Set Systems with Restricted Intersections modulo Prime Powers

László Babai[1]

Department of Computer Science, Department of Mathematics, University of Chicago, 1100 E. 58th Street, Chicago, Illinois 60637-1588
E-mail: laci@cs.uchicago.edu

Péter Frankl

Rényi Mathematical Institute, Budapest, Hungary H-1053, on leave from C.N.R.S., Paris, France

Samuel Kutin[2]

Department of Mathematics, University of Chicago, 1100 E. 58th Street, Chicago, Illinois 60637-1588
E-mail: kutin@cs.uchicago.edu

and

Daniel Štefankovič

Department of Computer Science, University of Chicago, 1100 E. 58th Street, Chicago, Illinois 60637-1588; and
Department of Computer Science, Comenius University, Bratislava, Slovakia
E-mail: stefanko@cs.uchicago.edu

We study set systems satisfying Frankl–Wilson-type conditions modulo prime powers. We prove that the size of such set systems is polynomially bounded, in contrast with V. Grolmusz's recent result that for non-prime-power moduli, no polynomial bound exists. More precisely we prove the following result.

Theorem. *Let p be a prime and $q=p^k$. Let $\mu_1, ..., \mu_s$ be distinct integers, $0 \leqslant \mu_i \leqslant q-1$. Let X be a set of n elements and let $A_1, A_2, ..., A_m$ be subsets of X with the following properties*:

- $|A_i| \not\equiv \mu_\ell \pmod q$ *for all* i, ℓ, $1 \leqslant i \leqslant m$, $1 \leqslant \ell \leqslant s$.
- *For all* i, j $(1 \leqslant i < j \leqslant m)$, *there exists* ℓ $(1 \leqslant \ell \leqslant s)$ *such that*

$$|A_i \cap A_j| \equiv \mu_\ell \qquad (\mathrm{mod}\ q).$$

Then

$$m \leqslant \binom{n}{D} + \binom{n}{D-1} + \cdots + \binom{n}{0},$$

where $D \leqslant 2^{s-1}$.

D is the minimum degree of polynomials "separating" the set $\{\mu_1, ..., \mu_s\}$ from the sizes of the A_i modulo q ($q = p^k$). Let $D(s, k)$ be the maximum value of D for fixed s and k (L and p vary). The asymptotic behavior of the function $D(s, k)$ turns out to be rather complicated; we establish polynomially related upper and lower bounds on $D(s, k)$. These bounds give us an idea of the limitations of the method. Our results extend a theorem of Deza, Frankl, and Singhi, who studied the case of prime moduli. The main point we make is that our bound implies a *polynomial bound* of the form $m \leqslant n^{c(s)}$ for all prime power moduli q. In this sense the theorem complements a remarkable recent result of Grolmusz that no bound of this form holds for any q which is not a prime power.

1. INTRODUCTION

Let q be a positive integer. Let $\mathbb{N}_q$ denote the set of integers $\{0, 1, ..., q-1\}$. Let $L \subset \mathbb{N}_q$. For an integer r we say that "$r \in L \pmod q$" if $r \equiv \mu \pmod q$ for some $\mu \in L$. We say that "$r \notin L \pmod q$" if $r \not\equiv \mu \pmod q$ for any $\mu \in L$.

Let X denote a set of n elements; we refer to X as the "universe." A set-system over X is a set $\mathscr{F}$ of subsets of X.

Definition 1.1. The set-system $\mathscr{F}$ is *L-avoiding* mod q if

- $|E| \notin L \pmod q$ for all $E \in \mathscr{F}$.

The set-system $\mathscr{F}$ is *L-intersecting* mod q if

- $|E \cap F| \in L \pmod q$ for all $E, F \in \mathscr{F}$, $E \neq F$.

We set $s = |L|$. We are interested in the maximum cardinality $m(n, s, q)$ of $\mathscr{F}$ in terms of n, s, and q. Here is the precise definition:

Definition 1.2. Let $m(n, s, q)$ denote the smallest integer such that the following holds:

For any set $L \subset \mathbb{N}_q$ of size $|L| = s$, if $\mathscr{F}$ is a set-system over a universe of n elements and $\mathscr{F}$ is L-avoiding mod q and L-intersecting mod q then $|\mathscr{F}| \leqslant m(n, s, q)$.

Notation. For integers n, s, we let $\binom{n}{\leqslant s}$ denote the sum $\sum_{i=0}^{s} \binom{n}{i}$.

Deza, Frankl, and Singhi [4] proved that for $q = p$ a prime,

$$m(n, s, p) \leqslant \binom{n}{\leqslant s}, \tag{1}$$

under the additional hypothesis that, for all $E \in \mathscr{F}$, $|E| \equiv \mu_0$ (mod p for some $\mu_0 \in \mathbb{N}_q \setminus L$. (This hypothesis was removed by Alon *et al.* in [1].)

This result followed the paper by Frankl and Wilson [5] which established the bound $|\mathscr{F}| \leqslant \binom{n}{s}$ for set systems that are L-intersecting mod p and μ_0-uniform, i.e., $|E| = \mu_0$ for all $E \in \mathscr{F}$.

Note that the right-hand side of inequality (1) is less than n^s.

Inequality (1) does not extend to non-prime moduli. It was shown in [6] (cf. [3, p. 117, Example 5.9.3–5.9.5]) that no bound of the form $m(n, s, q) \leqslant C_q n^s$ holds for $q = 6$ and for $q = p^2$ where $p \geqslant 7$ is a prime. In particular,

$$m(n, 3, 6) > cn^4 = cn^{s+1} \qquad (s = 3),$$

and

$$m(n, s, q) > c_q n^{s+p-5} \approx c_q n^{s+\sqrt{2s}},$$

where $q = p^2$, p is a prime, and $s = 1 + (q - p)/2$.

While this shows that in these cases, $m(n, s, q)\, n^{-s} \to \infty$ as $n \to \infty$ (q, s fixed), the rate of growth is still *polynomial*, i. e., of the form

$$O(n^{c(s)}) \tag{2}$$

for some function $c(s)$. (The value $c(s)$ does not depend on q or n.) In fact, in these examples, $c(s) < 2s$.

Recently Grolmusz [7] proved that much larger set systems satisfying the conditions exist if the modulus is an integer which is *not a prime power*: in this case,

$$m(n, q-1, q) \geqslant \exp\left(C(q) \frac{(\log n)^r}{(\log \log n)^{r-1}} \right), \tag{3}$$

where r is the number of distinct prime divisors of q and $C(q)$ is a function[3] of q. This growth rate is *superpolynomial* (as a function of n), so, for non-prime-power values of q, no bound of the form (2) can exist.

In this note we address the question of the order of magnitude of $m(n, s, q)$ for $q = p^k$ a prime power. Our main result establishes that in this case $m(n, s, q)$ is polynomially bounded, i. e., we find a bound of the form (2).

Theorem 1.1. *For $q = p^k$ a prime power, we have*

$$m(n, s, q) \leqslant \binom{n}{\leqslant 2^{s-1}}. \tag{4}$$

[3] Specifically, $C(q) = 1/(2^r p^r r^{r-1})$, where p is the largest prime divisor of q.

The proof of this theorem is given in Section 5. The later sections of the paper further refine this statement, but are not necessary to the proof of Theorem 1.1.

Combined with Grolmusz's theorem, this result settles the question of what moduli q make the expression $m(n, s, q)$ polynomially bounded as a function of n.

Corollary 1.1. *Let R be a set of integers $\geqslant 2$. Then the following are equivalent*:

(a) *There exists a function $c(s)$ such that for all $q \in R$ and all $n \geqslant 2$,*

$$m(n, s, q) \leqslant n^{c(s)}. \tag{5}$$

(b) *All members of R are prime powers.*

Indeed, if all members of R are prime powers, then, by Theorem 1.1, the bound (5) holds with $c(s) = 2^{s-1}$. On the other hand, if some $q \in R$ is not a prime power then, setting $s = q - 1$, Grolmusz's result shows that no exponent $c(s)$ can be valid.

For many values of $q = p^k$ and s, we obtain a stronger result. In Section 2.1, we define a "separating polynomial," as well as the quantity $D(s, k)$ (which is the minimum degree of such a polynomial over all choices of p and L).

Theorem 1.2. *For $q = p^k$ a prime power, we have*

$$m(n, s, q) \leqslant \binom{n}{\leqslant D(s, k)}. \tag{6}$$

The simple argument in Section 5 shows that

$$D(s, k) \leqslant \min \left\{ \left\lfloor \left(1 + \frac{s-1}{k}\right)^k \right\rfloor, 2^{s-1} \right\}. \tag{7}$$

Note in particular that $m(n, s, p^k) \leqslant n^{D(s,k)}$, and $D(s, k) \leqslant \min\{s^k, 2^{s-1}\}$. Furthermore, when $k = 1$, $D(s, 1) = s$, so our result includes the Deza–Frankl–Singhi inequality (1).

For small k, statement (7) is our best bound on $D(s, k)$. However, much better bounds hold when $\sqrt{s} < k < e^{s/2}$. By examining an optimization problem which may be of independent interest, we determine tight log-asymptotic bounds on $D(s, k)$ in terms of s and k. For example, when $k = \sqrt{s}$, (7) gives $D(s, k) = \exp(O(k \ln k))$, when in fact we prove $D(s, k) = \exp(\Theta(k))$. The log-asymptotic analysis is rather involved, and can be found in Sections 6 and 7.

This analysis gives us an idea of how far the Frankl–Wilson approach can be taken. There is a large gap between our upper bounds and the best known constructions of set systems (cf. Section 10).

Remark 1.1. In fact, our proof technique will show that $|\mathscr{F}| \leqslant \binom{n}{\leqslant D(s,k)}$ for set systems $\mathscr{F}$ satisfying a slightly more general property: for each $E \in \mathscr{F}$, there is a subset $L_E \subset \mathbb{N}_q$, $|L_E| \leqslant s$, so that $|E| \notin L_E \pmod{q}$, but $|E \cap F| \in L_E \pmod{q}$ for all $F \neq E$. The corresponding generalization of the Deza–Frankl–Singhi inequality can be found in [3, Example 5.10.1].

In Section 2, we define separating polynomials. In Section 3, we give a proof of Theorem 1.2 assuming the existence of these polynomials, which we then construct in Sections 4 and 5. In Sections 6 and 7 we obtain tight log-asymptotic bounds for the degrees of separating polynomials in terms of s and k; we also prove that 2^{s-1} is tight if we want a bound independent of k. In Section 8, we use higher incidence matrices, following [9] and [5], to obtain a stronger version of Theorem 1.2. In Section 9, we examine the special case $L = \{0, 1, ..., s-1\}$; in this case, we can improve the upper bound to $m \leqslant \binom{n}{\leqslant 2s}$. Finally, in Section 10, we list open questions.

2. DEFINITIONS

Throughout the rest of this paper we make the following assumptions:

Standard Assumptions.

- *p is a prime number, and $q = p^k$, $L \subset \mathbb{N}_q$, and $|L| = s$.*
- *α is an integer chosen so that $\alpha \notin L \pmod{q}$.*
- *$\mathscr{F}$ is a set-system over a universe of n elements.*
- *$\mathscr{F}$ is L-avoiding mod q and L-intersecting mod q.*

2.1. *Separating Polynomials*

Before we definine a separating polynomial, it will be helpful to introduce some p-adic terminology and notation.

