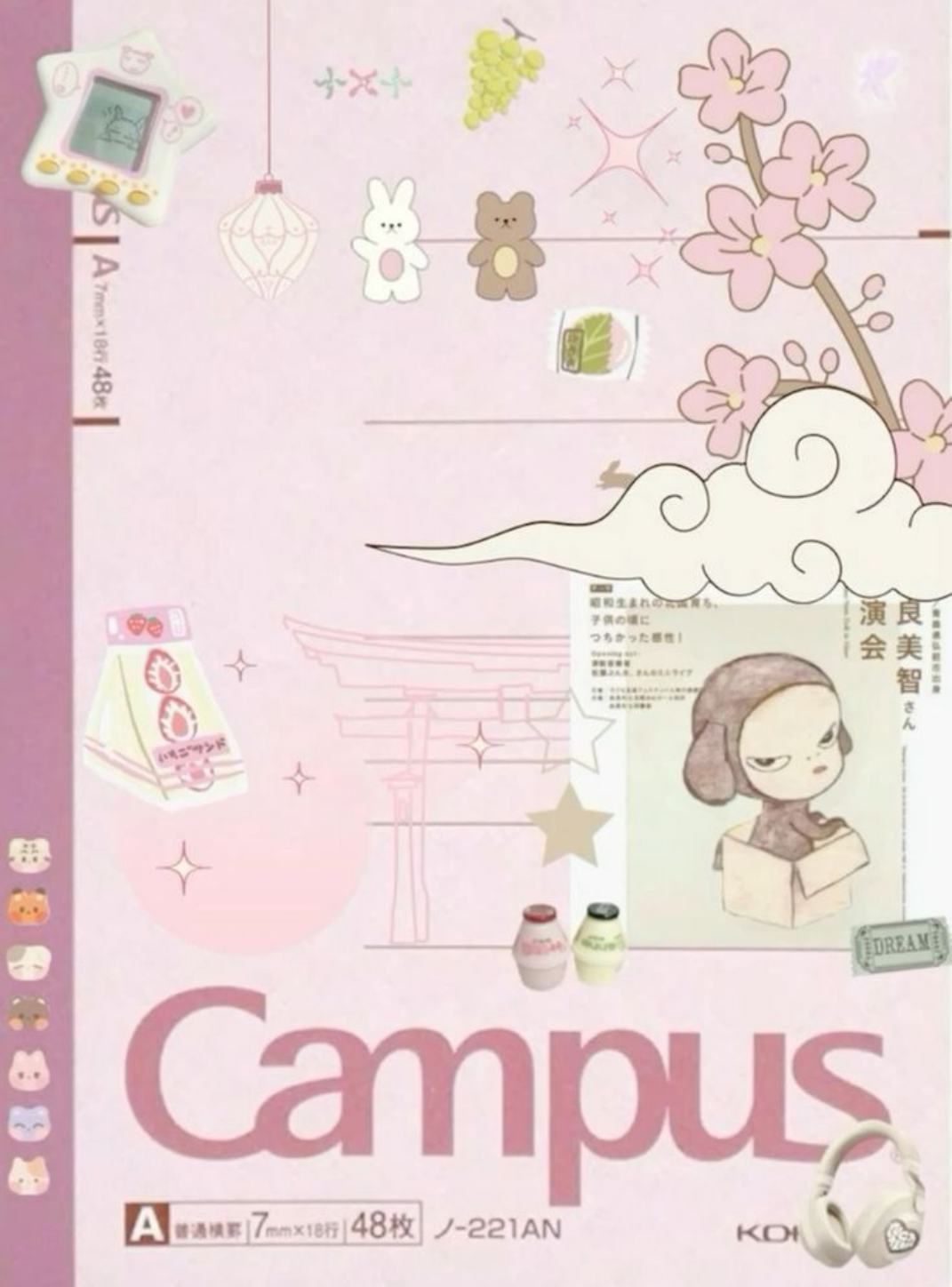

A 7mm×18行 48枚
良美智さん
演会
DREAM
Campus
A 普通横罫 7mm×18行 48枚 ノ-221AN

Goal: read 2024 paper - Improved PIR Schemes using Matching Vectors and Derivatives

Papers

2008

- well-known improvements to PIR
- locally-decodable codes
- Based on 1996 and 1994

Babai textbook (Frankl Wilson theorem)

Timeline

plan / theorem list by week 4/5

general understanding by week 7

presentation week 9

paper end of quarter.

What is an LDC?

* n = message length
N = codeword length

3-query LDC of sub-exponential length of size $\underbrace{\exp(\exp(O(\frac{\ln n}{\ln(\ln n)})))}_{N,\ \text{length of codeword}}$ (Yekhanin)

we can recover a bit by sampling 3 spots (3 queries) multiple times

e.x. Hadamard 2-query LDC

$x \in \{0,1\}^k$ and codeword is $C_x(y) = x \cdot y \bmod 2 \quad \forall y \in \{0,1\}^k$ (not always)

$x = [1,0]$:

y	x·y	$C_x(y)$
(0,0)	0	0
(0,1)	0	0
(1,0)	1	1
(1,1)	1	1

Codeword is $[0,0,1,1]$ (Depends on constant order of y)

Suppose codeword gets slightly corrupted $[0,0,1,1] \rightarrow [0,0,1,0]$

we want only 1 bit of message (say x_i) without decoding entire message

Process:

$y \in \{0,1\}^2$ and query: $C_x(y)$, $C_x(y+e_i)$ (unit want to recover)

$C_x(y) \oplus C_x(y+e_i) = x_i$ (mod 2 or $\mathbb{F}_2$)

(0,0) (0,1) (1,0) (1,1)

$[0, 0, 1, 1]$

$y = (0,0)$, $C_x(y) = 0$

$y + e_i = (1,0)$, $C_x(y+e_i) = 1$ $\Big\} > 1$

$y = (1,1)$ $C_x(y) = 0$

$y + e_i = (0,1)$ $C_x(y+e_i) = 0$ $\Big\} > 0$

wrong but after all 2-bit queries, this will be in the minority.

Ray-Chaudhuri – Wilson Theorem: let L be a set of s integers and F an L-intersecting k-uniform family of subsets of a set of n elements where $s \leq k$. Then $|F| \leq \binom{n}{s}$

e.x. $\{1,2,3,4,5,6\}$

$L = \{2,3\}$ $k = 3$, $n = 6$, $s = 2$

$|F| \leq \frac{6!}{2!4!} = \frac{6 \cdot 5}{2} = 15$

Frankl-Wilson Theorem: let p be a prime # and L a set of s distinct residues mod p, let k be an integer, $k \bmod p \notin L$

Assume $s + k \leq n$. let F be a family of subsets of a set of n elements s.t.

