin \mathcal{R}. \end{cases}$$

Identifying the set $\mathcal{R}$ with the sequence $\{e_n\}_{1\le n\le N}$ allow Dartyge and Sárközy to introduce for sets the well distribution measure and the correlation measure of

Mathematics Subject Classification (2010): Primary 11B75; Secondary 11T06.
Keywords: Pseudorandomness, complexity.

order k, where $k \in \mathbb{N}$, $k \geq 2$. These measures are respectively defined by

$$W(\mathcal{R}, N) = \max_{a,b,t} \Big| \sum_{j=0}^{t-1} e_{a+jb} \Big|, \qquad C_k(\mathcal{R}, N) = \max_{M,D} \Big| \sum_{n=1}^{M} e_{n+d_1} \cdots e_{n+d_k} \Big|,$$

where in $W(\mathcal{R}, N)$ the maximum is taken over all $a, b, t \in \mathbb{N}$ such that $1 \leq a \leq a + (t-1)b \leq N$ and in the $C_k(\mathcal{R}, N)$ the maximum is over all $M \in \mathbb{N}$, $D = (d_1, d_2, \ldots, d_k) \in \mathbb{Z}^k$ such that $0 \leq d_1 < d_2 < \ldots < d_k \leq N - M$. It is trivial to see that the well distribution measure and the correlation measure are less than N. So, a set $\mathcal{R}$ has **strong pseudorandom properties** if, for $k = O(\log(N))$, both $W(\mathcal{R}, N)$ and $C_k(\mathcal{R}, N)$ are of order $o(N)$, where o represents the Landau symbol o.

Some constructions of large families of subsets with strong pseudorandom properties can be found in [5], [6], [8]. In some applications it is also necessary to know that these families have a "rich structure", in particular one need to be sure that no subset of a given family is characterized by a few number of its elements.

In [1], Ahlswede, Khachatrian, Mauduit and Sárközy introduced the notion of complexity of families of binary sequences. This definition was adapted by Dartyge, Mosaki and Sárközy in [5] for families of subsets. We recall below this definition.

Definition 1.1. Let $\mathcal{F}$ be a family of subsets of $\{1, \ldots, N\}$. The family complexity $K(\mathcal{F})$ of $\mathcal{F}$ is the greatest $k \in \mathbb{N}$ such that for every $\mathcal{A} \in \{1, \ldots, N\}$ with $|\mathcal{A}| = k$ and every subset $\mathcal{B}$ of $\mathcal{A}$ there is an $\mathcal{R} \in \mathcal{F}$ such that $\mathcal{R} \cap \mathcal{A} = \mathcal{B}$.

In the same article, it is observed that for any given family of subsets $\mathcal{F}$, we have $K(\mathcal{F}) \leq \lfloor \log_2 \mathcal{F} \rfloor$, where $\log_2$ denotes the binary logarithm.

In [4] and [5], different results where obtained by mainly two different ways: exponential sums (also called character sums) and theorems of additive number theory. In this paper we propose new contributions on this problem, which can be found in Section 2. The proofs will use various ingredients: hyperplane arrangements, lattices, and sumsets estimates. This paper will be autocontained so Section 3 gives a brief introduction to hyperplane arrangements. In Section 4, we give the background necessary to understand the proofs related with geometry of numbers in Section 7. Sumsets estimates and the combinatorial Nullstellensatz are discussed in Section 5. The last sections are devoted to the proofs of our theorems.

1.1. Notation

Throughout the paper, we use the Landau symbols O and o and the Vinogradov symbol $\ll$. We recall that the assertions $U = O(V)$ and $U \ll V$ (sometimes we write this also as $V \gg U$) are both equivalent to the inequality $|U| \leq cV$ with some constant $c > 0$, while $U = o(V)$ means that $U/V \to 0$. In this paper, the dependence of the constants implied in the symbols $O, \ll$ are indicated using subscripts. As in the introduction, we denote by $\log_2$ the binary logarithm.

Let p be a prime number and $\mathbb{F}_p$ be the finite field of p elements. The elements of $\mathbb{F}_p$ are identified with the integers in the range $\{0, \ldots, p-1\}$. For an element $x \in \mathbb{F}_p$, we define $1/x = x^{-1} = x^{p-2}$. Abusing the notation, x can represent an integer or an indeterminate, so $\mathbb{F}_p[x]$ represents the ring of polynomials

with coefficients in $\mathbb{F}_p$ in the indeterminate x. We also consider $\mathbb{F}_p[x_1, \ldots, x_n]$, the ring of multivariate polynomials in $x_1, \ldots, x_n$ with coefficients in $\mathbb{F}_p$.

As in [4], we focus on sets $\mathcal{R} \subset \mathbb{F}_p$ of the following shape:

$$\mathcal{R} = \{n \in \{1, \ldots p\} : \exists h \in S \text{ such that } f(n) \equiv h \mod p\},$$

where $S \subset \mathbb{F}_p$ and $f(x) \in$ are given.

In this paper, we will use the following notation:

- $S_1 = \{1, \ldots, \Delta\}$ and $S_2 = \{1, 2^{-1}, \ldots, \Delta^{-1}\}$ with $\Delta < p/2$.
- $\mathcal{P}_1(d, p)$ is the set of polynomials of degree at most d with coefficients in $\mathbb{F}_p$.
- $\mathcal{P}_2(d, p)$ is the set of polynomials of degree at most d with coefficients in $\mathbb{F}_p$ without multiple roots.
- $\mathcal{P}_3(d, p)$ is the set of polynomials of degree at most d with coefficients in $\mathbb{F}_p$ which factorize in $\mathbb{F}_p$.
- $\mathcal{R}(f, S) = \{n \in \{0, \ldots, p-1\} \ : \ \exists h \in S, \ f(n) = h\}$.
- For $i = 1, 2, 3$, we define $\mathcal{F}_i(S, d) = \{\mathcal{R}(f, S_i) \ : \ f \in \mathcal{P}_i(d, p)\}$.
- For $i = 1, 2, 3$, $K(\mathcal{F}_i(S, d))$ is the complexity corresponding to $\mathcal{F}_i(S, d)$.

Notice that it is easy to see that $K(\mathcal{F}_3(S, d)) \leq K(\mathcal{F}_2(S, d)) \leq K(\mathcal{F}_1(S, d))$. So upper bounds in $K(\mathcal{F}_1(S, d))$ give upper bounds to all the complexities and that also applies to lower bounds to $K(\mathcal{F}_3(S, d))$.

2. Our results

Our aim is to study bounds on the value of the complexity of several families of sets. For example, in the case of the family $\mathcal{F}_3(S_2, d)$ with $d \geq 2$, it is proved in [4] that this bound is close to the correct order of magnitude. More precisely if $\beta = \Delta/p$, Theorem 1 of [4] implies

$$K(\mathcal{F}_3(S_2, d)) \gg_{\beta, d} \frac{\log p}{\log \log p}.$$

When the target S is formed by consecutive integers like S_1 the situation is completely different. This is our first result, we give an upper bound on $K(\mathcal{F}_1(S, d))$ independent of p and Δ.