We define the (p-adic) valuation $\mathrm{val}(t)$ of an integer t to be the exponent j such that p^j divides t, but p^{j+1} does not. Equivalently, $\mathrm{val}(t) = -\log_p |t|_p$, where $|t|_p$ represents the usual p-adic norm on $\mathbb{Z}$. We write $\mathrm{val}(0) = \infty$.

We list some useful properties of the valuation:

- $\mathrm{val}(t) \leqslant \infty$ and $\mathrm{val}(t) = \infty$ iff $t = 0$;
- $\mathrm{val}(tu) = \mathrm{val}(t) + \mathrm{val}(u)$;

- $\mathrm{val}(t+u) \geqslant \min\{\mathrm{val}(t), \mathrm{val}(u)\}$ (ultrametric inequality);
- If $\mathrm{val}(t) < \mathrm{val}(u)$ then $\mathrm{val}(t+u) = \mathrm{val}(t)$ (a consequence of the ultrametric inequality);
- $\mathrm{val}(p) = 1$.

With this p-adic terminology, we are now ready to define a *separating* polynomial:

DEFINITION 2.1. A polynomial h with integer coefficients *separates* a set $A \subset \mathbb{Z}$ from a set $B \subset \mathbb{Z}$ if

$$\max_{x \in A} \mathrm{val}(h(x)) < \min_{x \in B} \mathrm{val}(h(x)).$$

If $A = \{\alpha\}$, we say that h separates α from B.

Note that the definition is not symmetric in A and B.

Given our set L and number α, we are interested in polynomials which separate α from $L + q\mathbb{Z}$.

DEFINITION 2.2. We introduce the following quantities:

- Let $D(L, \alpha, q)$ denote the minimum possible degree of a polynomial separating α from $L + q\mathbb{Z}$.
- Let $D(s, k)$ be the maximum value of $D(L, \alpha, p^k)$, taken over all p, all $L \subset \mathbb{N}_q$ of size $|L| = s$, and all $\alpha \notin L (\mathrm{mod}\ q)$.
- Let $D(s) = \max_k \{D(s, k)\}$.

It is not obvious from the definition that any of these quantities even exists; we will show, however, that all of them are well-defined.

2.2. *An Optimization Problem*

Our attempt to determine $D(s, k)$ leads us to an interesting optimization problem:

DEFINITION 2.3. Let $S(s, k)$ denote the maximal value of $\sum_{i=1}^{k} s_i \ell_i$ where:

- $s_1, ..., s_k$ are nonnegative integers satisfying $\sum_{i=1}^{k} s_i = s$.
- For all t, $\ell_t = \lfloor \sum_{i=1}^{t-1} \ell_i s_i (1 - \frac{i}{t}) \rfloor + 1$.

We will show in Section 6.2 that $D(s, k) = S(s, k)$.

3. POLYNOMIALS WITH CONTROLLED p-ADIC BEHAVIOR

For completeness, we include in Section 3.1 the simple proof of the Deza–Frankl–Singhi inequality (1) given in [1] (cf. [3, p. 103]). This proof will motivate the basic idea of this paper. In Section 3.2, we extend the proof to prime power moduli.

The method presented herein uses spaces of multivariate polynomials in the spirit of [1, 2] (cf. [3, Chap. 5]). In contrast, the original proof in [4] (cf. [3, Example 7.4.15]) was based on the technique of higher incidence matrices introduced by Ray-Chaudhuri and Wilson [9] and successfully employed in the modular setting by [5]. This method yields important additional information; therefore, in Section 8, we explain this approach and extend it to prime power moduli as well.

3.1. *The Case of Prime Modulus*

Proof (of Inequality (1)). We use our standard assumptions from Section 2 with $q=p$. We need to prove that $|\mathscr{F}| \leqslant \binom{n}{\leqslant s}$.

Consider the univariate polynomial $h(t)=\prod_{\mu \in L}(t-\mu)$.

For each $E \in \mathscr{F}$, let v_E be the incidence vector $(\alpha_1, ..., \alpha_n)$ defined by setting $\alpha_i=1$ if $i \in E$ and $\alpha_i=0$ otherwise. Let f_E be the polynomial in n variables given by $f_E(x)=h(x \cdot v_E)$ where $x=(x_1, ..., x_n)$ and $x \cdot v_E$ denotes the dot product of these two vectors. We note that $f_E(v_F)=h(|E \cap F|)$.

The *multilinear reduction* of a monomial $\prod_{i \in I} x_i^{\ell_i} (\ell_i \geqslant 1)$ is the monomial $x_I := \prod_{i \in I} x_i$. The multilinear reduction $\bar{g}$ of a polynomial g is obtained by expanding g as a linear combination of monomials and performing the multilinear reduction of each monomial. We note that for any (0, 1)-vector v, we have $g(v)=\bar{g}(v)$.

Let us set $g_E=\overline{f_E}$. We note that $\deg(f_E) \leqslant s$ and therefore $\deg(g_E) \leqslant s$. Moreover, $g_E(v_F)=h(|E \cap F|)$.

We claim that the polynomials g_E $(E \in \mathscr{F})$ are linearly independent over $\mathbb{Q}$. Since the dimension of the space of multilinear polynomials of degree $\leqslant s$ in n variables is $\binom{n}{\leqslant s}$, we infer that $|\mathscr{F}| \leqslant \binom{n}{\leqslant s}$.

Suppose for a contradiction that there exists a nontrivial linear dependence $\sum_{E \in \mathscr{F}} \lambda_E g_E=0$ with $\lambda_E \in \mathbb{Q}$, where not all λ_E are zero. We may assume that all coefficients λ_E are integers, and some λ_F is not divisible by p. However, since $\sum_{E \in \mathscr{F}} \lambda_E f_E(v_F)=0$, we can write

$$\lambda_F h(|F|)=-\sum_{E \neq F} \lambda_E h(|E \cap F|).$$

Here, each term of the right hand side is divible by p; therefore p divides the left hand side. Now $h(|F|)$ is not divisible by p, therefore λ_F is, a contradiction with the choice of F. ∎

3.2. *Prime Power Moduli*

In place of the polynomial $h(t)=\prod_{\mu\in L}(t-\mu)$ in Section 3.1, we use a separating polynomial. The following lemma is implicit in [5].

(Recall that $q=p^k$.)

LEMMA 3.1. *Assume that, for any* $\alpha\notin L \pmod q$, *there exists a degree-d univariate polynomial* h_α *separating* α *from* $L+q\mathbb{Z}$. *Then* $|\mathscr{F}|\leqslant\binom{n}{\leqslant d}$.

The proof of this lemma will mimic the proof of the Deza–Frankl–Singhi inequality given above.

Proof. Let $r=\max_E\{\mathrm{val}(h_{|E|}(|E|))\}$. Without loss of generality, we assume $\mathrm{val}(h_{|E|}(|E|))=r$ for all E. (We can multiply $h_{|E|}$ by $p^{r-\mathrm{val}(h_{|E|}(|E|))}$.)

As above, let v_E be the incidence vector of $E\in\mathscr{F}$ and let f_E be the polynomial in n variables given by $f_E(x)=h_{|E|}(x\cdot v_E)$ where $x=(x_1, ..., x_n)$. We again note that $f_E(v_F)=h_{|E|}(|E\cap F|)$.

Still following the lines of the above proof, we set $g_E=\overline{f_E}$ (the multilinear reduction of f_E). We note that $\deg(f_E)\leqslant d$ and therefore $\deg(g_E)\leqslant d$. Moreover, $g_E(v_F)=h_{|E|}(|E\cap F|)$.

We claim that the polynomials g_E ($E\in\mathscr{F}$) are linearly independent over $\mathbb{Q}$. By the same dimension argument as above, we infer that $|\mathscr{F}|\leqslant\binom{n}{\leqslant d}$.

Suppose for a contradiction that there exists a nontrivial linear dependence $\sum_{E\in\mathscr{F}}\lambda_E g_E=0$ with $\lambda_E\in\mathbb{Q}$, where not all λ_E are zero. We may assume that all coefficients λ_E are integers, and some λ_F is not divisible by p. However, since $\sum_{E\in\mathscr{F}}\lambda_E f_E(v_F)=0$, we can write

$$\lambda_F h_{|F|}(|F|)=-\sum_{E\neq F}\lambda_E h_{|E|}(|E\cap F|).$$

Thus,

$$\mathrm{val}(\lambda_F h_{|F|}(|F|))\geqslant\min_{E\neq F}\{\mathrm{val}(\lambda_E h_{|E|}(|E\cap F|))\}.$$

For any $E\neq F$, $\mathrm{val}(h_{|E|}(|E\cap F|))>r$, so we have $\mathrm{val}(\lambda_F h_{|F|}(|F|))>r$. But $\mathrm{val}(\lambda_F)=0$, and $\mathrm{val}(h_{|F|}(|F|))=r$, so this is impossible. This contradiction proves the lemma. ∎

We now need to construct low-degree separating polynomials. When $k=1$, we can use the degree-s polynomial $h(t)=\prod_{\mu\in L}(t-\mu)$ for any $\alpha\notin L \pmod q$, as in Section 3.1. However, for $k>1$, we need to work a bit harder. The next two sections are concerned with the construction of a separating polynomial f. In Section 6, we show that this construction is optimal in the following sense: for any s, it is possible to choose q and L such that our construction is best possible.

4. TREES AND BOXES

4.1. *Rooted Trees*

Let us consider a rooted tree. Level i consists of all nodes at distance i from the root. (So the root itself constitutes level 0.) The *degree* of a node is the number of its children.

We define the "closeness" of two leaves as the level of their lowest common ancestor. ("Lowest" means at greatest distance from the root.)

We say that a tree is *levelwise regular* if nodes on the same level have the same degree. Given any rooted tree T, there is a unique (up to isomorphism) minimal levelwise regular rooted tree containing T, which we call the (levelwise regular) closure of T. Indeed, let s_i denote the maximum of the degrees of nodes on level i. Then the levelwise regular tree of degree s_i on level i is clearly the closure of T.

Note that the number of leaves of the closure of T is $\prod_{i=0}^{k-1} s_i$. (So the closure feels somewhat like a Cartesian product, justifying the term "box" to be introduced in Section 4.2.)

The following lemma gives a useful upper bound on the size of the closure of T.

LEMMA 4.1. *Let T be a rooted tree and let d be the number of levels of T on which some node has at least two children. Let s denote the number of leaves of T. Then the number of leaves of the closure of T is at most* $(1+\frac{s-1}{d})^d$.

Proof. Let s_i be the maximum of the degrees of the nodes on level i in T. Let $m=\prod_{i=0}^{k-1} s_i$, where k is the depth of T. As mentioned above, m is the number of leaves of the closure of T.

Let I denote the set $\{i: s_i>1\}$; then $|I|=d$. Since $m=\prod_{i\in I} s_i$, by the inequality between the arithmetic and the geometric means we deduce that $m\leqslant((\sum_{i\in I} s_i)/d)^d$.