1) $|E| \equiv k \bmod p$ for $E \in F$
2) $|E \cap F| \in L \bmod p$ for $E, F \in F$, $E \neq F$

# of uniform (mod p) sets that intersect within the rules are less than or equal to $\binom{n}{s}$

Then $|F| \leq \binom{n}{s}$

$\binom{5}{2} = \frac{5 \cdot 4}{2} = 10$

e.x. $\{1,2,3,4,5\}$ $n = 5$, $p = 3$

$L = \{1,0\}$ $k = 2$, $s = 2$

$F = \{ \{1,2\}, \{2,3\}, \{2,3\}, \{1,3\} \}$ (1, 2, 3, 4) $|F| \leq \binom{5}{2}$ ✓

2008 - 3-Query Locally Decodable Codes of Subexponential length by Klim Efremenko

1) Built off Yekhanin who created 3-query LDC, assuming ∃ an infinite number of mersenne primes

2) Efremenko created a 3-query that worked no matter what and was shorter: $\exp(\exp(\sqrt{\log n \cdot \log\log n}))$

3) Uses existance of super-polynomial size set-systems with restricted intersections by Grolmusz

1996 - Superpolynomial size set-systems with restricted intersections mod 6 and ramsey graphs

constructed:
- System $\mathcal{H}$ $|\mathcal{H}| = \exp(c \cdot \log^2(n) / \log(\log(n)))$ and $\mathcal{H} \subseteq P(A)$ where $|A| = n$
- and $\forall h_i, h_j \in \mathcal{H}$, $|h_i| \equiv_6 0$ but $|h_i \cap h_j| \not\equiv_6 0$

=> results generalize this to all non-prime-power m (6, 10, 15)
and negatively answers Frankl-Wilson (1981)

1994 - Represented Boolean functions as polynomials modulo composite numbers by BBR (over $\mathbb{Z}_m$)

MOD_m-degree of boolean function F is the smallest degree of any polynomial s.t. $F(\vec{x}) = 0$ iff $P(\vec{x}) = 0$

1) MOD-degree of strong OR of N variables is always N

weak representation

2) MOD_m-degree of OR of N variables $\leq O(\sqrt[r-1]{N})$ (r = prime factors of m)
↳ lower bound of $\Omega(\log N^{\frac{1}{r-1}})$

3) MOD_{p^e}-degree of $OR_N \geq \frac{N}{p^e - 1}$

weak rep in $\mathbb{Z}_{p^e}$ of degree d => strong rep in $\mathbb{Z}_p$ of degree = n $\leq d(p^e - 1)$

Tardos & Barrington — 1994 Supplement

Strong Representation: $f(x) = P(x)\ \forall x$, $P(x) = \{0,1\}$

One sided representation: $f(x) = 0$ iff $P(x) = 0$

↳ $f(x) = 1$ if $P(x) =$ anything

Weak representation: $\forall x,y\ P(x) \neq P(y)$ whenever $f(x) \neq f(y)$

↳ $\exists S \subseteq \mathbb{Z}_m$ s.t. $f(x) = 0$ iff $P(x) \in S$

Lemma 1: Let $q = p^e$ be a prime power and let P be a polynomial of degree d in n boolean variables over $\mathbb{Z}_q$. If P weakly represents a boolean f, then $\exists$ a polynomial P^* over $\mathbb{Z}_p$ of degree at most $(q-1)d$ strongly representing f.

Going from weak => one-sided without effect on degree (d)

One-sided in $\mathbb{Z}_{p^e}$ => strong in $\mathbb{Z}_p$ ($n \leq d \cdot (p^e - 1)$)

=> degree of weak/one-sided in $\mathbb{Z}_{p^e}$ ($d \geq \frac{n}{p^e - 1}$, b/c $\frac{n}{p^e-1} \cdot p^e - 1 = n$)

Corollary 2: If q is prime power p^e and $d(q-1) < n$, no polynomial of d can weakly represent OR over $\mathbb{Z}_p$

If we can, then $\exists$ $d(q-1)$ rep over $\mathbb{Z}_p$ which is less than n (contradiction), Strong OR_n is always degree n.

$d(p^e-1) \geq n$, for d = degree over $\mathbb{Z}_{p^e}$ degree $(\mathbb{Z}_{p^e}) \geq \frac{n}{p^e-1}$

Coro: let $e=1$ then $d(p-1) \geq n$ for d = degree over $\mathbb{Z}_p$

at minimum $p=2$ => $d \geq \frac{n}{(1)}$

COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA **20** (1) (2000) 71–85

SUPERPOLYNOMIAL SIZE SET-SYSTEMS WITH RESTRICTED INTERSECTIONS MOD 6 AND EXPLICIT RAMSEY GRAPHS

VINCE GROLMUSZ*

Received January 15, 1996
Revised August 2, 1999

Dedicated to the memory of Paul Erdős

We construct a system $\mathcal{H}$ of $\exp(c\log^2 n/\log\log n)$ subsets of a set of n elements such that the size of each set is divisible by 6 but their pairwise intersections are not divisible by 6. The result generalizes to all non-prime-power moduli m in place of $m=6$. This result is in sharp contrast with results of *Frankl* and *Wilson* (1981) for prime power moduli and gives strong negative answers to questions by *Frankl* and *Wilson* (1981) and *Babai* and *Frankl* (1992). We use our set-system $\mathcal{H}$ to give an explicit Ramsey-graph construction, reproducing the logarithmic order of magnitude of the best previously known construction due to *Frankl* and *Wilson* (1981). Our construction uses certain mod m polynomials, discovered by *Barrington*, *Beigel* and *Rudich* (1994).

1. Introduction

Generalizing the *Ray-Chaudhuri–Wilson* theorem [11], *Frankl* and *Wilson* [9] proved the following intersection theorem, one of the most important results in extremal set theory:

Theorem 1.1 (Frankl–Wilson). *Let $\mathcal{F}$ be a set-system over a universe of n elements. Suppose $\mu_0, \mu_1, \ldots, \mu_s$ are distinct residues modulo a prime p,*

* Part of this research was done while the author was visiting the Department of Computer Science at The University of Chicago.

Mathematics Subject Classification (1991): 05D05, 05D10, 68Q25

such that for all $F \in \mathcal{F}$,

$$|F| = k \equiv \mu_0 \pmod{p},$$

where $k+s \le n$, *and for any two distinct* $F, G \in \mathcal{F}$:

$$|F \cap G| \equiv \mu_i \pmod{p} \text{ for some } i,\ 1 \le i \le s.$$

Then

$$|\mathcal{F}| \le \binom{n}{s}. \tag{1}$$

This theorem has numerous applications in combinatorics and in geometry (e.g., the disproof of *Borsuk's conjecture* by *Kahn* and *Kalai* [10] (cf. [2], Sec. 5.6.), an explicit construction of Ramsey graphs, and geometric applications related to the Hadwiger-problem [9].)

Frankl and *Wilson* [9] asked whether inequality (1) remains true when the modulus p is replaced by a composite number m, or at least in the subcase $s = m-1$.

Frankl [8] answered the first of these questions (arbitrary $s \le m$) in the negative: he constructed faster growing set-systems for $m=6$, as well as for $m=p^2$, p prime. For $m=6$, *Frankl's* set-systems satisfy $s=3$ and $|\mathcal{F}| \approx cn^4$.

On the other hand, *Frankl* and *Wilson* [9] proved that inequality (1) remains in force when $s=m-1$ and m is a prime power.

In this paper we consider non-prime-power moduli m. For any such modulus, we give a very strong negative answer to both versions of the Frankl–Wilson question: we prove that for $s=m-1$, no upper bound of the form $n^{f(m)}$ exists. More precisely, we prove the following.

Theorem 1.2. *Let* m *be a positive integer, and suppose that* m *has* $r>1$ *different prime divisors:* $m = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r}$. *Then there exists* $c = c(m) > 0$, *such that for every integer* $h > 0$, *there exists an explicitly constructible uniform set-system* $\mathcal{H}$ *over a universe of* h *elements, such that*

(a) $|\mathcal{H}| \ge \exp\left(c \frac{(\log h)^r}{(\log\log h)^{r-1}}\right)$,

(b) $\forall H \in \mathcal{H}$: $|H| \equiv 0 \pmod{m}$,

(c) $\forall G, H \in \mathcal{H}, G \ne H$: $|G \cap H| \not\equiv 0 \pmod{m}$.