Theorem 2.1. *Suppose $d \geq 5$. Then the inequality $K(\mathcal{F}_1(S_1, d)) \leq (d+1)(d+2) \log_2 K(\mathcal{F}_1(S_1, d))$ holds.*

Remark 2.2. In fact our proof gives more precise bound than the one announced in Theorem 2.1 and this bound is valid for all degree $d \geq 1$ (see (6.2)). On the other hand, Theorem 2.1 is easier to apply. Also, it is possible to get tighter upper bounds using a computer. For example, for linear polynomials, the bound for $K(\mathcal{F}_1(S, 1))$ is 15, almost half. It has also been calculated for $K(\mathcal{F}_1(S, 2))$ giving 34 after two days in a cluster.

To prove the previous result, we need to devote Section 3 to introduce the concept of hyperplanes arrangement.

Theorem 1 of [4] gives an upper bound for $K(\mathcal{F}_3(S_2,d))$ only for degree $d \geq 2$ but in fact its proof provides also a result for $d=1$. Our next result gives another upper bound for $K(\mathcal{F}_3(S_2,1))$ which seems to be better in terms of Δ (but not in r) than the one we could obtained in [4]. This bound is based on geometry of numbers, thus proposes another approach to this problem.

Theorem 2.3. *Suppose that r and Δ satisfy*

$$\frac{r^{(3r-1)/2}\,p^{r-1}}{2^r\left(\frac{1}{2}\left(\frac{2p}{r-1}\right)^{1/r}-1\right)^{(r-1)^2}} \leq \Delta \leq \frac{p}{2}. \tag{2.1}$$

Then the inequality $K(\mathcal{F}_3(S_2,1)) \geq r$ holds.

When $p \to +\infty$, equation (2.1) imposes that $\Delta > (1+o(1))C_r p^{1-1/r}$ for some $C_r > 0$. With more computations, we could slightly improve (2.1) for example by using the remark after Lemma 7.1 but we preferred to avoid these complications.

In Theorem 2.1, we have proved that when S is the set of the Δ first numbers the complexity $K(\mathcal{F}_1(S,d))$ is bounded independently on p. In the other direction it seems to be very difficult to find the exact order of magnitude of this complexity in this particular case. In [5], Theorem 4.7, the Cauchy–Davenport theorem was essential to prove that if a general $S \subset \mathbb{F}_p$ has sufficiently many elements then $K(\mathcal{F}_1(S,d)) \geq d+2$. Our last result gives a slight improvement of this result.

Theorem 2.4. *For $d \geq 6$, $p \geq 7$ and $S \subset \mathbb{F}_p$ such that $(3p+13)/7 < |S| < p/2$ we have $K(\mathcal{F}_1(S,d)) \geq d+3$.*

Also, Theorem 4 in [4] gives a similar lower bound for the complexity $K(\mathcal{F}_2(S,d))$.

The proof of this last theorem uses some important results of additive number theory: the Cauchy–Davenport theorem and the combinatorial Nullstellensatz of Alon. We remark that the techniques of our proofs could lead to other improvements.

In the next three sections we recall some notions and results on hyperplane arrangements, lattices and sumsets estimates that we will need in our proofs.

3. Hyperplanes arrangement

Hyperplane arrangements are objects well studied in the field of combinatorial geometry, see [9]. We only introduce enough theory to understand the proof of Theorem 2.1, following the nice introduction given in [20].

Let r be a positive integer and $\mathbb{R}$ the field of real numbers. We denote by

$$\vec{a} = (a_1,\dots,a_r), \quad a_1,\dots,a_r \in \mathbb{R}$$

elements of $\mathbb{R}^r$, where $\mathbb{R}^r$ is a vector space of dimension r over the field $\mathbb{R}$. We also consider matrices with the usual operations involving matrices, namely multiplication, addition and transposition. Also, we need the topological concept of dimension of a set of elements in $\mathbb{R}^r$. Vectors in $\mathbb{R}^r$ are matrices with r rows and 1 column. The notation for the transposition of a matrix A is $A^\intercal$.

Definition 3.1. Given $\vec{a} \in \mathbb{R} - \{\vec{0}\}$ and $b \in \mathbb{R}$, the set $\{\vec{x} \in \mathbb{R}^r : \vec{a}^\intercal \vec{x} = b\}$ is called a hyperplane.

We also use $\vec{a} \cdot \vec{x}$ to denote $\vec{a}^\intercal \vec{x}$, which corresponds to the standard dot product, and the matrix form $A\vec{x} = \vec{b}$ to encode the finite set of hyperplanes $\mathcal{H} = \{\mathcal{H}_1, \ldots, \mathcal{H}_m\}$, where

$$\mathcal{H}_i = \Big\{\vec{x} \in \mathbb{R}^r : \sum_{j=1}^{r} a_{i,j}\, x_j = b_i\Big\}. \tag{3.1}$$

Definition 3.2. A set of hyperplanes in $\mathbb{R}^r$ partitions the space into relatively open convex polyhedral regions, called faces, of all dimensions. This partition is called a hyperplane arrangement.

We make a distinction between the two sides of a hyperplane. An element $\vec{p}$ of $\mathbb{R}^r$ is on the positive side of hyperplane $\mathcal{H}_i$, denoted by $\mathcal{H}_i^+$, if

$$\sum_{j=1}^{r} a_{i,j}\, p_j > b_i.$$

Similarly, we define $\vec{p} \in \mathbb{R}^r$ is on the negative side of hyperplane $\mathcal{H}_i$ and we denote it by $\mathcal{H}_i^-$.

For each $\vec{p} \in \mathbb{R}^r$ we define a sign vector of length m consisting of $1, 0, -1$ signs as follows:

$$sv(\vec{p})_i = \begin{cases} 1 & \text{if } \vec{p} \in \mathcal{H}_i^+, \\ -1 & \text{if } \vec{p} \in \mathcal{H}_i^-, \\ 0 & \text{if } \vec{p} \in \mathcal{H}_i, \end{cases}$$

where $i = 1, \ldots, m$, and m is the number of hyperplanes.

Definition 3.3. A face is a set of elements of $\mathbb{R}^r$ with the same sign vector. It is called a i-face if its dimension is $i \leq r$ and a cell if the dimension is r.

As a small comment, the dimension of a face is at least r minus the number of zeros in the sign vector of any of the points of the face. The number of faces of given dimension in a hyperplane arrangement is given in the following result.