Let T_j be the subtree of T containing all nodes at levels $\leqslant j$. By induction on j, we see that the number of leaves of T_j is at least $1+\sum_{i<j}(s_i-1)$. Hence $s-1\geqslant\sum_{i=0}^{k-1}(s_i-1)=\sum_{i\in I}(s_i-1)$, so $s-1+d\geqslant\sum_{i\in I} s_i$. Combining this inequality with the one in the preceding paragraph we obtain $m\leqslant(1+\frac{s-1}{d})^d$.

COROLLARY 4.1. *Under the conditions of Lemma* 4.1, *the number of leaves of the closure of T is at most* $\min\{\lfloor(1+\frac{s-1}{k})^k\rfloor, 2^{s-1}\}$, *where k is the depth of T.*

Proof. $(1+\frac{s-1}{x})^x$ is a strictly increasing function of x for $x>0$. Let m denote the number of leaves of the closure of T. Since $d\leqslant k$, we immediately obtain that $m \leqslant (1+\frac{s-1}{k})^k$, so $m \leqslant \lfloor (1+\frac{s-1}{k})^k \rfloor$. Also, since $s-1 \geqslant \sum_{i \in I} (s_i - 1) \geqslant \sum_{i \in I} 1 = d$, we obtain $m \leqslant (1+\frac{s-1}{s-1})^{s-1} = 2^{s-1}$.

Remark 4.1. Observe that $\lfloor (1+\frac{s-1}{k})^k \rfloor \leqslant s^k$, with equality holding when $k=1$.

4.2. *Tries and Boxes*

A *trie* over a finite alphabet Σ is a rooted tree whose edges are labeled by elements of Σ; all edges from a node to its children must receive different labels. (So in particular, the degree of each node is $\leqslant |\Sigma|$.)

Each leaf of a trie corresponds to a string (word) over Σ, obtained by reading labels along the path from the root to the leaf. It is clear that there is a 1–1 correspondence between tries over Σ and prefix-free sets of strings over Σ. (See for example [10, pp. 122–123] for this correspondence and its use in computer science.)

The closeness of two strings in a trie is the length of their longest common prefix. This is the same as the closeness of the corresponding leaves.

We define a *box* to be a levelwise regular trie. Every trie T over Σ can be embedded in a box over Σ, corresponding to the closure of the tree underlying T. We label the additional edges arbitrarily, making sure that there are no repeated labels among the edges from a node to its children. We shall call any of the resulting boxes a *closure* of T. While the topology of the closure is unique, the labeling is not.

5. PROOF OF A SIMPLE UPPER BOUND FOR $D(s, k)$

In this section we complete the proof of Theorem 1.2 by constructing a separating polynomial satisfying the requirements of Lemma 3.1. We refer to the notation and assumptions of Section 3.

Let us write each member of L as a k-digit number in base p (including leading zeros), writing the digits in reverse order (starting with the least significant digit). We then construct the corresponding trie over the alphabet $\Sigma = \{0, ..., p-1\}$. The *closeness* of $\mu, \nu \in N_q$ in the trie (as defined above) will be precisely $\operatorname{val}(\mu - \nu)$.

Let now $L_1 \subset \mathbb{N}_q$ be the set of integers corresponding to a closure of T. We have $L \subseteq L_1$. Moreover, by Lemma 4.1 and Corollary 4.1, we see that $|L_1| \leqslant D$ where

$$D = \min \left\{ \left\lfloor \left(1+\frac{s-1}{k}\right)^k \right\rfloor, 2^{s-1} \right\}.$$

We use the set L_1 to construct a separating polynomial of degree at most $|L_1| \leqslant D$. Since D depends only on s and k, this proves that $D(s, k) \leqslant D$.

LEMMA 5.1. *Let* $D = \min\{\lfloor (1 + \frac{s-1}{k})^k \rfloor,\ 2^{s-1}\}$. *Under our standard assumptions* (*Section* 2), *for any* $\mu_0 \in \mathbb{N}_q \backslash L$, *there exists a polynomial* h *of degree at most* D *which separates* $\mu_0 + q\mathbb{Z}$ *from* $L + q\mathbb{Z}$. *This implies that* $D(s, k) \leqslant D$.

Proof. Choose L_1 as above; write $L_1 = \{\mu_1, ..., \mu_m\}$. Note that it is possible that $\mu_0 \in L_1$. Define s_i as in Lemma 4.1 as applied to the trie T corresponding to L. For $i \leqslant k$, let $g_i = \prod_{j=i}^{k-1} s_j$ (so $g_k = 1$); then g_i is the number of leaves of a subtree whose root is at level i. For any leaf μ, g_i is also the number of leaves v such that $\mathrm{val}(\mu - v) \geqslant i$. Let $g = \sum_{i=1}^{k} g_i$.

Let $f(t) = \prod_{j=1}^{m} (t - \mu_j)$. If $t \equiv \mu_j \pmod q$ for some j, then, for any $i \leqslant k$, we have $|\{\mu_\ell\colon \mathrm{val}(t - \mu_\ell) \geqslant i\}| = g_i$, and therefore $\mathrm{val}(f(t)) \geqslant \sum_{i=1}^{k} g_i = g$. (We note that $\mathrm{val}(f(t))$ can be larger than g, and can even be ∞.)

If $t \not\equiv \mu_j \pmod q$ for any j, then, for $i < k$, we obtain that $|\{\mu_\ell\colon \mathrm{val}(t - \mu_\ell) \geqslant i\}|$ is either 0 or g_i (depending on whether or not the class of t modulo p^i is represented in L_1). Since $\mathrm{val}(t - \mu_i)$ can never be $\geqslant k$, we see that $\mathrm{val}(f(t)) \leqslant \sum_{i=1}^{k-1} g_i < g$. When $\mu_0 \notin L_1$, we are thus done with the construction, setting $h = f$.

Suppose, then, that $\mu_0 \in L_1$. In this case, let $h(t) = \prod_{\mu_j \neq \mu_0} (t - \mu_j)$. We claim that h separates $\mu_0 + q\mathbb{Z}$ from $L + q\mathbb{Z}$: more specifically, we claim that $\mathrm{val}(h(t)) > g - k$ for $t \in L \pmod q$, but $\mathrm{val}(h(t)) = g - k$ for $t \equiv \mu_0 \pmod q$.

Indeed, suppose $t \equiv \mu_j \pmod q$ where $\mu_j \neq \mu_0$. We know that

$$g \leqslant \mathrm{val}(f(t)) = \mathrm{val}(h(t)) + \mathrm{val}(t - \mu_0) \leqslant \mathrm{val}(h(t)) + (k-1),$$

so we have $\mathrm{val}(h(t)) \geqslant g - k + 1$ for any such t.

If, on the other hand, $t \equiv \mu_0 \pmod q$, then, by the counting argument above, $\mathrm{val}(h(t)) = \sum_{i=1}^{k-1} (g_i - 1) = g - k$.

By construction, $\deg h \leqslant \deg f = |L_1| \leqslant D$. ∎

Remark 5.1. Since $D \leqslant 2^{s-1}$, Lemma 5.1 proves that the quantities $D(L, \alpha, q)$, $D(s, k)$, and $D(s)$ from Section 2.1 all exist. In Section 6.3 we will prove that the bound $D(s) \leqslant 2^{s-1}$ is tight.

6. OPTIMAL SEPARATING POLYNOMIALS

In this section, we consider how to construct separating polynomials of minimum degree. For particular sets L, very low-degree polynomials may exist. However, if we want a bound on $D(s, k)$ as defined in Section 2.1, we need to consider all sets L for given s and k.

In Section 6.2, we relate this problem to the solution of the optimization problem defined in Section 2.2. In particular, we will show that $S(s, k) = D(s, k)$. Section 6.3 includes a short proof that $D(s) = \max_k \{D(s, k)\}$ is equal to 2^{s-1}, which shows that the construction in Section 5 is optimal.

We also give log-asymptotic estimates for $S(s, k)$. This requires a more detailed analysis, which appears in Section 7.

6.1. *Polynomials with Linear Factors*

We first prove that, when constructing separating polynomials, it suffices to consider products of linear terms:

LEMMA 6.1. *Assume that there exists a degree-d polynomial over $\mathbb{Z}$ separating α from B. Then there exists a polynomial separating α from B which is a product of at most d linear terms of the form $(x-\mu)$, $\mu \in B$.*

Proof. Let $a(x)$ be a maximal product of linear terms of the form $(x-\mu)$, $\mu \in B$ such that $\deg(a(x)) \leqslant d$ and $a(x)$ divides some polynomial of degree d separating α from B; we denote this polynomial $h(x)$, and write $h(x) = a(x)\, g(x)$. We show, by contradiction, that $a(x)$ separates α from B.

Take $v \in B$ such that $\mathrm{val}(a(v))$ is minimal. Assume that $a(x)$ does not separate α from B; then

$$\mathrm{val}(a(\alpha)) \geqslant \mathrm{val}(a(v)). \tag{8}$$

Let $g(x) - g(v) = (x - v)\, b(x)$. (Note that b has integer coefficients.) Then

$$h(x) = a(x)(x-v)\, b(x) + a(x)\, g(v).$$

We will show that $f(x) = a(x)(x-v)\, b(x)$ separates α from B.

By our choice of v, for any $\mu \in B$ we have $\mathrm{val}(a(\mu)\, g(v)) \geqslant \mathrm{val}(a(v)\, g(v)) = \mathrm{val}(h(v))$. By the ultrametric inequality:

$$\mathrm{val}(f(\mu)) = \mathrm{val}(h(\mu) - a(\mu)\, g(v)) \geqslant \min\{\mathrm{val}(h(\mu)), \mathrm{val}(h(v))\} > \mathrm{val}(h(\alpha)).$$

By (8) we have

$$\mathrm{val}(a(\alpha)\, g(v)) \geqslant \mathrm{val}(a(v)\, g(v)) = \mathrm{val}(h(v)) > \mathrm{val}(h(\alpha)),$$

and hence

$$\mathrm{val}(f(\alpha)) = \mathrm{val}(h(\alpha) - a(\alpha)\, g(v)) = \mathrm{val}(h(\alpha)).$$

Thus the polynomial $f(x)$ separates α from B, which contradicts the assumption that $a(x)$ is maximal. ∎

LEMMA 6.2. *Every polynomial $h(x)$ which has minimum degree subject to the condition that it separates α from $L+q\mathbb{Z}$ also separates $\alpha+q\mathbb{Z}$ from $L+q\mathbb{Z}$.*

Proof. Suppose, for some integer m, that $\mathrm{val}(h(\alpha+mq)) \neq \mathrm{val}(h(\alpha))$. Then let $f(x)=h(x+mq)-h(x)$; we have

$$\mathrm{val}(f(\alpha))=\min\{\mathrm{val}(h(\alpha+mq)), \mathrm{val}(h(\alpha))\} \leqslant \mathrm{val}(h(\alpha)).$$

For any $\mu \in L+q\mathbb{Z}$, we also have $\mu+mq \in L+q\mathbb{Z}$, so

$$\mathrm{val}(f(\mu)) \geqslant \min\{\mathrm{val}(h(\mu+mq)), \mathrm{val}(h(\mu))\} > \mathrm{val}(h(\alpha)).$$

Thus, $f(x)$ separates α from $L+q\mathbb{Z}$. But $\deg(f)<\deg(h)$, contradicting the minimality of h. We conclude that, for all $m \in \mathbb{Z}$, $\mathrm{val}(h(\alpha+mq))=\mathrm{val}(h(\alpha))$, proving the lemma. ∎

The next lemma shows that, for any k, α, and L, if p is sufficiently large, then there exists a minimum-degree polynomial separating α from $L+q\mathbb{Z}$ which is a product of linear terms of the form $(x-\mu)$, $\mu \in L$.