Remark 1.1. The value of c is roughly p_r^{-r}, where p_r is the largest prime divisor of m. The size of the sets in the set-system we construct is

$$|H| = h^{\frac{r-1}{2r-1} + o(1)}. \tag{2}$$

We note that for fixed m (m is not a prime power), the size of $\mathcal{H}$ grows faster than any polynomial of n. This is quite surprising, since previously it was believed that the failure of the attempts to prove a polynomial upper bound was due to the lack of techniques to handle non-prime-power composite moduli.

Our result gives a strong negative answer to a conjecture of *Babai* and *Frankl* ([2], Section 7.3, Conjecture C(r)). *Babai* and *Frankl* conjectured that conditions (b) and (c) of Theorem 1.2 imply

$$|\mathcal{H}| \le \binom{h}{m-1};$$

whereas our result shows that no bound of the form $h^{f(m)}$ exists for composite, non-prime power moduli m.

We can even strengthen statement (c) of Theorem 1.2 as follows:

Theorem 1.3. *Theorem 1.2 remains valid if we add the following condition:*

(d) $\forall G, H \in \mathcal{H}$, $G \neq H$ *and* $\forall i \in \{1, 2, \ldots, r\}$, *we have* $|G \cap H| \equiv 0 \pmod{p_i^{\alpha_i}}$ *or* $|G \cap H| \equiv 1 \pmod{p_i^{\alpha_i}}$.

Remark 1.2. *Theorem 1.3 implies that there exist super-polynomial size set-systems $\mathcal{H}$ such that the size of each set in $\mathcal{H}$ is divisible by m and the sizes of the pairwise intersections of the sets in $\mathcal{H}$ occupy at most $2^r - 1$ residue classes mod m out of the possible $m-1$ nonzero residue classes.*

In fact, this result can be further strengthened: 3 residue classes of intersection size suffice! This answers a question of Peter Frankl (private communication).

Corollary 1.1. *Let m be a positive integer, and suppose that m has $r > 1$ different prime divisors: $m = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r}$. Then there exists $c = c(m) > 0$, such that for every integer $h > 0$, there exists an explicitly constructible uniform set-system $\mathcal{H}$ over a universe of h elements such that*

(a) $|\mathcal{H}| \ge \exp\left(c \frac{(\log h)^2}{\log \log h}\right)$,

(b) $\forall H \in \mathcal{H}$: $|H| \equiv 0 \pmod{m}$,

(c) *the sizes of the pairwise intersections $|G \cap H|$ $(G, H \in \mathcal{H}, G \neq H)$ occupy only 3 residue classes mod m, none of which is 0.*

One of the striking applications of the Frankl–Wilson theorem for prime moduli was an explicit construction of graphs of size $\exp(c \log^2 n / \log \log n)$ without homogeneous subsets (cliques or anti-cliques) of size n. These are the

largest explicit Ramsey-graphs known to-date. As an application of our Theorem 1.2, we give an alternative construction of explicit Ramsey graphs of the same logarithmic order of magnitude, *i.e.*, of size $\exp(c' \log^2 n / \log\log n)$. (But our c' is less than their c).

A key ingredient of our construction is a low-degree polynomial constructed by *Barrington*, *Beigel* and *Rudich* [5], to represent the Boolean "OR" function mod m. Any reduction of the degree of such polynomials would yield improved explicit Ramsey graphs.

2. Preliminaries

Let $f:\{0,1\}^n \to \{0,1\}$ be a Boolean function and let m be a positive integer. *Barrington*, *Beigel* and *Rudich* [5] gave the following definition:

Definition 2.1. The polynomial P with integer coefficients *weakly represents* the Boolean function f modulo m if there exists an $S \subset \{0,1,2,\ldots,m-1\}$ such that for all $x \in \{0,1\}^n$,

$$f(x) = 0 \iff (P(x) \bmod m) \in S.$$

Here $(a \bmod m)$ denotes the smallest non-negative $b \equiv a \bmod m$.

We are interested in the smallest degree of polynomials representing f modulo m. Without loss of generality we may assume P is multilinear (since $x_i^2 = x_i$ over $\{0,1\}^n$).
Let $\mathrm{OR}_n:\{0,1\}^n \to \{0,1\}$ denote the n-variable OR-function:

$$\mathrm{OR}_n(x_1, x_2, \ldots, x_n) = \begin{cases} 0, & \text{if } x_1 = x_2 = \cdots = x_n = 0 \\ 1 & \text{otherwise.} \end{cases}$$

Suppose that the polynomial P weakly represents OR_n modulo a prime p. Without loss of the generality we may assume that for $x \in \{0,1\}^n$,

$$P(x) \equiv 0 \bmod p \iff x = (0,0,\ldots,0).$$

Then

$$f(x) = 1 - P^{p-1}(1 - x_1, 1 - x_2, \ldots, 1 - x_n)$$

is exactly the n-variable AND function, which can uniquely be written as a multilinear monomial

$$x_1 x_2 x_3 \ldots x_n.$$

Consequently, if the polynomial P weakly represents OR_n over $GF(p)$, then its degree is at least

$$\left\lceil \frac{n}{p-1} \right\rceil.$$

Tardos and *Barrington* [12] proved that the same conclusion holds if p is a prime power.

On the other hand, *Barrington, Beigel* and *Rudich* [5] proved that the conclusion fails for composite moduli with at least two distinct prime divisors:

Theorem 2.4 (Barrington, Beigel, Rudich). *Given $m = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r}$ where the p_i are distinct primes, there exists an explicitly constructible polynomial P of degree $O(n^{1/r})$ which weakly represents OR_n modulo m.*

For completeness, we reproduce here a short proof of this theorem.

Proof. Let $S_k(x)$ denote the k^{th} elementary symmetric polynomial, *i.e.* the sum of all multilinear monomials of degree k, formed from variables $x_1, x_2, \ldots, x_n$. For $x \in \{0,1\}^n$, the *weight* of x is defined as $\mathrm{wt}(x) = \sum_{i=1}^{n} x_i$. If $\mathrm{wt}(x) = \ell$, then

$$s_k(x) = \binom{\ell}{k}.$$

Since the value of $s_k(x)$ depends only on $\mathrm{wt}(x)$, with some abuse of the notation we shall write $s_k(x)$ as $s_k(j)$ where $j = \mathrm{wt}(x)$. Using this notation, one can formulate the following observation made in [5]:

Lemma 2.1. *[5] Let k be a positive integer, p be a prime and let e be the smallest integer satisfying $k < p^e$. Then $s_k(j) \equiv s_k(j + p^e) \pmod{p}$.*

Proof. We need to prove

$$\binom{j + p^e}{k} \equiv \binom{j}{k} \pmod{p}.$$

This is immediate from the identity

$$\binom{u+v}{t} = \sum_{w=0}^{t} \binom{u}{w} \binom{v}{t-w},$$

and the elementary fact that for any $1 \le \ell < p^e$, p is a divisor of $\binom{p^e}{\ell}$. ∎

Now, for $i = 1, 2, \ldots, r$, let e_i be the smallest integer that satisfies

$$p_i^{e_i} > \lceil n^{1/r} \rceil.$$

We define, for $i = 1, 2, \ldots, r$, the symmetric polynomial $G_i(x)$ by

$$G_i(x) = \sum_{j=1}^{p_i^{e_i} - 1} (-1)^{j+1} s_j(x).$$

One can easily prove (using the binomial expansion of $(1-1)^{p_i^{e_i}-1}$), that G_i correctly computes over the integers the OR function for inputs of weight at most $p_i^{e_i}-1$. Consequently, G_i correctly computes modulo p_i the OR function for inputs of weight at most $n^{1/r}$, and, additionally, $G_i \bmod p_i$ is periodic with period $p_i^{e_i}$.