Lemma 3.4 (Theorem 1.3 in [9]). *Given any set of hyperplanes $\mathcal{H} = \{\mathcal{H}_1, \ldots, \mathcal{H}_m\}$ in $\mathbb{R}^r$, then the number of i-faces in the correspondent hyperplane arrangement can be bounded by*

$$\sum_{j=0}^{i} \binom{r-j}{i-j} \binom{m}{r-j}.$$

4. Lattices

Lattices are just sets of vectors generated by linear combinations of a basis with integer coefficients. The fact that lattice are discrete sets introduce several new problems. One of them is the 'closest vector problem', which has important relations with cryptography. Here we review some results and definitions concerning the closest vector problem, all of which can be found in [10]. We also recommend the interested reader consulting [11], [17], [18].

Let $\{\vec{b}_1, \dots, \vec{b}_s\}$ be a set of linearly independent vectors in $\mathbb{R}^r$. The set

$$\mathcal{L} = \{c_1\vec{b}_1 + \cdots + c_s\vec{b}_s \mid c_1, \dots, c_s \in \mathbb{Z}\}$$

is an *s-dimensional lattice* with *basis* $\{\vec{b}_1, \dots, \vec{b}_s\}$. If $s = r$, the lattice $\mathcal{L}$ is of *full rank*. Associated to a lattice $\mathcal{L}$, it is possible to associate its volume, which can be calculated using any basis defining the lattice. For full rank lattices $\operatorname{vol}\mathcal{L} = |\det B|$, where B is the matrix which rows form a basis of $\mathcal{L}$.

One basic lattice problem is the closest vector problem (CVP): given a basis of a lattice $\mathcal{L}$ in $\mathbb{R}^r$ and a shift vector $\vec{t}$ in $\mathbb{R}^r$, the goal is finding a vector in the lattice $\mathcal{L}$ closest to the target vector $\vec{t}$ with respect to the Euclidean norm. It is well known that this problem is NP-hard when the dimension grows. However, it is solvable in deterministic polynomial time provided that the dimension of $\mathcal{L}$ is fixed (see [12], [15], for example).

For a slightly weaker task of finding a sufficiently close vector, the celebrated *LLL algorithm* of Lenstra, Lenstra and Lovász [13] provides a desirable solution, as noticed by [3]. Here, we state a weaker consequence as Lemma 4.1.

Lemma 4.1. *There exists a deterministic polynomial time algorithm which, when given an r-dimensional full rank lattice $\mathcal{L}$ and a shift vector $\vec{t}$, finds a lattice vector $\vec{u} \in \mathcal{L}$ satisfying the inequality*

$$4\,\|\vec{t} - \vec{u}\|^2 \le \|\vec{b}_1\|^2 + \cdots + \|\vec{b}_r\|^2,$$

where $\|\cdot\|$ denotes the Euclidean norm.

Thanks to Minkowski's second theorem and the upper bound for the Hermite constant, we have the following result. For information about these results and further references, see [16].

Lemma 4.2. *Let $\mathcal{L}$ be a r dimensional rank lattice, then there exists a basis $\{\vec{b}_1, \dots, \vec{b}_r\}$ of this lattice such that*

$$2^r\Big(\prod_{i=1}^{r} \|\vec{b}_i\|\Big) \le r^{r/2}(\operatorname{vol}\mathcal{L}).$$

Lemmas 4.1 and 4.2 give immediately a bound for the norm of an approximation of the CVP. Many other results on both exact and approximate finding of a closest vector in a lattice are discussed in [10], [11], [17].

To finish this section, we cite Lemma 1 in [11] because it will be necessary in our theorem.

Lemma 4.3. *Assume that $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^r$ are vectors with integer coefficients bounded in absolute value by M. Suppose that $n > r$, then there exists an integer relation,*

$$\sum_{i=1}^{n} \alpha_i \vec{v}_i = 0,$$

such that $|\alpha_i| \leq B$ where B is given by

$$\log_2 B = r \frac{\log_2 M + \log_2 n + 1}{n - r}.$$

Remark 4.4. We will apply this lemma with $r = 1$, that is when $\vec{v}_1, \ldots, \vec{v}_n$ are real numbers $v_1, \ldots, v_n$. In this case, it is easy to obtain a slight improvement for B. The idea of the proof of Lemma 4.3 is to consider all linear combinations $\sum_{i=1}^{r} \lambda_i \vec{v}_i$ with $0 \leq \lambda_i \leq B-1$. We obtain B^n vectors with coordinates of modulus $\leq n(B-1)M$. When $B^n > (2nBM+1)^r$ at least two vectors must be equal.

When $r = 1$ we can choose the coefficient λ_i such that $0 \leq |\lambda_i| \leq B-1$ and $\lambda_i v_i \geq 0$. Then $\sum_{i=1}^{n} \lambda_i v_i \in [-n(B-1)M, n(B-1)M]$.

Thus when $r = 1$, we can take

$$B = (nM)^{1/(n-1)}. \tag{4.1}$$

5. The Cauchy–Davenport inequality and the combinatorial Nullstellensatz

We recall in this section two theorems in additive number theory that we will use in the proof of Theorem 2.4. For any subsets $\mathcal{A}, \mathcal{B}$ of $\mathbb{F}_p$ and $r, s \in \mathbb{F}_p$, we will use the following usual notations in additive number theory:

$$\mathcal{A} + \mathcal{B} = \{a + b : a \in \mathcal{A},\ b \in \mathcal{B}\},$$
$$r + s\mathcal{A} = \{r + sa : a \in \mathcal{A}\}.$$

We begin with the Cauchy–Davenport theorem (see [21], Theorem 5.4, p. 223, and references therein).

Lemma 5.1 (Cauchy–Davenport inequality). *Let $\mathcal{A} \subset \mathbb{F}_p$ and $\mathcal{B} \subset \mathbb{F}_p$. We have*

$$|\mathcal{A} + \mathcal{B}| \geq \min(p, |\mathcal{A}| + |\mathcal{B}| - 1).$$

This theorem was already applied in [5] to prove that if a subset $S \subset \mathbb{F}_p$ is so that $|S|$ and $|\mathbb{F}_p \setminus S|$ are large enough then $K(\mathcal{F}_1(S, d)) \geq d + 2$ (see Theorem 4.7 in [5]). The other important ingredient of the proof of Theorem 2.4, is the combinatorial Nullstellensatz of Alon:

Lemma 5.2 (Combinatorial Nullstellensatz, Theorem 1.2 in [2], or Theorem 9.2 in [21]). *Let F be an arbitrary field, $P \in F[x_1, \ldots, x_n]$ be a polynomial of degree d which contains a non zero coefficient at $x_1^{d_1} \cdots x_n^{d_n}$ with $d_1 + \cdots + d_n = d$. Let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be subsets of F such that $|\mathcal{A}_i| > d_i$ for all $1 \leq i \leq n$. Then there exists $a_1 \in \mathcal{A}_1, \ldots, a_n \in \mathcal{A}_n$ such that $P(a_1, \ldots, a_n) \neq 0$.*

6. Proof of Theorem 2.1

In this section, we prove Theorem 2.1 using the concept of hyperplane arrangements. We introduce the following notation: for $\mathcal{A}, \mathcal{B}$ two sets, we denote by $\mathcal{A}-\mathcal{B}$ the set with the elements that belongs to $\mathcal{A}$ and do not belong to $\mathcal{B}$.