LEMMA 6.3. *For any k, α and L, and any sufficiently large p, there exists a minimum-degree polynomial separating α from $L+q\mathbb{Z}$ (where $q=p^k$) which is a product of linear terms of the form $(x-\mu)$, $\mu \in L$.*

Proof. Choose $p>2^s$. By Lemma 5.1, there is some polynomial $h(x)$ of minimum degree d separating α from $L+q\mathbb{Z}$, and $d \leqslant 2^{s-1}<p/2$. By Lemma 6.1, we may assume $h(x)$ is a product of linear factors:

$$h(x)=\prod_{i=1}^{d}(x-v_i)$$

with $v_i \in L+q\mathbb{Z}$. For each i, choose $\mu_i \in L$ such that $v_i \equiv \mu_i \pmod{q}$. We claim that the polynomial

$$f(x)=\prod_{i=1}^{d}(x-\mu_i)$$

also separates α from $L+q\mathbb{Z}$.

Indeed, since $\alpha \notin L+q\mathbb{Z}$, it is clear that $\mathrm{val}(f(\alpha))=\mathrm{val}(h(\alpha))$. Choose any $v \in L+q\mathbb{Z}$; since $p>2d$, there is some $\lambda \equiv v \pmod{q}$ such that, for all i, $\lambda \not\equiv v_i \pmod{p^{k+1}}$ and $\lambda \not\equiv \mu_i \pmod{p^{k+1}}$. Thus, for any i, $\mathrm{val}(\lambda-v_i)=\mathrm{val}(\lambda-\mu_i)$, and

$$\mathrm{val}(f(v)) \geqslant \mathrm{val}(f(\lambda))=\mathrm{val}(h(\lambda))>\mathrm{val}(h(\alpha))=\mathrm{val}(f(\alpha)).$$

This proves that f separates α from $L+q\mathbb{Z}$. ∎

Remark 6.1. We need the condition that p is sufficiently large in Lemma 6.3. For example, consider $q=27$, $\alpha=0$, and $L=\{1, 3, 6, 9, 12, 15, 18, 21\}$. The minimum-degree separating polynomial of the form $\prod(x-\mu)$ with $\mu \in L$ is

$$(x-9)(x-18)(x-3)(x-12)(x-21)(x-6)(x-15)(x-1)^4$$

which has degree 11. However, this is not a minimum-degree polynomial separating α from $L+q\mathbb{Z}$, since

$$(x-9)(x-18)(x-3)(x-12)(x-21)(x-6)(x-15)(x-1)(x-28)(x-55)$$

is a degree-10 separating polynomial.

6.2. *The Connection to the Optimization Problem*

We recall the optimization problem defined in Section 2.2:

Definition 6.1. Let $S(s, k)$ denote the maximal value of $\sum_{i=1}^{k} s_i \ell_i$ where:

- $s_1, ..., s_k$ are nonnegative integers with $\sum_{i=1}^{k} s_i = s$.
- For all t, $\ell_t = \lfloor \sum_{i=1}^{t-1} \ell_i s_i (1-\frac{i}{t}) \rfloor + 1$.

In this section, we show that $D(s, k) = S(s, k)$.

Lemma 6.4. *$D(s, k) \leqslant S(s, k)$.*

Proof. We are given some α and L; assume, without loss of generality, that $\alpha=0$. Let $L_i=\{\mu \in L \mid \mathrm{val}(\mu)=i\}$, and let $s_i=|L_i|$. We will attempt to construct a separating polynomial which is a product of linear terms of the form $(x-\mu)$, $\mu \in L$. Let $h(x)$ be such a polynomial, and write $h(x)=\prod_{\mu \in L}(x-\mu)^{a_\mu}$. Let $d_i=\sum_{\mu \in L_i} a_\mu$. We then have

$$\mathrm{val}(h(0)) = \sum_{i=0}^{k-1} i d_i$$

and, for any j, and any $v \in L_j + q\mathbb{Z}$:

$$\mathrm{val}(h(v)) = \sum_{i=0}^{j-1} i d_i + \sum_{i=j+1}^{k-1} j d_i + \sum_{\lambda \in L_j} a_\lambda \, \mathrm{val}(v-\lambda)$$

and hence

$$\mathrm{val}(h(v)) \geqslant \sum_{i=0}^{j} i d_i + \sum_{i=j+1}^{k-1} j d_i + (k-j)\, a_\mu \tag{9}$$

where $\mu \in L_j$ such that $v \equiv \mu (\mathrm{mod}\ q)$.

This implies that $h(x)$ will separate 0 from $L + q\mathbb{Z}$ if, for any $j \in \{0, ..., k-1\}$ and any $\mu \in L_j$,

$$a_\mu > \frac{1}{k-j} \sum_{i=j+1}^{k-1} (i-j)\, d_i . \tag{10}$$

(This condition will be necessary if we have equality in (9).)

We now attempt to construct $h(x)$ satisfying (9). Note that the expression (10) depends only on those d_i with $i > j$. Thus, when we construct $h(x)$, if we increase a_v for some v with $\mathrm{val}(v) \leqslant j$, this will not affect the validity of inequality (10). If some a_v with $\mathrm{val}(v) > j$ is increased, this can only increase a_μ, and hence the degree $d = \sum_{i=0}^{k-1} d_i$ can only increase. Hence it is optimal to make a_μ as small as possible for each μ in L_j, starting with $j = k-1$ and working down to $j = 0$.

The lower bound does not depend on a specific choice of $\mu \in L_j$, so for every $\mu \in L_j$ it is optimal to choose the same minimal value of a_μ, which we will denote ℓ_j; then, for all j, $d_j = \ell_j s_j$. The optimal choice for ℓ_j is

$$\ell_j = \left\lfloor \sum_{i=j+1}^{k-1} \left(1 - \frac{k-i}{k-j}\right) \ell_i s_i \right\rfloor + 1.$$

The degree of h(x) is $\sum_{i=0}^{k-1} \ell_i s_i$.

Since L can be any set of size s, we take the maximum over all $\{s_i\}$ with $s_0 + \cdots + s_{k-1} = s$. After a change of indexing we get the problem in the definition of $S(s, k)$, so our upper bound on $D(s, k)$ is exactly $S(s, k)$. ∎

Note that if we have equality in (9) then we have construct an optimal separating polynomial among polynomials which are products of linear terms of the form $(x-\mu)$, $\mu \in L$. There is equality in (9) iff $\mathrm{val}(\mu - v) = i$ for every $i \in \{0, ..., k-1\}$ and every distinct $\mu, v \in L_i$. If $p > s$ then, for any $s_0, ..., s_{k-1}$, it is possible to choose L satisfying this condition.

We are now in a position to show that $S(s, k)$ gives the best possible bound on $D(s, k)$.

Theorem 6.1. *For all s and k, $D(s, k) = S(s, k)$.*

Proof. We have already shown $D(s, k) \leqslant S(s, k)$. For any s, choose $p > s$ sufficiently large for Lemma 6.3 to apply. Choose $s_0, ..., s_k$ such that the sum in the definition of $S(s, k)$ achieves its maximum value. Chose L such that, for all i, $|L_i| = s_{k-i}$, and such that, for distinct $\mu, \nu \in L_i$, $\mathrm{val}(\mu - \nu) = i$. (As mentioned above, this is possible since $p > s$.)

Let $h(x)$ be a polynomial separating 0 from $L + q\mathbb{Z}$, and write $d = \deg(h)$. By Lemma 6.3, there is a degree-d polynomial separating 0 from $L + q\mathbb{Z}$ which is a product of factors $(x - \mu)$ with $\mu \in L$. By the proof of Lemma 6.4, and by our choice of L, any such polynomial has degree at least $S(s, k)$. This proves that $D(s, k) \geqslant S(s, k)$.

6.3. *Estimating $D(s, k)$*

By Theorem 6.1, we can find the minimum degree of a separating polynomial by finding $S(s, k)$, or optimizing the sum in Definition 2.3. In Section 7, we prove the following asymptotic estimates for $D(s, k) = S(s, k)$.

We use Knuth's "Θ" notation: two functions $f, g \colon \Omega \to \mathbb{R}$ are said to satisfy the relation $f = \Theta(g)$ if there exist positive constants c_1, c_2 such that

$$c_1 \, |f(x)| \leqslant |g(x)| \leqslant c_2 \, |f(x)|$$

for all but a finite number of values $x \in \Omega$. Our set Ω will typically be $\mathbb{N} \times \mathbb{N}$.

THEOREM 6.2. *For $k \leqslant \sqrt{s/e}$*

$$\ln S(s, k) = \Theta\left(k\left(2 + \ln \frac{s}{k^2}\right)\right).$$

For $\sqrt{s/e} \leqslant k \leqslant e^{s/2}$

$$\ln S(s, k) = \Theta\left(\sqrt{s\left(2 + \ln \frac{k^2}{s}\right)}\right).$$

For $e^{s/2} \leqslant k$

$$\ln S(s, k) = \Theta(s).$$

Furthermore, we can precisely answer another question: what is the best bound we can give depending only on s?

As in Section 2.1, let $D(s) = \max_k \{D(s, k)\} = \max_k \{S(s, k)\}$. We showed in Lemma 5.1 that $D(s)$ exists, and that it is at most 2^{s-1}. We now show that $D(s) = 2^{s-1}$.

THEOREM 6.3. *For any s there exist p, k, α and L, with $|L| = s$, such that the minimum degree of a polynomial separating α from $L + q\mathbb{Z}$ (where $q = p^k$) is exactly 2^{s-1}.*

Proof. Following Theorem 6.1, it suffices to find k and $s_1, ..., s_k$ such that

$$\sum_{i=1}^{k} s_i = s$$

$$\ell_t = \left\lfloor \sum_{i=1}^{t-1} \ell_i s_i \left(1 - \frac{i}{t}\right) \right\rfloor + 1$$

$$\sum_{i=1}^{k} \ell_i s_i = 2^{s-1}$$

We will inductively construct an infinite sequence $\{s_i\}$, $s_i \in \{0, 1\}$, as follows: let π_i denote the position of the ith one. Let $\pi_1 = 1$; then $s_1 = 1$, and $\ell_1 = 1$.

Suppose, for some j, we have found $\pi_1, ..., \pi_{j-1}$; these determine $s_1, ...,$ $s_{\pi_{j-1}}$, and we can use the expression for ℓ_t to compute $\ell_1, ..., \ell_{\pi_{j-1}}$. We will let $\pi_j = \sum_{i=1}^{j-1} \pi_i \, \ell_{\pi_i}$.