And now, by the Chinese Remainder Theorem, there exists a polynomial P which satisfies

$$P \equiv G_i \pmod{p_i}$$

for $i = 1,2,\ldots,r$, and the degree of P is the maximum of the degrees of polynomials G_i, $O(n^{1/r})$.

It is obvious that for $x \in \{0,1\}^n$, if $\mathrm{wt}(x) \neq 0$ then there exists an i, $1 \leq i \leq r$, such that $\mathrm{wt}(x) \not\equiv 0 \bmod p_i^{e_i}$, so $P(x) \not\equiv 0 \bmod p_1 p_2 \ldots p_r$. In addition, $P(0,0,\ldots,0) = 0$. Consequently, P weakly represents the OR function for all inputs in $\{0,1\}^n$ modulo $p_1 p_2 \ldots p_r$. Since $p_1 p_2 \ldots p_r$ is a divisor of m, if $P(x)$ is not 0 modulo $p_1 p_2 \ldots p_r$ then it is not 0 modulo m. Consequently, P weakly represents the OR function for all inputs in $\{0,1\}^n$ modulo m. ∎

Example. Let $m = 6$, and let

$$G_1(x) = \sum_{j=1}^{2^3-1} (-1)^{j+1} s_j(x),$$

and

$$G_2(x) = \sum_{j=1}^{3^2-1} (-1)^{j+1} s_j(x).$$

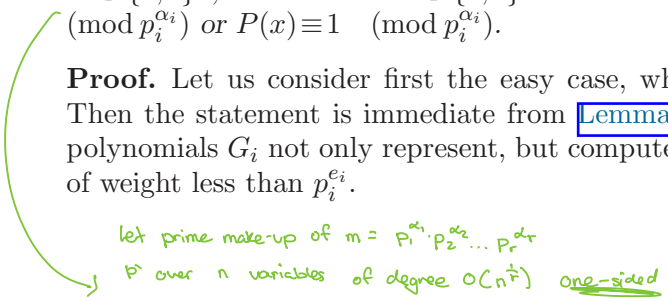
Then

$$P(x) = 3G_1(x) + 4G_2(x)$$

weakly represents OR_{71} modulo 6 (or modulo 6ℓ for any integer ℓ), and its degree is only 8.

Corollary 2.2. *Let $m = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r}$. Then there exists an explicitly constructible polynomial P' with n variables and of degree $O(n^{1/r})$ which is equal to 0 on $x = (0,0,\ldots,0) \in \{0,1\}^n$, it is nonzero mod m for all other $x \in \{0,1\}^n$, and for all $x \in \{0,1\}^n$ and for all $i \in \{1,\ldots,r\}$, $P(x) \equiv 0 \pmod{p_i^{\alpha_i}}$ or $P(x) \equiv 1 \pmod{p_i^{\alpha_i}}$.*

Proof. Let us consider first the easy case, when $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 1$. Then the statement is immediate from Lemma 2.1 and from the fact that polynomials G_i not only represent, but compute the OR function for inputs of weight less than $p_i^{e_i}$.



In the general case, let us observe that G_i is either 0 or 1 modulo p_i on $\{0,1\}^n$. Then we need the modulus-amplifying polynomials R_i of degree $2\alpha_i$ of *Beigel* and *Tarui* [3], with the following properties:

$$N \equiv 0 \pmod{p_i} \Longrightarrow R_i(N) \equiv 0 \pmod{p_i^{\alpha_i}}$$

and

$$N \equiv 1 \pmod{p_i} \Longrightarrow R_i(N) \equiv 1 \pmod{p_i^{\alpha_i}}.$$

Now, set $G_i' = R_i \circ G_i$ and construct P' by applying the Chinese Remainder Theorem to the G_i'. ∎

3. The Lower Bound

proof that size of HEZL is sublinear in terms of n→∞. elements available

Proof of Theorem 1.2

Let $P(z_1, z_2, \ldots, z_n)$ be a polynomial of degree d which satisfies that $P(0,0,0,\ldots,0) = 0$, and for every $(z_1, z_2, \ldots, z_n) \in \{0,1\}^n$

$$P(z_1, z_2, \ldots, z_n) \equiv 0 \pmod{m} \Longleftrightarrow z_1 = z_2 = \ldots = z_n = 0.$$

An explicit construction of such P of degree $d = O(n^{1/r})$ was given in Theorem 2.4.
Let $Q(z_1, z_2, \ldots, z_n) = P(1 - z_1, 1 - z_2, \ldots, 1 - z_n)$. Then $Q(1,1,1,\ldots,1) = 0$, and for all $z \in \{0,1\}^n$ we have

(3) $$Q(z) \equiv 0 \pmod{m} \Longleftrightarrow z_1 = z_2 = \ldots = z_n = 1.$$

using property we construct AND function Q(x)

Using the polynomial Q we state our main Lemma:

Lemma 3.2. *For every integer $n > 0$, there exists a uniform set-system $\mathcal{H}$ over a universe of $2(m-1)n^{2d}/d!$ elements which is explicitly constructible from the polynomial Q and satisfies*

(a) $|\mathcal{H}| = n^n$,
(b) $\forall H \in \mathcal{H}\colon |H| \equiv 0 \pmod{m}$,
(c) $\forall G, H \in \mathcal{H}, G \neq H\colon |G \cap H| \not\equiv 0 \pmod{m}$.

Lemma 3.2 easily yields Theorem 1.2 setting $d = \Theta(n^{1/r})$ and using elementary estimations for the binomial coefficients.

Proof of Lemma 3.2. Q can be written as

$$Q(z_1, z_2, \ldots, z_n) = \sum_{i_1, i_2, \ldots, i_\ell} a_{i_1, i_2, \ldots, i_\ell} z_{i_1} z_{i_2} \ldots z_{i_\ell},$$

where $0 \le \ell \le d$, and $a_{i_1,i_2,\ldots,i_\ell}$ is integer, $1 \le i_1 < i_2 < \cdots < i_\ell \le n$. Let us define

$$\tilde{Q}(z_1, z_2, \ldots, z_n) = \sum_{i_1,i_2,\ldots,i_\ell} \tilde{a}_{i_1,i_2,\ldots,i_\ell} z_{i_1} z_{i_2} \ldots z_{i_\ell}, \tag{4}$$

where $\tilde{a}_{i_1,i_2,\ldots,i_\ell} = (a_{i_1,i_2,\ldots,i_\ell} \bmod m)$ is the smallest, positive integer, congruent to $a_{i_1,i_2,\ldots,i_\ell}$ modulo m, for $1 \le i_1 < i_2 < \cdots < i_\ell \le n$.

We note, that (3) is satisfied for $\tilde{Q}$, but $\tilde{Q}(1,1,1,\ldots,1)$ is not necessarily 0.

Let the function $\delta : \{0,1,\ldots,n-1\} \times \{0,1,\ldots,n-1\} \to \{0,1\}$ be defined as

$$\delta(u,v) = \begin{cases} 1, & \text{if } u = v, \\ 0 & \text{otherwise.} \end{cases}$$

Let $A = (a_{xy})$ be an $n^n \times n^n$ matrix $(x, y \in \{0,1,2,\ldots,n-1\}^n)$.

We define the entry a_{xy} as follows:

$$a_{xy} = \tilde{Q}(\delta(x_1,y_1), \delta(x_2,y_2), \ldots, \delta(x_n,y_n)) \bmod m. \tag{5}$$

We note that $a_{xx} = \tilde{Q}(1,1,\ldots,1) \equiv 0 \pmod m$. Conversely, if $a_{xy} \equiv 0 \pmod m$ then $x = y$.