We want to prove that the complexity measure $K(\mathcal{F}_i(S,d)),\ i=1,2,3$ is upper bounded, so we are going to show that $\mathcal{A}=\{1,\dots,k\}$ is possible to be partitioned in two different sets $\mathcal{B}$ and $\mathcal{C}$ such that there doesn't exist $\mathcal{R}\in\mathcal{F}_1(S_1,d)$ with

$$\mathcal{B}\subset\mathcal{R} \quad \text{and} \quad \mathcal{C}\subset\{0,\dots,p-1\}-\mathcal{R}, \tag{6.1}$$

when k is big enough.

First, we notice that for each partition $\mathcal{B}$ and $\mathcal{C}$ satisfying $\mathcal{A}=\mathcal{B}\cup\mathcal{C}$, $\mathcal{B}\cap\mathcal{C}=\emptyset$ and Equation (6.1), there exists a polynomial f of degree d,

$$f(x)=f_0+f_1x+\cdots+f_dx^d,$$

such that

- $f(i) \bmod p\in S_1$ for $i\in\mathcal{B}$,
- $f(i) \bmod p\in\{0,\dots,p-1\}-S_1$ for $i\in\mathcal{C}$.

Since for any $i\in\{1,\dots,k\}$, $f_0,\dots,f_d\in\{0,\dots,p-1\}$, we have

$$f(i)\in\left[0,(p-1)\frac{k^{d+1}-1}{k-1}\right]\subset[0,pk^{d+1}[\,;$$

an equivalent way to state this is to define the following hyperplanes:

$$\mathcal{H}_{i,j}=\left\{\vec{x}\in\mathbb{R}^{d+1}\mid\sum_{\ell=0}^{d}x_\ell\,i^\ell=jp+\Delta+0.5\right\},\quad 1\le i\le k,\ 0\le j\le k^{d+1}-1,$$

$$\mathcal{G}_{i,j}=\left\{\vec{x}\in\mathbb{R}^{d+1}\mid\sum_{\ell=0}^{d}x_\ell\,i^\ell=jp+0.5\right\},\quad 1\le i\le k,\ 0\le j\le k^{d+1}-1,$$

and notice that $f(i) \bmod p\in S_1 \iff (f_0,\dots,f_d)\in\mathcal{H}_{i,j}^-\cap\mathcal{G}_{i,j}^+$ for some value $0\le j\le k^{d+1}-1$.

By Lemma 3.4, the number of cells of dimension $d+1$ is at most

$$\sum_{j=0}^{d+1}\binom{2k^{d+2}}{d+1-j}\le\frac{(d+2)k^{(d+1)(d+2)}2^{d+1}}{(d+1)!},$$

which is the same of the number of sign vectors without 0's. On the other hand, the number of partitions $\mathcal{A}=\mathcal{B}\cup\mathcal{C}$ with $\mathcal{B}\cap\mathcal{C}=\emptyset$ is equal to 2^k and for each partition there must be a different sign vector without 0's.

This means that,

$$2^k\le\frac{(d+2)\,k^{(d+1)(d+2)}\,2^{d+1}}{(d+1)!}. \tag{6.2}$$

When $d\ge5$, this implies that $2^k\le k^{(d+1)(d+2)}$. This finishes the proof.

7. Proof of Theorem 2.3

As a comment, the combination of Lemma 7.1 with the main theorem in [19] gives a weaker result than Theorem 2.3, but we are able to improve the constant. Before proving the theorem, we need the following technical lemma.

Lemma 7.1. *Given different $a_1, \dots, a_r \in \mathbb{F}_p$, the set*

$$\left\{ x \in \mathbb{F}_p \mid \frac{b_1}{a_1 x + 1} + \cdots + \frac{b_r}{a_r x + 1} \bmod p = 0,\ |b_1|, \dots, |b_r| \le k \text{ , not all zero} \right\}$$

has at most $(r-1)(2k+1)^r/2$ elements.

Proof. First of all, we remark that the polynomials $(a_i x + 1)$ have different roots. Without losing generality, we suppose that $(a_i x + 1) \neq 0$ for $i = 1, \dots, r$, otherwise we take the corresponding $b_i = 0$.

The set of values of x can be defined as the values,

$$\sum_{i=1}^{r} b_i \prod_{j \neq i}^{r} (a_j x + 1) = 0 \mod p. \tag{7.1}$$

When $b_1, \dots, b_r$ are fixed and not all zero, the number of solutions is at most $r-1$, by Lagrange theorem. It is easy to check that the equality in (7.1) is not identically zero just substituting $x = a_i^{-1}$ for $i = 1, \dots, r$ and notice that one of the values is different from zero. To finish the proof it remains to observe that we don't need to consider all $|b_1|, \dots, |b_r| \le k$ not all equal to zero but only the $b_1, \dots, b_k$ such that the number of index i with $b_i > 0$ is bigger that the number of i such that $b_i < 0$ and this gives $(2k+1)^r/2$ for such $(b_1, \dots, b_r)$. This remark finishes the proof. □

Remark 7.2. With more care, we can obtain a slight improvement of Lemma 7.1. We can use more precisely the fact that $b_1, \dots, b_k$ are not all equal to zero by doing the following partition:

$$\bigcup_{\ell=1}^{r} \bigcup_{1 \le i_1 < \dots < i_\ell \le r} \left\{ x \in \mathbb{F}_p : \sum_{j=1}^{\ell} \frac{b_j}{a_{i_j} + 1} \equiv 0 \mod p, 0 < |b_1|, \dots, |b_\ell| \le k \right\}.$$

Then we use the same argument as before but with polynomials of degree at most $\ell - 1$. This gives the upper bound

$$\frac{1}{2} \sum_{\ell=1}^{r} \binom{r}{\ell} (\ell - 1)(2k)^\ell = ((r-1)(2k+1)^r - r(2k+1)^{r-1} + 1)/2.$$

In order to simplify some computations, we have chosen to use the upper bound given in Lemma 7.1, but we include this remark here because it may have an independent interest.