We conclude that, for any $j > 1$,

$$\ell_{\pi_j} = \left\lfloor \sum_{i=1}^{j-1} \ell_{\pi_i} - \frac{1}{\pi_j} \sum_{i=1}^{j-1} \pi_i \ell_{\pi_i} \right\rfloor + 1 = \sum_{i=1}^{j-1} \ell_{\pi_i}.$$

Thus $\ell_{\pi_j} = 2^{j-2}$ for $j > 1$.

Finally, let $k = \pi_s$. Then

$$\sum_{i=0}^{k} s_i \ell_i = \sum_{j=1}^{s} \ell_{\pi_j} = 2^{s-1}.$$

This proves the theorem. ∎

7. THE OPTIMIZATION PROBLEM

In this section we determine the logarithmic order of magnitude of $S(s, k)$ as stated in Theorem 6.2. In Section 7.1, we introduce two simplified versions of the optimization problem, without the floor function, and show that the optima remain within a constant factor. In Section 7.2, we examine a related maximization problem. In Sections 7.3 and 7.4, we prove upper and lower bounds on $S(s, k)$. Finally, in Section 7.5, we complete the proof of Theorem 6.2.

7.1. *The Problem without the Floor Function*

We first restate the definition in a simpler form, without the floor function. There are two approaches we could take: overestimating ℓ_i, or underestimating ℓ_i.

Definition 7.1. Let $Z(s, k)$ denote the maximal value of $\sum_{i=1}^{k} s_i \check{\ell}_i$ where:

- $s_1, ..., s_k$ are nonnegative integers with $\sum_{i=1}^{k} s_i = s$.
- $\check{\ell}_i = \sum_{j=1}^{i-1} \check{\ell}_j s_j (1 - \frac{j}{i}) + 1 \quad (1 \leqslant i \leqslant k)$.

Definition 7.2. Let $T(s, k)$ denote the maximal value of $\sum_{i=1}^{k} s_i \hat{\ell}_i$ where

- $s_1, ..., s_k$ are nonnegative integers with $\sum_{i=1}^{k} s_i = s$
- $\hat{\ell}_i = \sum_{j=1}^{i-1} \hat{\ell}_j s_j (1 - \frac{j}{i}) \quad (1 \leqslant i \leqslant k)$.

For any sequence $\{s_i\}$, we clearly have $\hat{\ell}_i \leqslant \ell_i \leqslant \check{\ell}_i$, so we can conclude that $T(s, k) \leqslant S(s, k) \leqslant Z(s, k)$.

We now make precise statements about $Z(s, k)$ and $T(s, k)$:

Lemma 7.1.

$$\sum_{i=1}^{k} s_i \check{\ell}_i = \sum_{m=1}^{k} \sum_{1 \leqslant a_1 < a_2 < \cdots < a_m \leqslant k} \left(1 - \frac{a_1}{a_2}\right)\left(1 - \frac{a_2}{a_3}\right) \cdots \left(1 - \frac{a_{m-1}}{a_m}\right) s_{a_1} s_{a_2} \cdots s_{a_m} \tag{11}$$

Proof. By induction on k. The lemma holds for $k = 1$. For the induction step we need to show

$$s_{k+1} \check{\ell}_{k+1} = s_{k+1} + \sum_{m=1}^{k} \sum_{1 \leqslant a_1 < a_2 < \cdots < a_m \leqslant k} \left(1 - \frac{a_1}{a_2}\right) \cdots \left(1 - \frac{a_{m-1}}{a_m}\right)\left(1 - \frac{a_m}{k+1}\right) s_{a_1} \cdots s_{a_m} s_{k+1}.$$

Using the recurrence for $\check{\ell}_{k+1}$ and the induction hypothesis we obtain:

$$\check{\ell}_{k+1} - 1 = \sum_{i=1}^{k} \check{\ell}_i s_i \frac{k+1-i}{k+1} = \frac{1}{k+1} \sum_{j=1}^{k} \sum_{i=1}^{j} \check{\ell}_i s_i$$
$$= \frac{1}{k+1} \sum_{j=1}^{k} \sum_{m=1}^{j} \sum_{1 \leqslant a_1 < \cdots < a_m \leqslant j} \left(1 - \frac{a_1}{a_2}\right) \cdots \left(1 - \frac{a_{m-1}}{a_m}\right) s_{a_1} \cdots s_{a_m}$$
$$= \frac{1}{k+1} \sum_{m=1}^{k} \sum_{1 \leqslant a_1 < \cdots < a_m \leqslant k} \left(1 - \frac{a_1}{a_2}\right) \cdots$$
$$\left(1 - \frac{a_{m-1}}{a_m}\right) s_{a_1} \cdots s_{a_m} (k+1-a_m)$$

which proves the lemma. ∎

LEMMA 7.2.

$$\sum_{i=1}^{k} s_i \hat{\ell}_i = \sum_{m=1}^{k} \sum_{1 = a_1 < a_2 < \cdots < a_m \leqslant k} \left(1 - \frac{a_1}{a_2}\right)\left(1 - \frac{a_2}{a_3}\right) \cdots$$
$$\left(1 - \frac{a_{m-1}}{a_m}\right) s_{a_1} s_{a_2} \cdots s_{a_m}. \tag{12}$$

(Note that in (12) we have $1 = a_1$ whereas in (11) we have $1 \leqslant a_1$.)

Proof. Similar to the proof of Lemma 7.1. ∎

We will use these expressions to determine upper bounds for $Z(s, k)$ and matching lower bounds for $T(s, k)$. However, it is worth observing that this first step of using $Z(s, k)$ and $T(s, k)$ instead of $S(s, k)$ costs at most a factor of 3.

LEMMA 7.3.

$$Z(s, k) \leqslant 3T(s, k),$$

and thus

$$\tfrac{1}{3} Z(s, k) \leqslant S(s, k) \leqslant Z(s, k).$$

Proof. Let $s_1, ..., s_k$ be the optimal assignment for $Z(s, k)$. Note that $s_1 > 0$. If not we can set $s'_i = s_{i+1}$, $1 \leqslant i < k$ and obtain $Z(s', k) > Z(s, k)$. Now

$$
\begin{aligned}
Z(s,k)-T(s,k) &= \sum_{m=1}^{k} \sum_{1<a_1<a_2<\cdots<a_m\leqslant k} \left(1-\frac{a_1}{a_2}\right)\left(1-\frac{a_2}{a_3}\right)\cdots \\
&\qquad \left(1-\frac{a_{m-1}}{a_m}\right) s_{a_1}s_{a_2}\cdots s_{a_m} \\
&\leqslant \frac{2}{s_1}\cdot \sum_{m=1}^{k} \sum_{1<a_1<a_2<\cdots<a_m\leqslant k} \left(1-\frac{1}{a_1}\right)\left(1-\frac{a_1}{a_2}\right)\cdots \\
&\qquad \left(1-\frac{a_{m-1}}{a_m}\right) s_1 s_{a_1}s_{a_2}\cdots s_{a_m} \\
&\leqslant \frac{2}{s_1}\cdot T(s,k)\leqslant 2\cdot T(s,k) \quad \blacksquare \qquad (13)
\end{aligned}
$$

7.2. *Estimating* $(1-a_1/a_2)\cdots(1-a_m/a_{m+1})$

To estimate $Z(s,k)$, we will need to understand the following maximization problem, which arises from Lemma 7.1:

Definition 7.3. Let $M(m,k)$ denote the maximal value of

$$\left(1-\frac{a_1}{a_2}\right)\left(1-\frac{a_2}{a_3}\right)\cdots\left(1-\frac{a_{m-1}}{a_m}\right)\left(1-\frac{a_m}{a_{m+1}}\right) \tag{14}$$

where $1\leqslant a_1<a_2\cdots<a_m<a_{m+1}\leqslant k$ and each $a_i\in\mathbb{Z}$.

We now prove a series of bounds on $M(m,k)$.

Lemma 7.4. *For* $m\leqslant k/(2e)$

$$M(m,k)>\sqrt{\frac{2m}{k}}\left(\frac{\ln\dfrac{k}{2m}}{2m}\right)^m.$$

Proof. Let $t=(k/2m)^{1/m}$ and let the sequence $\{a_i\}$ be roughly a geometric progression with quotient t:

$$a_{m+1-i}=\begin{cases} k & \text{for}\quad i=0 \\ \left\lceil \dfrac{a_{m+2-i}}{t}\right\rceil & \text{for}\quad 1\leqslant i\leqslant m \end{cases}$$

Clearly $a_{m+1-i} \geqslant \frac{k}{t^i}$. We have

$$1-\frac{a_i}{a_{i+1}} \geqslant 1-\frac{\dfrac{a_{i+1}}{t}+1}{a_{i+1}} \geqslant 1-\frac{1}{t}-\frac{1}{a_2} \geqslant \frac{\ln t}{\sqrt{t}}-\frac{t^{m-1}}{k}=\frac{\ln t}{\sqrt{t}}\left(1-\frac{t^{m-1/2}}{k \ln t}\right).$$

Using

$$\frac{t^{m-1/2}}{k \ln t} \leqslant \frac{t^m}{k \ln t}=\frac{k/2m}{k\,\dfrac{1}{m} \ln(k/2m)}=\frac{1}{2 \ln(k/2m)} \leqslant \frac{1}{2},$$

we obtain

$$1-\frac{a_i}{a_{i+1}} \geqslant \frac{\ln t}{2\sqrt{t}}>0.$$

(Note that this also proves that the a_i are distinct.) Thus,

$$M(m,k) \geqslant \left(\frac{\ln t}{2\sqrt{t}}\right)^m=\sqrt{\frac{2m}{k}}\left(\frac{\ln \dfrac{k}{2m}}{2m}\right)^m. \quad \blacksquare$$

Lemma 7.5. *For $m \leqslant k$*

$$M(m,k) \geqslant \frac{1}{(m+1)!}.$$

Proof. Take $a_i=i$ for $1 \leqslant i \leqslant m+1$. $\blacksquare$

The following weak upper bound on $M(m,k)$ does not use the fact that the a_i are integers.