By equation (4), the polynomial $\tilde{Q}(z)$ is a sum of the monomials of the form $z_{i_1} z_{i_2} \ldots z_{i_\ell}$ $(\ell \le d)$. We wish to keep all coeffcients equal to 1; therefore we shall say that the monomial $z_{i_1} z_{i_2} \ldots z_{i_\ell}$ $(\ell \le d)$ occurs with *multiplicity* $\tilde{a}_{i_1,i_2,\ldots,i_\ell}$ in this sum. Note that each multiplicity is a nonnegative integer $\le m-1$.

Consequently, the matrix A is a sum of the matrices $B_{i_1,i_2,\ldots,i_\ell} = (b_{x,y}^{i_1,i_2,\ldots,i_\ell})$, corresponding to the monomial $z_{i_1} z_{i_2} \ldots z_{i_\ell}$ in the following way:

$$b_{x,y}^{i_1,i_2,\ldots,i_\ell} = \delta(x_{i_1}, y_{i_1}) \delta(x_{i_2}, y_{i_2}) \ldots \delta(x_{i_\ell}, y_{i_\ell}).$$

This matrix occurs in the sum with multiplicity $\tilde{a}_{i_1,i_2,\ldots,i_\ell}$.

It is easy to verify that $B_{i_1,i_2,\ldots,i_\ell}$ is permutationally equivalent to the matrix

$$\begin{pmatrix} J_1 & & & 0 \\ & \ddots & & \\ 0 & & & \\ & & & J_{n^\ell} \end{pmatrix} \tag{6}$$

where the diagonal blocks J_i are all-ones matrices of size $n^{n-\ell} \times n^{n-\ell}$, and there are exactly n^ℓ pairwise disjoint all-ones blocks in $B_{i_1,i_2,\ldots,i_\ell}$. "Permutationally equivalent" means that there exists a permutation such that if it

is applied both to the rows and to the columns of the matrix, the result is equal to (6). Let us refer to these all-ones blocks of $B_{i_1,i_2,\ldots,i_\ell}$ as *B-blocks*. We shall say that each B-block of $B_{i_1,i_2,\ldots,i_\ell}$ occurs with multiplicity $\tilde{a}_{i_1,i_2,\ldots,i_\ell}$.

By equation (4), A can be written in the following form:

$$A = \sum_{i_1,i_2,\ldots,i_\ell} \tilde{a}_{i_1,i_2,\ldots,i_\ell} B_{i_1,i_2,\ldots,i_\ell}. \tag{7}$$

Lemma 3.3. *Taking multiplicities into account,*

(a) *every cell of the main diagonal of A is covered by the same number of B-blocks, and this number is divisible by m;*
(b) *for any pair of cells of the main diagonal of A, the number of those B-blocks which cover both members of the pair, is not divisible by m.*

Proof. We note that the number of B-blocks covering cell (x,y) is a_{xy}. Now statement (a) follows by equation (3), observing that for all x,

$$a_{xx} = \tilde{Q}(1,1,\ldots,1) \equiv 0 \pmod{m}.$$

For part (b), we note that the B-blocks are square submatrices, symmetric to the diagonal; therefore a B-block covers the cells (x,x) and (y,y) exactly if it covers the cell (x,y). The number of B-blocks covering both (x,x) and (y,y) is therefore $a_{xy} \not\equiv 0 \pmod{m}$, again by equation (3). ∎

Corollary 3.3. *There exists an explicitly constructible hypergraph $\mathcal{G}$ with n^n vertices and fewer than $2(m-1)n^{2d}/d!$ edges, such that every vertex is contained in the same number of edges, and this number is divisible by m; while for any two vertices, the number of edges, containing both of the vertices, is not divisible by m. (We allow multiple edges and take multiplicities into account.)*

Proof. From Lemma 3.3, choose the cells of the diagonal of A for the vertices and the intersections of the B-blocks with the diagonal for edges (with the corresponding multiplicity).

The number of edges is

$$h := \tilde{Q}(n,n,\ldots,n) = \sum_{\ell \le d} \sum \tilde{a}_{i_1,i_2,\ldots,i_\ell} n^\ell \le (m-1) \sum_{\ell \le d} \binom{n}{\ell} n^\ell$$

$$< (m-1) \sum_{\ell \le d} n^{2\ell}/\ell! < 2(m-1)n^{2d}/d!,$$

assuming, as we may, that $n \ge 2d$. ∎

We note that the number of edges containing each vertex is

$$\tilde{Q}(1,1,\ldots,1) \leq (m-1)\left(\binom{n}{d}+\binom{n}{d-1}+\ldots\binom{n}{0}\right) < 2(m-1)\binom{n}{d}.$$

Now we are ready to complete the proof of Lemma 3.2.

Let us consider the dual of the hypergraph of Corollary 3.3, i. e., let the universe be the set of B-blocks, and if a B-block was present a times in the hypergraph $\mathcal{G}$, then it will correspond to a different points (or elements) in the universe. Consequently, our universe is a set (rather than a multiset). The size of the universe is $h<2(m-1)n^{2d}/d!$.

The diagonal cells of A correspond to the members of the set-system $\mathcal{H}$: the set corresponding to cell (x,x) consists of exactly those B-blocks which cover (x,x). Therefore $|\mathcal{H}|=n^n$.

Since every diagonal cell of A is covered by the same number of B-blocks, the resulting $\mathcal{H}$ is a uniform set system. As discussed previously, this number (the size of the members of $\mathcal{H}$) is $\tilde{Q}(1,1,\ldots,1)\leq(m-1)\sum_{\ell=0}^{d}\binom{n}{d}<2(m-1)\binom{n}{d}$.

From Corollary 3.3, statements (a), (b), (c) of Lemma 3.2 follow. ∎

Remark 3.3. *We note from the foregoing that the number of vertices of $\mathcal{H}$ is $h:=\tilde{Q}(n,n,\ldots,n)$, and the number of vertices of each member of $\mathcal{H}$ is $\tilde{Q}(1,1,\ldots,1)$. We note that $\tilde{Q}(n,n,\ldots,n)\leq n^d\tilde{Q}(1,1,\ldots,1)$.*

To prove the estimate on the size of the members of $\mathcal{H}$ in terms of h (the number of vertices of $\mathcal{H}$) given in Remark 1.1, we first add dummy vertices to increase h to its upper bound $h':=n^d\tilde{Q}(1,1,\ldots,1)$ stated above. Now, since this quantity is still $\leq 2(m-1)n^{2d}/d!$, we see, using the bound $d=O(n^{1/r})$ guaranteed by Theorem 2.4, that

$$n^d \geq (h')^{\frac{r}{2r-1}+o(1)}$$

and therefore the size of the members of $\mathcal{H}$ is

$$\tilde{Q}(1,1,\ldots,1) \leq (h')^{\frac{r-1}{2r-1}+o(1)},$$

as claimed in equation (2). ∎

Proof of Theorem 1.3. The statement is immediate if the polynomial P' of Corollary 2.2 is used for the construction of the set-system $\mathcal{H}$ in the proof of Theorem 1.2 in the place of the polynomial P. ∎

Proof of Corollary 1.1. Let $m' = p_1^{\alpha_1} p_2^{\alpha_2}$, and apply Theorem 1.3 for constructing a set-system $\mathcal{H}$ for h and this m'. The intersections occupy only 3 residue classes modulo m'. Now replace every point of the universe by m/m' new points; the new points will be the members of exactly the same sets of the set-system as the old point. The statement follows. ∎

4. An Application: Ramsey Graphs

The set-system $\mathcal{H}$ of Theorem 1.2 yields new families of explicit Ramsey-graphs.