Proof of Theorem 2.3. To prove the result about the complexity, we are going to select a set $\mathcal{A} = \{a_1, \dots, a_r\}$, reorder the elements of $\mathcal{A}$ and divide it in two different sets,

$$\mathcal{B} = \{a_1, \dots, a_k\} \quad \text{and} \quad \mathcal{C} = \{a_{k+1}, \dots, a_r\};$$

our aim is to show the existence of $\mathcal{R}$ as in (6.1). Take now an element $y \in \mathbb{F}_p$ and consider the lattice $\mathcal{L}$ of volume p^{r-1} defined by the rows of the following matrix:

$$\begin{pmatrix} (a_1y+1)^{-1} & (a_2y+1)^{-1} & \dots & (a_ry+1)^{-1} \\ p & 0 & \dots & 0 \\ 0 & p & \dots & 0 \\ & & \ddots & \\ 0 & 0 & & 0 \\ 0 & 0 & \dots & p \end{pmatrix}, \tag{7.2}$$

and the following target vector:

$$\vec{t} = (\underbrace{\Delta/2, \dots, \Delta/2}_{j}, \underbrace{3\Delta/2, \dots, 3\Delta/2}_{k-j}).$$

The matrix defined in (7.2) has r columns and $r+1$ rows, so the rows are a generator system but not a basis. The rows of the matrix

$$\begin{pmatrix} 1 & (a_1y+1)(a_2y+1)^{-1} & \dots & (a_1y+1)(a_ry+1)^{-1} \\ 0 & p & \dots & 0 \\ & & \ddots & \\ 0 & 0 & & 0 \\ 0 & 0 & \dots & p \end{pmatrix},$$

are a basis of the lattice $\mathcal{L}$, and the determinant is p^{r-1}. Notice that if we show the existence of a vector $\vec{u}$ in the lattice such that $\|\vec{u} - \vec{t}\| < \Delta/2$, then

- for some value $c \in \mathbb{F}_p$, we have

$$\vec{u} = \big(c(a_1y+1)^{-1} \bmod p, c(a_2y+1)^{-1} \bmod p, \dots, c(a_ry+1)^{-1} \bmod p\big);$$

- for $i = 1, \dots, k$, we have

$$|c(a_iy+1)^{-1} \bmod p| < \Delta,$$

and for $i = k+1, \dots, r$,

$$|c(a_iy+1)^{-1} \bmod p| > \Delta.$$

By the previous two properties, it is only necessary to take the polynomial $f = f_0 + f_1x = c^{-1} + c^{-1}yx$.

Now, we outline the idea to prove that the vector $\vec{u}$ exists. By Lemma 4.2, there is a basis $\{\vec{b}_1, \dots, \vec{b}_r\}$ satisfying

$$2^r \Big(\prod_{i=1}^{r} \|\vec{b}_i\| \Big) \le r^{r/2} (\operatorname{vol} \mathcal{L}).$$

We will show that selecting a well-chosen $y \in \mathbb{F}_p$, all the vectors in $\mathcal{L}$ have norm bigger than U with

$$U = \frac{1}{r}\Big(\frac{1}{2}\Big(\frac{2p}{r-1}\Big)^{1/r} - 1\Big)^{r-1}.$$

Applying Lemma 4.1, we have that $\|\vec{t} - \vec{u}\| \leq r^{(r+1)/2}\, p^{r-1}\, 2^{-r-1}\, U^{1-r} \leq \Delta/2$, and this is exactly the hypothesis of theorem 2.3.

So, we have to characterize the values $y \in \mathbb{F}_p$ such that $\mathcal{L}$ has a vector of norm shorter that U, so let call this vector $\vec{h} \in \mathcal{L}$.

The vector $\vec{h}$ is of the form

$$\vec{h} = \big(c'(a_1 y + 1)^{-1} \bmod p, c'(a_2 y + 1)^{-1} \bmod p, \ldots, c'(a_r y + 1)^{-1} \bmod p\big),$$

for some $c' \in \mathbb{F}_p$ and $|c'(a_i y + 1)^{-1} \bmod p| < U$ for $i = 1, \ldots, r$.

Now, we apply Lemma 4.3 and Remark 4.4 to values $c'(a_1 y + 1)^{-1} \bmod p, \ldots, c'(a_r y + 1)^{-1} \bmod p$, so there exist $\alpha_1, \ldots, \alpha_r$ satisfying

$$c'(\alpha_1 (a_1 y + 1)^{-1} + \cdots + \alpha_r (a_r y + 1)^{-1}) = 0 \bmod p \tag{7.3}$$

and

$$\max\{|\alpha_1|, \ldots, |\alpha_r|\} \leq (rU)^{1/(r-1)}.$$

Lemma 7.1 bounds the number of values $y \in \mathbb{F}_p$ such that (7.3) holds. This number is less than $(r-1)(2(rU)^{1/(r-1)} + 1)^{r-1}/2 < p$, so there exists a value y not satisfying this equation. This finishes the proof. □

8. Proof of Theorem 2.4

Let $S \subset \mathbb{F}_p$ satisfying the hypotheses of Theorem 2.4. We now use the notation $S^c = \mathbb{F}_p \setminus S$. Let $\mathcal{A} = \{a_1, \ldots, a_{d+3}\} \subset \mathbb{F}_p$ and $\{\mathcal{B}, \mathcal{C}\}$ a partition of $\mathcal{A}$. We have to find a polynomial $f(x) \in \mathbb{F}_p[x]$ of degree d such that

$$f(a_i) \in \begin{cases} S & \text{if } a_i \in \mathcal{B}, \\ S^c & \text{if } a_i \in \mathcal{C}. \end{cases}$$

To simplify some expressions, we will write for $1 \leq i \leq d+3$, $S_i = S$ if $a_i \in \mathcal{B}$, $S_i = S^c$ if $a_i \in \mathcal{C}$. If all the sets S_i are equal then it is sufficient to take f as a constant polynomial. Otherwise we may suppose that

$$S_{d+2} = S_{d+3}, \tag{8.1}$$

and that there exist $1 \leq i < j \leq d+1$ such that

$$S_i \neq S_j. \tag{8.2}$$

We will construct f with the Lagrange interpolation polynomials. For $1 \leq i \leq d+1$ we define

$$L_i(x) = \prod_{\substack{j \neq i \\ 1 \leq j \leq d+1}} \frac{x - a_j}{a_i - a_j}.$$

The polynomial $f(x)$ must be of the following shape:

$$f(x) = \sum_{i=1}^{d+1} x_i\, L_i(x),$$

with $x_i \in S_i$ for $1 \le i \le d+1$. We have to show that we can choose $x_1, \dots, x_{d+1}$ such that $f(a_{d+2}) \in S_{d+2}$ and $f(a_{d+3}) \in S_{d+3}$.

We will distinguish two situations according to the fact that the ratios $\frac{L_i(a_{d+3})}{L_i(a_{d+2})}$ are the same for many i or not.