Lemma 7.6.

$$M(m,k) \leqslant (1-k^{-1/m})^m.$$

Proof. Using the inequality between the arithmetic and geometric means we obtain:

$$\left(1-\frac{a_1}{a_2}\right)\cdots\left(1-\frac{a_m}{a_{m+1}}\right) \leqslant \left(\frac{m-\sum_{i=1}^{m}\dfrac{a_i}{a_{i+1}}}{m}\right)^m$$

$$\leqslant \left(1-\left(\frac{a_1}{a_{m+1}}\right)^{1/m}\right)^m \leqslant (1-k^{-1/m})^m. \quad \blacksquare \qquad (15)$$

LEMMA 7.7.

$$M(m,k) \leqslant \left(\frac{e \ln \dfrac{ek}{m}}{m} \right)^m .$$

Proof. For $m=1$ the upper bound is valid because it is greater than 1. We have

$$M(m+1,k) = \max_{m<\ell<k} M(m,\ell)\left(1-\frac{\ell}{k}\right).$$

Hence, to prove the upper bound it suffices to show that, for any $m, k \in \mathbb{N}$ and any $\ell \in \mathbb{R}$ such that $m+1 \leqslant \ell \leqslant k-1$:

$$\frac{\left(\dfrac{e \ln \dfrac{e\ell}{m}}{m} \right)^m \left(1-\dfrac{\ell}{k}\right)}{\left(\dfrac{e \ln \dfrac{ek}{m+1}}{m+1} \right)^{m+1}} \leqslant 1.$$

Let $L(\ell, m, k)$ denote the left-hand side of the above inequality. We have

$$L(\ell,m,k) = \frac{(m+1)^{m+1}}{m^m} \frac{1}{e} \frac{\left(\ln \dfrac{e\ell}{m}\right)^m}{\left(\ln \dfrac{ek}{m+1}\right)^{m+1}} \left(1-\frac{\ell}{k}\right)$$

$$= \frac{(m+1)^{m+1}}{m^m} \frac{1}{e} \left(1 + \frac{\ln\left(1+\dfrac{1}{m}\right)}{\ln \dfrac{e\ell}{m+1}} \right)^m \frac{\left(\ln \dfrac{e\ell}{m+1}\right)^m}{\left(\ln \dfrac{ek}{m+1}\right)^{m+1}} \left(1-\frac{\ell}{k}\right). \tag{16}$$

Let $a=\frac{e\ell}{m+1}$ and $b=\frac{ek}{m+1}$. Using $1-x\leqslant\ln\frac{1}{x}$ and $\ln(1+x)\leqslant x$ we obtain

$$\begin{aligned}L(\ell, m, k) &\leqslant \frac{(m+1)^{m+1}}{m^m}\,\frac{1}{e}\left(1+\frac{\frac{1}{m}}{\ln a}\right)^m \frac{(\ln a)^m}{(\ln b)^{m+1}}\ln\frac{b}{a} \\ &\leqslant \frac{(m+1)^{m+1}}{m^m}\,\frac{1}{e}\left(1+\frac{1}{m}\right)^m\left(\left(\frac{\ln a}{\ln b}\right)^m-\left(\frac{\ln a}{\ln b}\right)^{m+1}\right).\end{aligned}$$

Clearly $0\leqslant\frac{\ln a}{\ln b}\leqslant 1$ and, for $0\leqslant x\leqslant 1$, we have $x^m-x^{m+1}\leqslant\frac{m^m}{(m+1)^{m+1}}$. Thus

$$L(\ell, m, k)\leqslant\frac{1}{e}\left(1+\frac{1}{m}\right)^m\leqslant 1. \quad \blacksquare$$

7.3. *Upper Bounds on $Z(s, k)$*

The mth symmetric polynomial $\sigma_m(s_1, ..., s_k)$ is the coefficient of x^{k-m} in $\prod_{i=1}^{k}(x+s_k)$. We define the m-th symmetric mean as follows:

$$p_m(s_1, ..., s_k)=\left(\frac{\sigma_m(s_1, ..., s_k)}{\binom{k}{m}}\right)^{1/m}.$$

Note that $p_1(s_1, ..., s_k)$ is the arithmetic mean and $p_k(s_1, ..., s_k)$ is the geometric mean. In [8, p. 52] the following generalization of the inequality between the arithmetic and geometric means is shown:

Lemma 7.8. *For non-negative $s_1, ..., s_k$:*

$$p_1(s_1, ..., s_k)\geqslant p_2(s_1, ..., s_k)\geqslant\cdots\geqslant p_k(s_1, ..., s_k).$$

For each of the inequalities, equality holds iff all the s_i are equal.

Lemma 7.9.

$$Z(s, k)\leqslant\sum_{m=1}^{k}\binom{k}{m}\left(\frac{s}{k}\right)^m M(m-1, k). \tag{17}$$

For $s\leqslant k$

$$Z(s, k)\leqslant\sum_{m=1}^{s}\binom{s}{m}M(m-1, k). \tag{18}$$

Proof. Choose $\{s_i\}$ maximizing $\sum_{i=1}^{k} s_i \check{\ell}_i$. Then, by Lemma 7.1,

$$Z(s,k) = \sum_{m=1}^{k} \sum_{1 \leqslant a_1 < \cdots < a_m \leqslant k} \left(1 - \frac{a_1}{a_2}\right) \cdots \left(1 - \frac{a_{m-1}}{a_m}\right) s_{a_1} \cdots s_{a_m}$$

$$\leqslant \sum_{m=1}^{k} M(m-1,k)\, \sigma_m(s_1, ..., s_k). \tag{19}$$

Using Lemma 7.8 we obtain

$$\sigma_m(s_1, ..., s_k) \leqslant \binom{k}{m} \left(\frac{s}{k}\right)^m.$$

If $s \leqslant k$, we can use the fact that the s_i are integers. In this case the function $\sigma_m(s_1, ..., s_k)$ attains its maximum when s of the s_i are 1, and hence

$$\sigma_m(s_1, ..., s_k) \leqslant \binom{s}{m}. \quad \blacksquare$$

Lemma 7.10.

$$Z(s,k) \leqslant sk \max_{0 \leqslant m \leqslant k} \left(\frac{es}{m}\right)^m M(m,k).$$

Proof. Using (17) and $\binom{k}{m} \leqslant \left(\frac{ek}{m}\right)^m$, and then replacing m with $m+1$, we obtain

$$Z(s,k) \leqslant \sum_{m=1}^{k} \left(\frac{es}{m}\right)^m M(m-1,k)$$

$$= \sum_{m=0}^{k-1} \left(\frac{es}{m}\right)^m \left(\frac{m}{m+1}\right)^{m+1} es \frac{1}{m} M(m,k)$$

$$\leqslant sk \max_{0 \leqslant m \leqslant k} \left(\frac{es}{m}\right)^m M(m,k).$$

Lemma 7.11. *For* $k \leqslant \sqrt{s/e}$

$$Z(s,k) \leqslant sk \left(\frac{e^2 s}{k^2}\right)^k.$$

Proof. Using Lemma 7.7, and the inequality $\ln x \leqslant x-1$, we obtain

$$M(m,k) \leqslant \left(\frac{e \ln \dfrac{ek}{m}}{m} \right)^m \leqslant \left(\frac{ek}{m^2} \right)^m$$

and hence by Lemma 7.10

$$Z(s,k) \leqslant sk \max_{0 \leqslant m \leqslant k} \left(\frac{e^2 sk}{m^3} \right)^m.$$

The maximum of $(\frac{e^2 sk}{m^3})^m$ is attained when $m = (sk/e)^{1/3}$ and hence if $k \leqslant (sk/e)^{1/3}$ then

$$Z(s,k) \leqslant sk \left(\frac{e^2 s}{k^2} \right)^k. \quad \blacksquare$$

LEMMA 7.12. *For $c \geqslant 1$ the solution of*

$$\ln y + 2y - \frac{1}{y} = c \tag{20}$$

is in the range $[(c - \ln \frac{c}{2})/2, (c - \ln \frac{c}{2})/2 + \varDelta]$, where

$$\varDelta = \frac{1}{c - \ln \dfrac{c}{2}} - \frac{1}{2} \ln \left(1 - \frac{\ln \dfrac{c}{2}}{c} \right). \tag{21}$$

Proof. Let $f(y) = \ln y + 2y - \frac{1}{y}$. Then

$$f\left(\frac{c - \ln \dfrac{c}{2}}{2} \right) - c = \ln \left(1 - \frac{\ln \dfrac{c}{2}}{c} \right) - \frac{2}{c - \ln \dfrac{c}{2}} < 0.$$

Note that $f'(y) = \frac{1}{y} + 2 + \frac{1}{y^2} > 2$ and hence

$$f\left(\frac{c - \ln \dfrac{c}{2}}{2} + \varDelta \right) - c \geqslant 0. \quad \blacksquare$$

LEMMA 7.13. *For* $k \geqslant \sqrt{s/e}$

$$Z(s, k) \leqslant ks \cdot e^{2e\sqrt{s(1 + 1/2 \ln k^2/s)}}.$$

Proof. Using Lemma 7.10 and Lemma 7.7 we obtain

$$Z(s, k) \leqslant ks \max_{0 \leqslant m \leqslant k} \left(\frac{e^2 s \ln \dfrac{ek}{m}}{m^2} \right)^m. \tag{22}$$

Let $x = \frac{m}{ek}$. We have

$$Z(s, k) \leqslant ks \max_{0 \leqslant x \leqslant 1/e} \left(\frac{\dfrac{s}{k^2} \ln 1/x}{x^2}. \right)^{x \cdot ek} \tag{22}$$

The maximum value of

$$\left(\frac{\dfrac{s}{k^2} \ln 1/x}{x^2} \right)^x$$

is attained when

$$\ln \ln \frac{1}{x} + 2 \ln \frac{1}{x} - \frac{1}{\ln \dfrac{1}{x}} = 2 + \ln \frac{k^2}{s}.$$

By Lemma 7.12, when this occurs, we have

$$\frac{c - \ln \dfrac{c}{2}}{2} \leqslant \ln \frac{1}{x} \leqslant \frac{c - \ln \dfrac{c}{2}}{2} + \Delta,$$

where $c = 2 + \ln \frac{k^2}{s}$ and Δ is given by (21). This implies that

$$\frac{1}{x^2} \leqslant e^{c - \ln c/2 + 2\Delta} \qquad \text{and} \qquad x \leqslant e^{(-c + \ln c/2)/2} = \sqrt{\frac{1}{e^2} \cdot \frac{s}{k^2} \left(1 + \frac{1}{2} \ln \frac{k^2}{s} \right)}$$

and hence

$$\max\left(\frac{\frac{s}{k^2}\ln 1/x}{x^2}\right)^x \leqslant \left(\frac{s}{k^2}\left(\frac{c-\ln\frac{c}{2}}{2}+\varDelta\right)e^{c-\ln c/2+2\varDelta}\right)^x$$

$$=\left(\frac{c-\ln\frac{c}{2}+2\varDelta}{c}\,e^{2\varDelta+2}\right)^x. \tag{23}$$

Since we know that $c \geqslant 2$ and $\varDelta \leqslant 1$, and therefore that $(c-\ln\frac{c}{2}+2\varDelta)/c \leqslant 2$, we can conclude that

$$\left(\frac{c-\ln\frac{c}{2}+2\varDelta}{c}\,e^{2\varDelta+2}\right)^x \leqslant e^{2\sqrt{s/k^2(1+1/2\ \ln k^2/s)}}.$$

Substituting into inequality (22) we obtain

$$Z(s,k) \leqslant ks\cdot e^{2e\sqrt{s(1+1/2\ \ln k^2/s)}}. \quad \blacksquare$$

7.4. *Lower Bounds on* $T(s,k)$

LEMMA 7.14. *For any* $n \in \mathbb{Z}$, $1 \leqslant n \leqslant \min\{s,k\}$

$$T(s,k) \geqslant M(n-1,k)\left(\frac{s}{n}-1\right)^n. \tag{24}$$

Proof. In (12) consider only the part of the sum where $m=n$. Choose $1=a_1<a_2<\cdots<a_m \leqslant k$ such that

$$M(m-1,k)=\left(1-\frac{a_1}{a_2}\right)\left(1-\frac{a_2}{a_3}\right)\cdots\left(1-\frac{a_{m-1}}{a_m}\right).$$

For $1 \leqslant i \leqslant m$ let $s_{a_i}=\lfloor s/m\rfloor$. For this choice of $\{s_i\}$, the contribution of one term in the sum (12) is at least the right-hand side of (24). $\blacksquare$

LEMMA 7.15. *For* $k \geqslant \sqrt{2s}$, $\ln k \leqslant s/2$

$$T(s,k) \geqslant \frac{1}{e^3}\,e^{0.03\sqrt{s(2+\ln k^2/2s)}}.$$

Proof. If $\sqrt{s(2+\ln\frac{k^2}{2s})} \leqslant 100$, the result is trivial. We therefore assume for the remainder of the proof that $\sqrt{s(2+\ln\frac{k^2}{2s})} > 100$.