Theorem 4.5 (Frankl–Wilson, 1981). *For $t \geq 3$, there exists an explicitly constructible graph on $\exp\left(c\frac{(\log t)^2}{(\log\log t)}\right)$ vertices which does not contain either a complete graph or an independent set of size t.*

The constant c given in [9] is $c = \frac{1}{4}$. Our construction yields $c = \frac{2}{81}$ only. In addition to giving a novel proof of Theorem 4.5, we extend it to the case of several colors:

Theorem 4.6. *For $r \geq 2$, $t \geq 3$, there exists an explicitly constructible r-coloring of the edges of the complete graph on $\exp\left(c_r \frac{(\log t)^r}{(\log\log t)^{r-1}}\right)$ vertices such that no color contains a complete graph on t vertices. Here $c_r = c/p_r^{2r} \sim c(r\ln r)^{-2r}$, where p_r is the r^{th} prime, and $c > 0$ is an absolute constant.*

The *existence* of graphs with $ct2^{t/2}$ vertices without a complete subgraph or an independent set of size t was proved in Erdős's celebrated 1947 paper [7]. Erdős's probabilistic proof can be easily adapted to yield the *existence* of an r-coloring of the edges of the complete graph on $c(r)tr^{t/2}$ vertices, without a monochromatic complete subgraph on t vertices. (The exact formula is $\lfloor (t/e)r^{(t-1)/2-1/t} \rfloor$ so we can take $c(r) = 1/(er)$.)

Proof of Theorem 4.6. Let $m = p_1 p_2 \ldots p_r$, where p_i is the i^{th} prime. Let K be a complete graph on vertex-set $\mathcal{H}$, where $\mathcal{H}$ is a set-system with the properties stated in Theorem 1.2, with $h = \lfloor t^{1/p_r} \rfloor$. We define an r-coloring of the edges of K by colors $1, 2, \ldots, r$ as follows: edge UV, where $U, V \in \mathcal{H}$, has color i if

$$i = \min_{j \in \{1,2,\ldots,r\}} \{j : p_j \text{ does not divide } |U \cap V|\}.$$

Descriptions of a function $f(x)$

$O(g(x))$ = bounded above by some $M \cdot g(x)$ $\iff$ $\frac{f(x)}{g(x)} < \infty$ as $M \to \infty$

$o(g(x))$ = bounded above by some $m \cdot g(x)$ as $m \to 0$

↳ $o(1)$ = less than any m

$\Omega(g(x))$ = bounded below by some $c \cdot g(x)$ as $c \to \infty$ (at least as fast as $g(x)$)

$\omega(g(x))$ = bounded below by some $c \cdot g(x)$ as $c \to \infty$ (faster than $g(x)$)

Overview of Grolmusz

Frankl-Wilson says we can build a set-system $\mathcal{F}$ over universe of n elements

s.t. given $\underbrace{u_0, u_1, \dots, u_s}_{S}$, distinct residues mod p:

1) $\forall F \in \mathcal{F}$, $|F| = k \equiv_p u_0$ $\quad (s \le k \le n)$ // uniform sets s.t. size of $|S| \le |F| \le n$

2) $\forall F, G \in \mathcal{F}$, $|F \cap G| \in S$ mod p

Then $|\mathcal{F}| \le \binom{n}{s}$

What happens when we go to a composite mod? *not insanely fast, but faster than $\binom{n}{s}$

Case 1: mod 6 w/ arbitrary $s \le 6$ ⟶ faster growing system sets

Case 2: mod p^2 w/ arbitrary $s \le p^2$ ⟶ faster growing system sets

Case 3: mod p^e w/ $s = m-1$ ⟶ $|\mathcal{F}| \le \binom{n}{s}$

negative ans over ↓

Grolmusz case: non-prime-power composite m and $s = m-1$ ⟶ $|\mathcal{F}| > \binom{n}{s}$

↳ **Theorem 1.2**

let m be a positive integer w/ $r > 1$ prime divisors (cannot be 1 bc then it's a prime power)

$\exists$ function $c(m) = c > 0$ s.t. $\forall$ int $h > 0$, we can construct a set system $\mathcal{H}$ over h elements s.t.

1) $\forall H \in \mathcal{H}$, $|H| \equiv 0$ mod m

2) $\forall H, G \in \mathcal{H}$, $|H \cap G| \not\equiv 0$ mod m (somewhere in s)

3) $|\mathcal{H}| \ge \exp\left(c \cdot \frac{(\log h)^r}{(\log\log h)^{r-1}}\right)$ $|\mathcal{H}|$ is bounded below by huge number as $h \to \infty$

* $c = \frac{1}{p_r^{r-r}}$ where p_r is largest prime divisor

* $|H| = h^{\frac{r-1}{2r-1} + o(1)}$, $|H|$ is bounded as $h \to \infty$

Pre-Lims

Definition 2.1: The polynomial P weakly represents boolean function f mod m if $\exists$ $S \subset \{0, 1, 2, \ldots, m-1\}$ s.t. $\forall x \in \{0,1\}^n$, $f(x) = 0 \iff (P(x) \bmod m) \in S$

WLOG: P is multilinear b/c $x_i^2 = x_i$ over $\{0,1\}^n$

e.x. $P = x_i^3 x_j + x_j^2 x_i + x_i$

$= x_i \cdot x_i x_j + x_j x_i + x_i$

$= x_i \cdot x_j +$

Summary: for prime / prime power modulo p. Min degree for a OR_n representation is $\frac{n}{p-1}$

fails for composite non-prime power modulo m.

Let $OR_n: \{0,1\}^n \to \{0,1\}$ denote n-variable OR function

$$OR_n(x_1, x_2, \ldots x_n) = \begin{cases} 0, & \text{if } \vec{x} = \vec{0} \\ 1, & \text{o/w} \end{cases}$$

Suppose $\exists$ P s.t. P weakly represents OR_n modulo prime p

$OR_n(\vec{x}) = 0 \iff P(\vec{x}) \bmod p \equiv 0 \quad (S = \{0\})$ by def of (2.1) weak rep

aka $P(\vec{x}) \equiv 0 \bmod p \iff \vec{x} = \vec{0}$

Then define $f(x) = 1 - P^{p-1}(\vec{1} - \vec{x})$ is n-variable AND function

if $\vec{x} = \vec{1}$, $P(\vec{1} - \vec{1}) \equiv 0 \bmod p \Rightarrow f(x) = 1 - 0^{p-1}$

if $\vec{x} \neq \vec{1}$, $P(\vec{1} - \vec{x}) \not\equiv 0 \bmod p \Rightarrow f(x) = 1 - (\text{something})^{p-1} = 0$

$\Rightarrow f(x) = x_1 \cdot x_2 \cdot \ldots \cdot x_n$ P_{AND} is of degree n

Therefore if P_{OR} weakly represents OR_n mod p, denote degree to be some k, then P_{AND} representing $1 - (P_{OR})^{p-1}$ is of degree n.

$1 - (\text{Degree } k)^{p-1}$ is degree n

$\Rightarrow n \leq k \cdot p - 1$ b/c of potential cancellations

$$k \geq \frac{n}{p-1}$$

Theorem 2.4

Given $m = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ where p_i are distinct primes, $\exists$ polynomial P of degree $O(n^{1/r})$ which weakly represents OR_n mod m.

Proof $S_k(x)$ denotes k^{th} elementary symmetric polynomial

↳ $S_1(x) = x_1 + x_2 + x_3$

$S_2(x) = x_1x_2 + x_1x_3 + x_2x_3$ } $n=3$

$\forall x \in \{0,1\}^n$, weight of x ($wt(x)$) $= \sum_{i=1}^{n} x_i$ (how many 1s in x)

$\Rightarrow S_k(x) = \binom{wt(x)}{k}$, $S_k(x) = \sum x_j x_i x_n x_h$ (order matters etc)

$S_k(x) \Rightarrow S_k(wt(x)) \Rightarrow S_k(j)$

since $S_k(x)$ depends on $wt(x)$ $\Rightarrow$ we can rewrite $S_k(wt(x))$

lemma 2.1: $k \in \mathbb{Z}^+$, p is prime, e is smallest integer s.t. $k < p^e$.