8.1. First case

We suppose that there exist $1 \le i < j < k \le d+1$ such that we have

$$\frac{L_i(a_{d+3})}{L_i(a_{d+2})} = \frac{L_j(a_{d+3})}{L_j(a_{d+2})} = \frac{L_k(a_{d+3})}{L_k(a_{d+2})} =: \lambda. \tag{8.3}$$

We may suppose that $i = 1$, $j = 2$, $k = 3$ even if it means to change the order of the a_i. Since a_{d+2} and a_{d+3} are in $\mathbb{F}_p \setminus \{a_1, \dots, a_{d+1}\}$, $\lambda \ne 0$. The conditions $f(a_{d+2}) \in S_{d+2}$ and $f(a_{d+3}) \in S_{d+3}$ are equivalent to the system

$$\begin{cases} x_1L_1(a_{d+2}) + x_2L_2(a_{d+2}) + x_3L_3(a_{d+2}) & = w_1 - \sum_{i=4}^{d+1} x_i\, L_i(a_{d+2}), \\ \lambda(x_1L_1(a_{d+2}) + x_2L_2(a_{d+2}) + x_3L_3(a_{d+2})) & = w_2 - \sum_{i=4}^{d+1} x_i\, L_i(a_{d+3}), \end{cases} \tag{8.4}$$

with $x_i \in S_i$ for $1 \le i \le d+1$ and $w_1 \in S_{d+2}$, $w_2 \in S_{d+3}$. This implies that

$$w_2 - \lambda w_1 + \sum_{i=4}^{d+1} x_i(\lambda L_i(a_{d+2}) - L_i(a_{d+3})) = 0. \tag{8.5}$$

For $i = 4$ to $d+1$ we write for brevity $\mu_i := \lambda L_i(a_{d+2}) - L_i(a_{d+3})$.

• First we suppose that there exists $i_0 \in \{4, \dots, d+1\}$ such that $\mu_{i_0} \ne 0$. By the Cauchy–Davenport inequality (Lemma 5.1) applied two times, we have

$$\begin{aligned} \Big|S_{d+3} - \lambda S_{d+2} + \sum_{i=4}^{d+1} \mu_i S_i\Big| &\ge \min\Big(p, |S_{d+3}| + \Big| - \lambda S_{d+2} + \sum_{i=4}^{d+1} \mu_i S_i\Big| - 1\Big) \\ &\ge \min\Big(p, |S_{d+3}| + |S_{d+2}| + \Big|\sum_{i=4}^{d+1} \mu_i S_i\Big| - 2\Big) \\ &\ge \min(p, |S_{d+3}| + |S_{d+2}| + |S_{i_0}| - 2) \ge p, \end{aligned}$$

since $3|S| - 2 \ge p$ under the hypotheses of Theorem 2.4. Thus there exist $w_1 \in S_{d+2}$, $w_2 \in S_{d+3}$, $(x_4, \dots, x_{d+1}) \in \prod_{i=4}^{d+1} S_i$ such that (8.5) holds.

To solve completely (8.4) it remains to find $(x_1, x_2, x_3) \in S_1 \times S_2 \times S_3$ such that

$$\sum_{i=1}^{3} x_i\, L_i(a_{d+2}) = w_1 - \sum_{i=4}^{d+1} x_i\, L_i(a_{d+2}). \tag{8.6}$$

Since $L_i(a_{d+2}) \neq 0$, by the Cauchy–Davenport inequality we see that

$$\Big|\sum_{i=1}^{3} L_i(a_{d+2})S_i\Big| \geq 3|S| - 2 \geq p$$

under the hypotheses of Theorem 2.4. Thus there exist $(x_1, x_2, x_3) \in S_1 \times S_2 \times S_3$ satisfying (8.6).

• Now we suppose that $\mu_i = 0$ for all $4 \leq i \leq d+1$ that is $L_i(a_{d+3}) = \lambda L_i(a_{d+2})$ for all $1 \leq i \leq d+1$. This implies that $x - a_{d+3}$ divides all the polynomials $L_i(x) - \lambda L_i(a_{d+2})$. But we know that $1 = \sum_{i=1}^{d+1} L_i(x)$. This gives

$$\sum_{i=1}^{d+1}(L_i(x) - \lambda L_i(a_{d+2})) = 1 - \lambda,$$

we deduce that $\lambda = 1$. In this case, $f(a_{d+2}) = f(a_{d+3})$ for any choice of the x_i. Since $K(\mathcal{F}_1(S, d)) \geq d + 2$ by Theorem 4.7 of [5], we can find a polynomial f solving this situation.

8.2. Second case

It remains to handle the case when there are no i, j, k satisfying (8.3). Since $d \geq 6$, in this case there exist $1 \leq i < j < k \leq d+1$ such that the ratios $\frac{L_i(a_{d+3})}{L_i(a_{d+2})}$, $\frac{L_j(a_{d+3})}{L_j(a_{d+2})}$, $\frac{L_k(a_{d+3})}{L_k(a_{d+2})}$ are distinct and at least one of the sets S_i, S_j or S_k is S^c. We may then suppose that $i = 1, j = 2, k = 3$ and $S_1 = S^c$.

Let $(x_4, \ldots, x_{d+1}) \in S_4 \times \cdots \times S_{d+1}$ be fixed. We define

$$w_1 = -\sum_{i=4}^{d+1} x_i\, L_i(a_{d+2}) \text{ and } w_2 = -\sum_{i=4}^{d+1} x_i\, L_i(a_{d+3}).$$

We will prove that there exist $(x_1, x_2, x_3) \in S_1 \times S_2 \times S_3$ such that

$$x_1L_1(a_{d+2}) + x_2L_2(a_{d+2}) + x_3L_3(a_{d+2}) \in w_1 + S_{d+2},$$
$$x_1L_1(a_{d+3}) + x_2L_2(a_{d+3}) + x_3L_3(a_{d+3}) \in w_2 + S_{d+3}.$$

For this we will apply the polynomial method. Let $\mathcal{R}_i = w_i + S_{d+1+i}$ for $i = 1, 2$. We consider the polynomial

$$P(x_1, x_2, x_3) = \prod_{v \in \mathcal{R}_1^c} \Big(\sum_{i=1}^{3} L_i(a_{d+2})x_i - v\Big) \prod_{v \in \mathcal{R}_2^c} \Big(\sum_{i=1}^{3} L_i(a_{d+3})x_i - v\Big).$$

By Lemma 5.2, it is sufficient to prove that there exist k_1, k_2, k_3 such that

$$\begin{cases} k_1 + k_2 + k_3 &= |\mathcal{R}_1^c| + |\mathcal{R}_2^c|, \\ k_1 &\leq |S^c| - 1, \\ \max(k_2, k_3) &\leq |S| - 1, \end{cases} \tag{8.7}$$

and the coefficient of P in $x_1^{k_1}x_2^{k_2}x_3^{k_3}$ is non-zero. Since $k_1+k_2+k_3$ must be equal to the degree of P, no term v appears in the coefficient of $x_1^{k_1}x_2^{k_2}x_3^{k_3}$. Thus this coefficient is the same as the coefficient of $x_1^{k_1}x_2^{k_2}x_3^{k_3}$ in the polynomial

$$Q(x_1,x_2,x_3)=\Big(\sum_{i=1}^{3}L_i(a_{d+2})x_i\Big)^{|\mathcal{R}_1^c|}\Big(\sum_{i=1}^{3}L_i(a_{d+3})x_i\Big)^{|\mathcal{R}_2^c|}.$$