Using Lemma 7.14 and Lemma 7.4 we obtain that for any $m \in \mathbb{N}$, $m \leqslant \min\{k/(2e), s\}$,

$$T(s,k) \geqslant \sqrt{\frac{2m}{k}} \left(\frac{\ln \frac{k}{2m}}{2m} \right)^m \left(\frac{s}{m} - 1 \right)^m \geqslant \sqrt{\frac{2m}{k}} \left(\frac{s \ln \frac{k}{2m}}{2m^2} \right)^m e^{-m^2/s - m}.$$

Let $y = \frac{2m}{k}$. We have

$$T(s,k) \geqslant \sqrt{y} \left(\frac{\frac{2s}{k^2} \ln \frac{1}{y}}{y^2} \right)^{y \cdot k/2} e^{-m^2/s - m}. \tag{25}$$

We can approximate the maximum of the right-hand side of (25) with the help of Lemma 7.12 from Section 7.3. If x is such that

$$\ln \frac{1}{x} = \frac{c - \ln \frac{c}{2}}{2}$$

where $c = 2 + \ln \frac{k^2}{2s}$ (note that $c \geqslant 2$), then

$$\frac{\frac{2s}{k^2} \ln \frac{1}{x}}{x^2} = \frac{2s}{k^2} \cdot \frac{c - \ln \frac{c}{2}}{2} e^{c - \ln c/2} = e^2 \frac{c - \ln \frac{c}{2}}{c} \geqslant e^2 \left(1 - \frac{1}{2e} \right) \geqslant e^{1.75}.$$

Hence, if we choose for m the value

$$m = \left\lfloor \frac{k}{2} x \right\rfloor = \left\lfloor \frac{1}{e} \sqrt{\frac{s}{2} \left(1 + \frac{1}{2} \ln \frac{k^2}{2s} \right)} \right\rfloor \tag{26}$$

we have $y \leqslant x$, so

$$\frac{s \ln \frac{k}{2m}}{2m^2} \geqslant e^{1.75}.$$

Also, since $\sqrt{s(2 + \ln \frac{k^2}{2s})} > 100$, we have $m \geqslant 18$, and hence

$$m \geqslant \frac{18}{19e} \sqrt{\frac{s}{2} \left(1 + \frac{1}{2} \ln \frac{k^2}{2s} \right)} \geqslant \frac{18}{19e} \sqrt{\frac{s}{2}}.$$

Since $k \geqslant \sqrt{2s}$, we know $m \leqslant k/(2e)$. Since $\ln k \leqslant \frac{1}{2}s$, we get $m \leqslant \frac{1}{2e}s$ and $\ln \frac{k^2}{2s} \leqslant \sqrt{s(2+\ln \frac{k^2}{2s})}$. (The statement $m \leqslant \frac{1}{2e}s$ follows from (26) for $s \geqslant 4$; when $s<4$, we conclude from (26) that $m=0$.) Therefore,

$$T(s,k) \geqslant \sqrt{\frac{2m}{k}}\, e^{1.75m} e^{-m^2/s-m} \geqslant \sqrt{\frac{18}{19e}} \left(\frac{2s}{k^2}\right)^{1/4} e^{(1.75-1/2e-1)\,m}$$

$$\geqslant \sqrt{\frac{18}{19e}}\, e^{-(1.75-1/2e-1)} e^{1/2e(1.75-1/2e-1)\sqrt{s(2+\ln k^2/2s)}-1/4\,\ln k^2/2s}$$

$$\geqslant \frac{1}{e^3}\, e^{0.03\sqrt{s(2+\ln k^2/2s)}}. \quad \blacksquare$$

LEMMA 7.16. *For* $k \leqslant \sqrt{2s}$

$$T(s,k) \geqslant \frac{1}{ke^6}\left(\frac{es}{k}\right)^k.$$

Proof. If we chose $m=k$, and let $a_i=i$ and $s_i=\lfloor \frac{s}{k} \rfloor$ for $1 \leqslant i \leqslant m$ we obtain

$$T(s,k) \geqslant \frac{1}{k!}\left(\frac{s}{k}-1\right)^k.$$

Using $k! \leqslant ke(k/e)^k$, we get

$$T(s,k) \geqslant \frac{1}{ke}\left(\frac{e}{k}\right)^k \left(\frac{s}{k}\right)^k \left(1-\frac{k}{s}\right)^k \geqslant \frac{1}{ke}\left(\frac{es}{k^2}\right)^k e^{-k^2/s-k} \geqslant \frac{1}{ke^6}\left(\frac{es}{k^2}\right)^k.$$

(The final step is straightforward for $s \geqslant 6$, and can be checked case-by-case for smaller s.) $\blacksquare$

7.5. *Proof of Theorem* 6.2

We are now ready to prove Theorem 6.2.

Proof (of Theorem 6.2). From Lemmas 7.11, 7.13, 7.15, 7.16 we have

$$k \leqslant \sqrt{s/e} \Rightarrow \ln S(s,k) \leqslant k\left(2+\ln \frac{s}{k^2}\right)+\ln s+\ln k \tag{27}$$

$$k \geqslant \sqrt{s/e} \Rightarrow \ln S(s,k) \leqslant \sqrt{2}e\sqrt{s\left(2+\ln \frac{k^2}{s}\right)}+\ln(sk) \tag{28}$$

$$k \leqslant \sqrt{2s} \Rightarrow \ln S(s,k) \geqslant k\left(1+\ln\frac{s}{k^2}\right) - 6 - \ln k \tag{29}$$

$$k \geqslant \sqrt{2s},\ k \leqslant e^{s/2} \Rightarrow \ln S(s,k) \geqslant 0.03\sqrt{s\left(2+\ln\frac{k^2}{2s}\right)} - 3. \tag{30}$$

For $k = e^{s/2}$, inequality (30) gives $\ln S(s,k) = \Omega(s)$. We know that $S(s,k) \leqslant 2^{s-1}$ and that $S(s,k)$ is increasing in s and k. Hence for $k \geqslant e^{s/2}$ we have $\ln S(s,k) = \Theta(s)$. ∎

8. INDEPENDENT SET SYSTEMS

We now introduce the higher incidence matrices used in the original proof of the Deza–Frankl–Singhi inequality in [4] (following [5, 9]). We define the notion of an s^*-independent set system, and show that, if a set system $\mathscr{F}$ satisfies our standard assumptions of Section 2, then $\mathscr{F}$ is D^*-independent for D as in Theorem 1.2. Our presentation closely follows that in [3, Chapter 7].

Let $\mathscr{F}$ and $\mathscr{T}$ be families of subsets of a universe X of n points; we define the $(\mathscr{F}, \mathscr{T})$-*inclusion matrix* $I(\mathscr{F}, \mathscr{T})$ to be an $|\mathscr{F}| \times |\mathscr{T}|$ matrix, indexed by $\mathscr{F}$ and $\mathscr{T}$, where the entry indexed by (F, T) is 1 if $T \subseteq F$ and 0 otherwise. (Note that this matrix is only defined up to relabeling of its rows and columns, since we do not specify any labeling.)

We use $I(\mathscr{F}, s)$ as a shorthand for $I(\mathscr{F}, \binom{X}{s})$, in which the columns correspond to all possible subsets of X of size s. We call this matrix the *s-inclusion matrix* of $\mathscr{F}$. A related matrix is the *s-intersection matrix* $A_s(\mathscr{F}) = I(\mathscr{F}, s)\, I(\mathscr{F}, s)^T$; it is easy to see that the entry of $A_s(\mathscr{F})$ indexed by E and F is $\binom{|E \cap F|}{s}$.

The s^*-*inclusion matrix* of $\mathscr{F}$ is the matrix $I^*(\mathscr{F}, s)$, with dimensions $|\mathscr{F}| \times \binom{n}{\leqslant s}$, obtained by concatenating the t-inclusion matrices for $t \leqslant s$:

$$I^*(\mathscr{F}, s) = [\, I(\mathscr{F}, s) | I(\mathscr{F}, s-1) | \cdots | I(\mathscr{F}, 0)\,].$$

Definition 8.1. A family $\mathscr{F}$ is *s-independent* if

- the matrix $I(\mathscr{F}, s)$ has full row-rank $|\mathscr{F}|$.

A family $\mathscr{F}$ is s^*-*independent* if

- the matrix $I^*(\mathscr{F}, s)$ has full row-rank $|\mathscr{F}|$.

Note that s-independence implies s^*-independence, but the converse is not true. Also note that $\mathrm{rk}(A_s(\mathscr{F})) \leqslant \mathrm{rk}(I(\mathscr{F}, s))$, so if $A_s(\mathscr{F})$ is non-singular then $\mathscr{F}$ is s-independent.

It is immediate that any s-independent family has size at most $\binom{n}{s}$, and any s^*-independent family has size at most $\binom{n}{\leqslant s}$. Thus, the lemma below is strictly stronger than Lemma 3.1.

LEMMA 8.1. *Let $\mathscr{F}$ satisfy our standard assumptions. Assume that, for any $\alpha \notin L \pmod{q}$, there exists a degree-d univariate polynomial h_α separating α from $L + q\mathbb{Z}$. Then $\mathscr{F}$ is d^*-independent.*

The proof follows [5] for the case of prime modulus.

Proof. Suppose that $I^*(\mathscr{F}, d)$ does not have full row rank; then there is some nontrivial linear combination of the rows that sums to 0, or, equivalently, some nonzero $\vec{v} \in \mathbb{Q}^{|\mathscr{F}|}$, $\vec{v} = (\lambda_E)_{E \in \mathscr{F}}$, such that $\vec{v} I^*(\mathscr{F}, d) = \vec{0}$. We may assume that all λ_E are integers, and that, for some $F \in \mathscr{F}$, p does not divide λ_F.

For all $s \leqslant d$, we have $\vec{v} I(\mathscr{F}, s) = \vec{0}$, so $\vec{v} A_s(\mathscr{F}) = \vec{0}$. In particular, looking at the entry of $\vec{v} A_s(\mathscr{F})$ corresponding to F, we get that $\sum_{E \in \mathscr{F}} \lambda_E \binom{|E \cap F|}{s} = 0$ for any $s \leqslant d$.