$\Rightarrow S_k(j) \equiv S_k(j+p^e) \mod p$

Proof WTS: $\binom{j+p^e}{k} \equiv \binom{j}{k} \mod p$

$\binom{j+p^e}{k} = \sum_{\omega=0}^{k} \binom{j}{\omega}\binom{p^e}{k-\omega}$ identity

$p \mid \binom{p^e}{l}$, $1 \le l < p^e$ fact

$\sum_{\omega=0}^{k} \binom{j}{\omega}\binom{p^e}{k-\omega} - \binom{j}{k} = \underbrace{\sum_{\omega=0}^{k-1} \binom{j}{\omega}\binom{p^e}{k-\omega}}_{\text{divisible by } p} + \binom{j}{k}\binom{p^e}{0} - \binom{j}{k} \equiv 0 \mod p$ ☑

Construction: $\forall i$, e_i is smallest integer s.t. $\lceil n^{1/r} \rceil < p_i^{e_i}$ (p_i is some prime [p], $\lceil n^{1/r} \rceil$ is a pos integer [k])

$$G_i(x) = \sum_{j=1}^{p_i^{e_i}-1} (-1)^{j+1} s_j(x) = \sum_{j=1}^{p_i^{e_i}-1} (-1)^{j+1} \binom{wt(x)}{j}$$

fact: $(1-1)^\omega = \sum_{j=0}^{\omega} \binom{\omega}{j}(-1)^j = 0$

Case 1: $wt(x) = 0$, ($\vec{x} = (0,0,\dots,0)$) $\Rightarrow$ $G_i(x) = 0$

Case 2: $0 < wt(x) < p_i^{e_i-1}$ $\Rightarrow$ $G_i(x) = 1$ (use fact)

G_i correctly computes OR mod p_i for inputs of weight $\le p_i^{e_i-1}$

$\Rightarrow$ guaranteed to work for inputs of weight $\le n^{1/r}$

G_i is of degree $n^{1/r}$ given inputs are of weight

Chinese Remainder Theorem says $\exists$ $P \equiv G_i \mod p_i$ $\Rightarrow$ P is of degree $M \cdot (n^{1/r})$ some M $O(n^{1/r})$

Corollary 2.2: $m = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$. $\exists$ P with n variables and of degree $O(n^{1/r})$ s.t.

1) $P(\vec{x}) \equiv 0 \mod m \iff \vec{x} = \vec{0}$ Theorem 2.4 states this

2) $\forall \vec{x} \in \{0,1\}^n$, $\forall p_i^{\alpha_i}$, $P(\vec{x}) \equiv \{0,1\} \mod p_i^{\alpha_i}$

Lower Bounds

Proof of Theorem 2.1

Let $P(z_1, z_2, \dots, z_n)$ be a polynomial of degree d which satisfies
$P(0, 0, \dots, 0) \equiv 0 \mod m \iff \vec{z} = \vec{0}$

Explict construction given in Theorem 2.4. (degree $O(n^{1/r})$)

Let $Q(z_1, z_2, \dots, z_n) = P(1-z_1, 1-z_2, \dots 1-z_n)$

$Q(1, 1, \dots 1) = 0$

Lemma 3.2: $\forall$ integer $n>0$, $\exists$ uniform set system $\mathcal{H}$ over $2(m-1)n^{2d}/d!$ elements

1) $|\mathcal{H}| = n^n$
2) $\forall H \in \mathcal{H}$, $|H| \equiv 0 \mod m$
3) $\forall G, H \in \mathcal{H}$, $|G \cap H| \not\equiv 0 \mod m$

if $d = \Theta(n^{1/r})$, 3.2 $\Rightarrow$ 1.2

Proof $Q(z_1, z_2, \dots, z_n) = \sum_{i_1, i_2, \dots, i_d} a_{i_1, i_2, \dots i_d} \cdot z_{i_1} z_{i_2} \dots z_{i_d}$

$\tilde{Q}(z_1, z_2, \dots, z_n) = \sum_{i_1, i_2, \dots, i_d} \tilde{a}_{i_1, i_2, \dots i_d} \cdot z_{i_1} z_{i_2} \dots z_{i_d}$

$\tilde{a}_{i_1, i_2, \dots, i_d}$: smallest positive int congruent to $a_{i_1, i_2, \dots, i_d} \mod m$

$\tilde{Q}(\vec{1}) \equiv 0 \mod m$ but $\tilde{Q}(\vec{0})$ is not nessecessarily 0

$\delta : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $u, v \in \{0, 1, 2, \dots, n-1\}$

$\delta : \{0, 1, \dots, n-1\} \times \{0, 1, \dots, n-1\} \to \{0, 1\}$ defined by $\delta(u, v) = \begin{cases} 1, & u=v \\ 0, & \text{o/w} \end{cases}$

every available vector of size n so n^n many

Let $A \in M_{n^n \times n^n}$ s.t. rows are $x \in \{0, 1, \dots, n-1\}^n$ and collomns $y \in \{0, 1, \dots, n-1\}^n$

entry $a_{xy} = \tilde{Q}(\delta(x_1, y_1), \dots \delta(x_n, y_n))$ 1 if $x=y$ & 0 o/w

$\tilde{Q}(\vec{z})$ is a sum of monomials $z_{i_1} z_{i_2} \dots z_{i_\ell}$ $(\ell \le d)$

↓ m copies = 0 copies b/c mod ↓

break it into multiplicities (less than m-1) so coefficient is always 1

Matrix A

Define $B_{i_1,\dots,i_\ell}$ as $b_{xy}^{i_1,\dots i_\ell} = \delta(x_{i_1}, y_{i_1}) \cdot \dots \cdot \delta(x_{i_\ell}, y_{i_\ell})$

$$A = \sum_{i_1, i_2, \dots i_\ell} \tilde{a}_{i_1, i_2, \dots, i_\ell} \cdot B_{i_1, i_2, \dots, i_\ell}$$

Let $n=2$. $x, y \in \{0,1\}^2 = \{0,1\}$ or $\{1,0\}$

$$A = \begin{matrix} \\ \{0,1\} \\ \{1,0\} \end{matrix} \begin{matrix} \begin{matrix} \{0,1\} & \{1,0\} \end{matrix} \\ \begin{bmatrix} \tilde{Q}(1,1) & \tilde{Q}(0,0) \\ \tilde{Q}(0,0) & \tilde{Q}(1,1) \end{bmatrix} \end{matrix} \Rightarrow \tilde{Q} = \text{some polynomial (sum of monomials)}$$

Say case $\tilde{Q} = v_1 v_2 + v_1 v_2 + v_1 + v_2$ (B^{12}, B^{12}, B^1, B^2; the $v_1 v_2$ pair is a multiplicity of 2)

$$A = \begin{bmatrix} B'^2(1,1) + B^{12}(1,1) + B^1(1,1) + B^2(1,1) & B'^2(0,0) + B^{12}(0,0) + B^1(0,0) + B^2(0,0) \\ B'^2(0,0) + B^{12}(0,0) + B^1(0,0) + B^2(0,0) & B'^2(1,1) + B^{12}(1,1) + B^1(1,1) + B^2(1,1) \end{bmatrix} = B^{12} + B^{12} + B^1 + B^2$$

$$= 2B^{12} + B^1 + B^2$$

Technically

	$\{1,0\}$	$\{0,1\}$	$\{1,1\}$	$\{0,0\}$
$\{1,0\}$	$\tilde{Q}(1,1)$	$\tilde{Q}(0,0)$	$\tilde{Q}(1,0)$	$\tilde{Q}(0,1)$
$\{0,1\}$	$\tilde{Q}(0,0)$	$\tilde{Q}(1,1)$	$\tilde{Q}(0,1)$	$\tilde{Q}(1,0)$
$\{1,1\}$	$\tilde{Q}(1,0)$	$\tilde{Q}(0,1)$	$\tilde{Q}(1,1)$	$\tilde{Q}(0,0)$
$\{0,0\}$	$\tilde{Q}(0,1)$	$\tilde{Q}(1,0)$	$\tilde{Q}(0,0)$	$\tilde{Q}(1,1)$