We use the multinomial formula:

$$Q(x_1,x_2,x_3)=\sum_{\substack{k_{11}+k_{12}+k_{13}=|\mathcal{R}_1^c|\\ k_{21}+k_{22}+k_{23}=|\mathcal{R}_2^c|}}\prod_{\substack{i=1,2\\ j=1,2,3}}\frac{|\mathcal{R}_i^c|!}{k_{ij}!}\left(x_jL_j(a_{d+1+i})\right)^{k_{ij}}.$$

Of course here and in the sequel the integers k_{ij} and k_i are ≥ 0. We collect the coefficients of the different monomials $x_1^{k_1}x_2^{k_2}x_3^{k_3}$:

$$Q(x_1,x_2,x_3)=|\mathcal{R}_1|!\,|\mathcal{R}_2|!\sum_{k_1+k_2+k_3=|\mathcal{R}_1^c|+|\mathcal{R}_2^c|}x_1^{k_1}x_2^{k_2}x_3^{k_3}\,c(k_1,k_2,k_3),$$

with

$$c(k_1,k_2,k_3)=\sum_{\substack{k_{1\ell}+k_{2\ell}=k_\ell \text{ for } \ell=1,2,3\\ k_{11}+k_{12}+k_{13}=|\mathcal{R}_1^c|}}\prod_{\substack{i=1,2\\ j=1,2,3}}\frac{L_j(a_{d+1+i})^{k_{ij}}}{k_{ij}!}. \tag{8.8}$$

We now prove that there exist k_1,k_2,k_3 satisfying (8.7) such that $c(k_1,k_2,k_3)\neq 0$. We detect the condition $k_{11}+k_{12}+k_{13}=|\mathcal{R}_1^c|$ with the circle method:

$$\int_0^1 \mathrm{e}\left(\alpha(k_{11}+k_{12}+k_{13}-|\mathcal{R}_1^c|)\right)\mathrm{d}\,\alpha=\begin{cases}1 & \text{if } k_{11}+k_{12}+k_{13}=|\mathcal{R}_1^c|,\\ 0 & \text{otherwise,}\end{cases}$$

with the notation $\mathrm{e}(t)=\exp(2i\pi t)$. Next we insert this formula in (8.8) and use the multinomial formula:

$$c(k_1,k_2,k_3)=\frac{1}{k_1!k_2!k_3!}\int_0^1\prod_{i=1}^{3}\left(L_i(a_{d+2})\,\mathrm{e}(\alpha)+L_i(a_{d+3})\right)^{k_i}\mathrm{e}\left(-\alpha|\mathcal{R}_1^c|\right)\mathrm{d}\,\alpha.$$

We replace in the integral $\mathrm{e}(\alpha)$ by z:

$$c(k_1,k_2,k_3)=\frac{1}{k_1!\,k_2!\,k_3!}\,I(|\mathcal{R}_1^c|,k_1,k_2,k_3)$$

with

$$I(n,k_1,k_2,k_3)=\frac{1}{2i\pi}\int_{|z|=1}\prod_{i=1}^{3}\left(L_i(a_{d+2})z+L_i(a_{d+3})\right)^{k_i}\frac{\mathrm{d}\,z}{z^{n+1}}.$$

It is probably possible to obtain, *via* the saddle point method, an asymptotic formula when n,k_i are going to infinity and the k_i are well chosen. Here we only

need to find appropriate k_i such that $I(n,k_1,k_2,k_3)\neq 0$. It can be achieved with a simple recurrence argument on the n,k_1,k_2,k_3.

If $k_1+k_2+k_3=n$, then by the Cauchy formula we immediately see that

$$I(n,k_1,k_2,k_3)=\prod_{i=1}^{3}L_i(a_{d+2})^{k_i}\neq 0. \tag{8.9}$$

Next we prove the following.

Lemma 8.1. *Let n, k_1, k_2 and k_3 be some positive integers such that $I(n,k_1,k_2,k_3)\neq 0$. Then at least two terms among $I(n,k_1+1,k_2,k_3)$, $I(n,k_1,k_2+1,k_3)$, and $I(n,k_1,k_2,k_3+1)$ are not equal to 0.*

Proof. By a direct computation we observe that

$$L_2(a_{d+2})I(n,k_1,k_2,k_3+1)-L_3(a_{d+2})I(n,k_1,k_2+1,k_3)=D_{23}I(n,k_1,k_2,k_3), \tag{8.10}$$

with

$$D_{23}=L_2(a_{d+2})L_3(a_{d+3})-L_3(a_{d+2})L_2(a_{d+3}).$$

We obtain two analogous formulae by replacing in (8.10) L_2, L_3 by L_i, L_j for some $i\neq j$ in $\{1,2,3\}$, the determinant D_{23} being then replaced by

$$D_{ij}=L_i(a_{d+2})L_j(a_{d+3})-L_j(a_{d+2})L_i(a_{d+3}).$$

Since the ratio $\frac{L_i(a_{d+3})}{L_i(a_{d+2})}$ are distinct, the determinants D_{ij} are non-zero and the lemma follows. □

We will handle only the case $S_{d+2}=S_{d+3}=S$ which is the most difficult since $|S|\leq|S^c|$. In this case, we have $|\mathcal{R}_1^c|=|\mathcal{R}_2^c|=|S^c|$. By adding at most 3 elements in $\mathcal{R}_1^c$ and at most 1 element in $\mathcal{R}_2$ we may suppose that 4 divides $|\mathcal{R}_1^c|$ and 2 divides $|\mathcal{R}_2^c|$. In any case we have then with these eventually modified sets $|\mathcal{R}_1^c|\leq|S^c|+3$ and $|\mathcal{R}_2^c|\leq|S^c|+1$. We start out with $k_1=|\mathcal{R}_1^c|/2-4$, $k_2=k_3=(|\mathcal{R}_1^c|-k_1)/2$. By (8.9), $I(|\mathcal{R}_1^c|,k_1,k_2,k_3)\neq 0$.

Next we apply Lemma 8.1 in the following way. We suppose that in the m-th step, we have found m_1, m_2, m_3 such that $m_1+m_2+m_3=m$ and $I(|\mathcal{R}_1^c|,k_1+m_1,k_2+m_2,k_3+m_3)\neq 0$.