Now, let $h(x) = h_{|F|}(x)$, and $r = \operatorname{val}(h(|F|))$. We can write $h(x) = \sum_{s=0}^{d} \gamma_s \binom{x}{s}$ where $\gamma_s \in \mathbb{Z}$. So, by the above paragraph, we must have

$$\sum_{E \in \mathscr{F}} \lambda_E h(|E \cap F|) = \sum_{E \in \mathscr{F}} \lambda_E \sum_{s=0}^{d} \gamma_s \binom{|E \cap F|}{s}$$

$$= \sum_{s=0}^{d} \gamma_s \sum_{E \in \mathscr{F}} \lambda_E \binom{|E \cap F|}{s} = 0.$$

We can thus write

$$\lambda_F h(|F|) = - \sum_{E \neq F} \lambda_E h(|E \cap F|),$$

and therefore

$$\operatorname{val}(\lambda_F h(|F|)) \geqslant \min_{E \neq F} \{\operatorname{val}(\lambda_E h(|E \cap F|))\}.$$

For any $E \neq F$, $\operatorname{val}(h(|E \cap F|)) > r$, so we have $\operatorname{val}(\lambda_F h_{|F|}(|F|)) > r$. But $\operatorname{val}(\lambda_F) = 0$, and $\operatorname{val}(h(|F|)) = r$, so this is impossible. This contradiction proves that no such nonzero $\vec{v}$ exists, and hence $\mathscr{F}$ is d^*-independent.

We say that a set system $\mathscr{F}$ is m-uniform if $|F| = m$ for every $F \in \mathscr{F}$. The above result gives us a slightly stronger version of Theorem 1.2 when $\mathscr{F}$ is m-uniform, thanks to a lemma of Frankl and Wilson:

LEMMA 8.2 [5]. *If $\mathscr{F}$ is m-uniform and s^*-independent for some $s<m$, then $\mathscr{F}$ is also s-independent.*

THEOREM 8.1. *Let $\mathscr{F}$ be an m-uniform set-system, $q=p^k$, and $L\subset\mathbb{N}_q$ with $|L|=s<\log m$. Suppose $m\notin L(\operatorname{mod} q)$, and $\mathscr{F}$ is L-intersecting mod q. Then*

$$|\mathscr{F}|\leqslant\binom{n}{2^{s-1}}.$$

Proof. By Lemma 5.1, there is a polynomial of degree $d\leqslant 2^{s-1}$ which separates m from $L+q\mathbb{Z}$. By Lemma 8.1, $\mathscr{F}$ is d^*-independent. Since $d<m$, we apply Lemma 8.1 and conclude that $\mathscr{F}$ is d-independent. Thus $|\mathscr{F}|\leqslant\binom{n}{d}$. ∎

9. THE CASE $L=\{0, ..., s-1\}$

For any s, let $c(s)$ be the least integer so that, for any prime power q and any $n\geqslant 2$, $m(n, s, q)\leqslant n^{c(s)}$. We have shown that $c(s)$ exists, and that $c(s)\leqslant 2^{s-1}$. However, the best known lower bound on $c(s)$ is $s+O(\sqrt{s})$ [6]. Is $c(s)$ polynomial in s? Is it linear in s?

When $L=\{0, ..., s-1\}$, we can improve our bound on $|\mathscr{F}|$ to $\binom{n}{\leqslant 2s}$.

LEMMA 9.1. *If $j<k$, and $0<a<p$, then $L_1=\{0, ..., ap^j-1\}$ is a box.*

Proof. Every congruence class modulo p^j is represented in L_1, so, for $0\leqslant i<j$, each node at level i has exactly p children. Each node at level j has exactly a children, and, for $i>j$, each node at level i has exactly one child. Thus, by definition, L_1 is a box. ∎

COROLLARY 9.1. *If $L=\{0, ..., s-1\}$, and $\mathscr{F}$ satisfies our standard assumptions, then $\mathscr{F}$ is $(2s)^*$-independent, and hence $|\mathscr{F}|\leqslant\binom{n}{\leqslant 2s}$.*

Proof. Choose the largest j such that $p^j<s$. Then either $s<p^{j+1}<s+p^j$, or $s<ap^j<s+p^j$ with $a<p$. In either case, there is a box of size at most $s+p^j<2s$ containing L. As in Lemma 5.1, we can construct the desired polynomials of degree less than $2s$; we then apply Lemma 8.1. ∎

If we add an additional uniformity condition, we can do even better:

THEOREM 9.1. *Let $L=\{0, ..., s-1\}$, and suppose $\mathscr{F}$ satisfies our standard assumptions. Suppose further that, for all $E \in \mathscr{F}$, $|E| \equiv s \pmod q$. Then $\mathscr{F}$ is s-independent, and hence $|\mathscr{F}| \leqslant \binom{n}{s}$.*

Proof. Consider the matrix $A_s(\mathscr{F})$. Any diagonal element is of the form $\binom{|E|}{s} = \binom{ap^k+s}{s}$ for some a. For each $j>0$, the number of multiples of p^j in the set $\{ap^k+1, ..., ap^k+s\}$ is exactly $\lfloor s/p^j \rfloor$, so $\mathrm{val}((ap^k+s)\cdots(ap^k+1)) = \mathrm{val}(s!)$. Thus, no diagonal element is a multiple of p.

Now, consider an off-diagonal element $\binom{|E \cap F|}{s} = \binom{ap^k+t}{s}$ for some a and some $t<s$. For each $j>0$, the number of multiples of p^j in the set $\{ap^k+j-s+1, ..., ap^k+j\}$ is at least $\lfloor s/p^j \rfloor$. Also, the number of multiples of p^k is 1, while $\lfloor s/p^k \rfloor = 0$. Thus, $\mathrm{val}((ap^k+j)\cdots(ap^k+j-s+1)) > \mathrm{val}(s!)$. We conclude that every off-diagonal element is a multiple of p.

We have shown that $A_s(\mathscr{F})$ is non-singular over $\mathbb{F}_p$, so it must be non-singular over $\mathbb{Q}$. This implies that $\mathscr{F}$ is s-independent. ∎

The above theorem slightly generalizes a result of Frankl and Wilson in [5], who prove this result for $s=p^k-1$.

10. OPEN QUESTIONS

Our work raises several open questions. The most striking of these is the large gap between our upper bounds and the best-known constructions.

We recall that, for $L \subset \mathbb{Z}$, a family $\mathscr{F}$ of sets is L-avoiding $\bmod p^k$ if $|E| \notin (L+p^k\mathbb{Z})$ for all $E \in \mathscr{F}$, and that $\mathscr{F}$ is L-intersecting $\bmod p^k$ if $|E \cap F| \in (L+p^k\mathbb{Z})$ for all $E, F \in \mathscr{F}$, $E \neq F$.

Question 1. Does there exist a constant c such that, for every prime power p^k, and every $L \subset \mathbb{Z}$, every set system $\mathscr{F}$ on n points which is L-intersecting mod p^k and L-avoiding mod p^k contains at most n^{cs} sets, where $s=|L|$?

In Theorem 1.1, we show that any such set system has size at most $n^{2^{s-1}}$. The largest known set system [6] has size roughly $n^{s+\sqrt{2s}}$.

Our work suggests sets L which might yield large families $\mathscr{F}$. If there is an integer α such that all sets $E \in \mathscr{F}$ have size $|E| \equiv \alpha \pmod{p^k}$, then the upper bound obtained by our methods is maximized when L satisfies the following condition: for distinct $\mu, \nu \in L$, if $p^i \mid (\mu-\nu)$, then $p^i \mid (\mu-\alpha)$. For example, we could take $\alpha=0$, and $L=\{1, p, p^2, ..., p^{k-1}\}$.

For the particular L above, we have $k=s$, so, by Theorem 6.2, we have $|\mathscr{F}| \leqslant n^{\exp(O(\sqrt{s \ln s}))}$.

Question 2. Does there exist a set system $\mathscr{F}$ on n points which is L-intersecting mod p^2 and L-avoiding mod p^2 and which contains more than n^{cs} sets, where $s = |L|$?

When the modulus is p^2, our methods show that $|\mathscr{F}|$ is at most n^D, where $D = \lfloor (1 + (s-1)/2)^2 \rfloor \approx s^2/4$. However, the largest known example is that mentioned above ([6]), of size roughly $n^{s+\sqrt{2s}}$. Can we construct a larger example? Can we reduce our upper bound to n^{cs}?

Again, our work suggests a set L which might yield a larger set system. If we assume $|E| \equiv 0 \pmod{p^2}$ for every $E \in \mathscr{F}$, then we need our set L to contain both multiples of p and non-multiples of p, and our upper bounds are maximized when L does not contain numbers μ and ν where $\mu \equiv \nu \not\equiv 0 \pmod{p}$. For example, we could take $L = \{1, 2, ..., p-1, p, 2p, ..., (p-1)p\}$. Sets in our system would have to intersect in one of two ways, one yielding a multiple of p, the other a set of size at most $p-1 \pmod{p^2}$.

We recall that $D(s, k)$ denotes the maximum, taken over all sets L with $|L| = s$, all primes p, and all $\alpha \notin L + p^k\mathbb{Z}$, of the minimum degree of a polynomial $f(x)$ such that, for some r, $p^r \mid f(\mu)$ for any $\mu \in L + p^k\mathbb{Z}$, but $p^r \nmid f(\alpha)$.

Question 3. Determine $D(s, s)$ asymptotically.

Question 4. How difficult is it to compute $D(s, k)$ precisely, or to estimate it within a constant factor?

The trivial approach, based on Definition 2.3, uses roughly $\binom{s+k}{k}$ steps. Ideally, we would wish to use only $(\log D(s, k))^{\text{constant}}$ number of steps.

ACKNOWLEDGMENTS

The authors gratefully acknowledge stimulating discussions with Vince Grolmusz and Tom Hayes.

REFERENCES

1. N. Alon, L. Babai, and H. Suzuki, Multilinear polynomials and Frankl–Ray-Chaudhuri–Wilson type intersection theorems, *J. Combin. Theory Ser. A* **58** (1991), 165–180.
2. L. Babai, On the non-uniform Ray-Chaudhuri–Wilson inequality, *Combinatorica* **8** (1988), 133–135.
3. L. Babai and P. Frankl, "Linear Algebra Methods in Combinatorics, with Applications to Geometry and Computer Science," preliminary version 2, Univ. of Chicago, Chicago, 1992.
4. M. Deza, P. Frankl, and N. M. Singhi, On functions of strength t, *Combinatorica* **3** (1981), 331–339.

5. P. Frankl and R. M. Wilson, Intersection theorems with geometric consequences, *Combinatorica* **1** (1981), 357–368.
6. Peter Frankl, Constructing finite sets with given intersections, *in* "Combinatorial Mathematics (Marseille–Luminy, 1981)," Annals of Discrete Mathematics, Vol. 17, pp. 289–291, North-Holland, Amsterdam, 1983.
7. V. Grolmusz, Superpolynomial size set-systems with restricted intersections mod 6 and explicit Ramsey graphs, *Combinatorica* **20**, No. 1 (2000), 71–86.
8. G. H. Hardy, J. E. Littlewood, and G. Pólya, "Inequalities," 2nd ed., Cambridge University Press, Cambridge, UK, 1952.
9. D. K. Ray-Chaudhuri and R. M. Wilson, On t-designs, *Osaka J. Math.* **12** (1975), 737–744.
10. Thomas A. Standish, "Data Structure Techniques," Addison–Wesley, Reading, MA, 1980.