We select $i_0\in\{1,2,3\}$ such that $m_{i_0}=\max(m_1,m_2,m_3)$. By Lemma 8.1 there exist $(\delta_1,\delta_2,\delta_3)\in\{(1,0,0),(0,1,0),(0,0,1)\}$ such that $\delta_{i_0}=0$ and $I(|\mathcal{R}_1^c|,m_1+\delta_1,m_2+\delta_2,m_3+\delta_3)\neq 0$.

After $|\mathcal{R}_2^c|$ such iterations, we find 3 positive integers n_1,n_2,n_3 such that $n_1+n_2+n_3=|\mathcal{R}_2^c|$, $\max(n_1,n_2,n_3)\leq|\mathcal{R}_2^c|/2$ and $I(|\mathcal{R}_1^c|,k_1+n_1,k_2+n_2,k_3+n_3)\neq 0$.

Our choice of k_1,k_2,k_2 yields to the upper bounds:

$$k_1+n_1\leq\frac{|\mathcal{R}_1^c|+|\mathcal{R}_2^c|}{2}-4\leq|S^c|-2,$$

$$\max(k_2+n_2,k_3+n_3)\leq\frac{|\mathcal{R}_1^c|}{4}+\frac{|\mathcal{R}_2^c|}{2}+2\leq\frac{3|S^c|+13}{4}<|S|,$$

under the conditions of Theorem 2.4. We may apply Lemma 5.2. This ends the proof of Theorem 2.4.

Remark 8.2. When there exist two indexes $i < j$ such that $S_i = S_j = S^c$, it is possible to obtain a better condition for the size on S. In this case we can change the order and suppose that $S_{d+2} = S_{d+3} = S^c$. Then in this last part we can start with $k_1 = k_2 = k_3 = |S|/3$ and apply Lemma 8.1 $|S|$ times. The corresponding k_i, n_i would satisfy $k_i + n_i \leq |S|/3 + |S|/2 = 5|S|/6$ which is sufficient.

The situation when $S_i = S^c$ for exactly one index i is the most difficult to handle.

Acknowledgments. The authors are very grateful to the referee for the comments which greatly improved the paper. We also want to thank the organizers of the workshop "Pseudorandomness in number theory (CIRM, July 2014)", where this collaboration started.

References

[1] Ahlswede, R., Khachatrian, L. H., Mauduit, C. and Sárközy, A.: A complexity measure for families of binary sequences. *Period. Math. Hungar.* **46** (2003), no. 2, 107–118.

[2] Alon, N.: Combinatorial Nullstellensatz. *Combin. Probab. Comput.* **8** (1999), no. 1-2, 7–29.

[3] Babai, L.: On Lovász' lattice reduction and the nearest lattice point problem. *Combinatorica* **6** (1986), no. 1, 1–13.

[4] Balasubramanian, R., Dartyge, C. and Mosaki, É.: Sur la complexité de familles d'ensembles pseudo-aléatoires. *Ann. Inst. Fourier (Grenoble)* **64** (2014), no. 1, 267–296.

[5] Dartyge, C., Mosaki, É. and Sárközy, A.: On large families of subsets of the set of the integers not exceeding N. *Ramanujan J.* **18** (2009), no. 2, 209–229.

[6] Dartyge, C. and Sárközy, A.: Large families of pseudorandom subsets formed by power residues. *Unif. Distrib. Theory* **2** (2007), no. 2, 73–88.

[7] Dartyge, C. and Sárközy, A.: On pseudo-random subsets of the set of the integers not exceeding N. *Period. Math. Hungar* **54** (2007), no 2, 183–200.

[8] Dartyge, C., Sárközy, A. and Szalay, M.: On the pseudo-randomness of subsets related to primitive roots. *Combinatorica* **30** (2010), no. 2, 139–162.

[9] Edelsbrunner, H.: *Algorithms in combinatorial geometry.* EATCS Monographs on Theoretical Computer Science 10, Springer-Verlag, Berlin, 1987.

[10] Grötschel, M., Lovász, L. and Schrijver, A.: *Geometric algorithms and combinatorial optimization.* Algorithms and Combinatorics: Study and Research Texts 2, Springer-Verlag, Berlin, 1988.

[11] Joux, A. and Stern, J.: Lattice reduction: a toolbox for the cryptanalyst. *J. Cryptology* **11** (1998), no. 3, 161–185.

[12] Kannan, R.: Minkowski's convex body theorem and integer programming. *Math. Oper. Res.* **12** (1987), no 3, 415–440.

[13] Lenstra, A. K., Lenstra, H. W. and Lovász, L.: Factoring polynomials with rational coefficients. *Math. Ann.* **261** (1982), no. 4, 515–534.

[14] Mauduit, Ch. and Sárközy, A.: On finite pseudorandom binary sequences. I. Measure of pseudorandomness, the Legendre symbol. *Acta Arith.* **82** (1997), no. 4, 365–377.

[15] Micciancio, D. and Voulgaris, P.: A deterministic single exponential time algorithm for most lattice problems based on voronoi cell computations. *SIAM J. Comput.* **42** (2013), no. 3, 1364–1391.

[16] Nguyen, P. Q.: Hermite's constant and lattice algorithms. In *The LLL algorithm. Survey and applications. Information security and Cryptography*, 19–69. Springer, 2010.

[17] Nguyen, P. Q. and Stern, J.: Lattice reduction in cryptology: an update. In *Algorithmic number theory (Leiden, 2000)*, 85–112. Lecture Notes in Comput. Sci. 1838, Springer, Berlin, 2000.

[18] Nguyen, P. Q. and Stern, J.: The two faces of lattices in cryptology. In *Cryptography and lattices (Providence, RI, 2001)*, 146–180. Lecture Notes in Comput. Sci. 2146, Springer, Berlin, 2001.

[19] Niederreiter, H. and Sloan, I. H.: Lattice rules for multiple integration and discrepancy. *Math. Comp.* **54** (1990), no. 189, 303–312.

[20] Sleumer, N. H.: *Hyperplane arrangements: Construction, visualization and applications.* Master's thesis, Swiss Federal Institute of Technology, 2000.

[21] Tao, T. and Vu, V. H.: *Additive combinatorics.* Cambridge Studies in Advanced Mathematics 105, Cambridge University Press, Cambridge, 2006.

Received September 8, 2015.

Cécile Dartyge: Institut Elie Cartan, UMR 7502, Université de Lorraine, BP 239, 54506 Vandœuvre-lès-Nancy Cedex, France.
E-mail: cecile.dartyge@univ-lorraine.fr

Domingo Gómez-Pérez: Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Avda. Los Castros s/n, 39400 Santander, Spain.
E-mail: domingo.gomez@unican.es

During the preparation of this paper the second author was partially supported by project MTM2014-55421-P from the Ministerio de Economía y Competitividad. This work was also supported by the ANR-FWF bilateral project MuDeRa "Multiplicativity: Determinism and Randomness" (France-Austria).