Lemma 3.3: B-Blocks covering each cell = a_{xy}

1) Diagnol of A is when $x=y$ so $a_{xy} \equiv 0 \mod m$

2) $\forall x, y \quad a_{xy} \not\equiv 0 \mod m$

only looks at certain inputs

$\tilde{Q}$ = a shit ton of monomials w/ some coefficient represented as $\tilde{a}_{i_1,\dots,i_\ell} \cdot z_1 \dots z_\ell$ (coefficient · inputs)

$\Rightarrow \tilde{a}$ blocks of $B_{1,\dots,\ell}$

slots 1 where those inputs agree and 0 o/w.

if all B-Blocks output 1, $\equiv 0 \mod m$

$\not\equiv 0 \mod m$ otherwise

Corollary 3.3: $\exists$ hyper graph G with n^n vertices and fewer than $2(m-1)n^{2d}/d!$ edges s.t. every vertex is contained in equal number of edges ($\equiv 0 \bmod m$) and any two vertices share number of vertices ($\not\equiv 0 \bmod m$)

Chooses cells of A's Diagnal (n^n) as vertices an intersections of the B-blocks for edges

↳ Edge $B_{i_1,\dots,i_d}$ connects all vertices that have inputs $i_1,\dots,i_d$ in common.

$$\text{Edges} = n \cdot = \tilde{Q}(n,n,\dots,n) = \sum_{\ell \le d} \sum' \hat{a}_{i_1,i_2,\dots,i_\ell} \cdot n^\ell$$

(multiplicity; $\sum'$: all monomials (degree of Q))

give agreement at only ℓ places, what vector can be picked?

Assume $\ell = 0, 2$

$\forall$ edge connects $\{a,\_,b,\_,\dots\}$

but what are values of a,b? $\exists$ n options for either $\Rightarrow n^2$ options (n^ℓ)

$$\le (m-1) \sum_{\ell \le d} \binom{n}{\ell} n^\ell$$

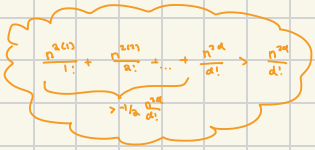
(less than $m-1$ multiples of each)

$$< (m-1) \sum_{\ell \le d} n^{2\ell}/\ell!$$

($> n^{2d}/d!$; end term dominates)

we care about end behavior

$$< (m-1)\, 2n^{2d}/d! \quad \text{assuming } n \ge 2d$$



$$\#\text{ of edges per vertex} = \tilde{Q}(1,1,\dots,1) \le (m-1)\left[\binom{n}{d} + \binom{n}{d-1} + \dots + \binom{n}{0}\right] < 2(m-1)\binom{n}{d}$$

(max multiplicity; how many others match in $\le d$ places?; same as before, rest contributes less than half of sum so $2\binom{n}{d} > \binom{n}{d} + \dots$)

$\forall$ vertex has $\tilde{Q}(1,1,\dots,1)$ edges $\equiv 0 \bmod m$

$\forall$ two vertices has $\tilde{Q}(1,\dots,0,\dots,1) \not\equiv 0 \bmod m$

(at least one zero somewhere)

Set Construction (Dual Hyper Graph of $G =: \mathcal{H}$)

Elements: Edges of G ← B-Blocks (all variations)

Sets: Vertices of G ← different vectors

elements $= h := \tilde{Q}(n, n, \dots, n)$

$(h < 2(m-1)n^{2d}/d!)$

types of B-Blocks (i.e. z_1, z_8 1st and 8th input)

times how much "variety"

e.x. $\tilde{Q}(z_1, \dots, z_n) = \underbrace{z_1 z_2}_{\text{B-Blocks}} + z_8 z_6 z_4$

so $\tilde{Q}(1, \dots, 1) = 2$, all "types of B-Blocks"

$\tilde{Q}(n, \dots, n) = n^2 + n^1$, all "variations"

sets $= n^n$ the dif vectors

$\Rightarrow$ $\mathcal{H}$ is a collection of n^n sets over $\tilde{Q}(n, n, \dots, n)$ elements

---

Rmk 3.3

Universe of $\tilde{Q}(n, \dots, n)$ elements w/ size of $\forall H \in \mathcal{H}$ $= \tilde{Q}(1, 1, \dots, 1)$

$\Rightarrow \tilde{Q}(n, \dots, n) \le n^d \tilde{Q}(1, 1, \dots, 1)$ (maximizes all types of B-Blocks)

$h \xrightarrow{\text{add dummies}} h' = n^d \tilde{Q}(1, 1, \dots, 1) \le 2(m-1)n^{2d}/d!$ (max multiplicity from G construction)

$n^d \ge h'^{\frac{r}{2r-1} + o(1)}$ not sure how tbh but $\tilde{Q}(1, \dots, 1) \le (h')^{\frac{r-1}{2r-1} + o(1)}$

Conclusion of Lemma 3.2

$|\mathcal{H}| = n^n$

$\forall H \in \mathcal{H}$, $|H| = \tilde{Q}(1, 1, \dots, 1) \equiv 0 \mod m$ (a_{xx}) (diagonal)

$\forall H, G \in \mathcal{H}$, $|H \cap G| \not\equiv 0 \mod m$ (a_{xy}) (off diagonal)

Conclusion of Grolmusz

From CR_n reps mod m, we get degree $n^{1/r}$ polynomials.

This allows us to form $P(z_1, \dots, z_n)$ s.t. $P(\vec{z}) \equiv 0 \bmod m \iff \vec{z} = \vec{0}$

and therefore $Q(\vec{z}) = P(1 - \vec{z}) \equiv 0 \bmod m$ iff $\vec{z} = \vec{1}$.

Lemma 3.2

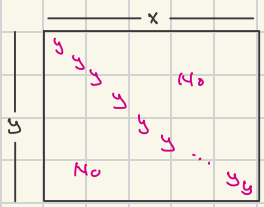
$\forall n > 0$, $\exists$ uniform set system $\mathcal{H}$ over $2(m-1)n^{2d}/d!$ elements s.t.

1) $|\mathcal{H}| = n^n$
2) $\forall H \in \mathcal{H}$, $|H| \equiv 0 \bmod m$
3) $\forall G, H \in \mathcal{H}$, $|H \cap G| \not\equiv 0 \bmod m$

a bunch of monomials of some multiplicity

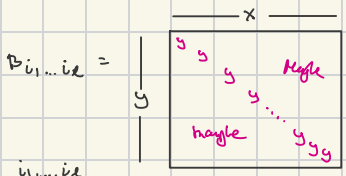
$$\tilde{Q}(z_1, \dots, z_n) = \sum_{i_1, \dots, i_n} \tilde{a}_{i_1, \dots, i_n} \cdot (z_{i_1} \cdot \dots \cdot z_{i_n})$$

$x, y \in \{0, 1, 2, \dots, n-1\}^n$



$\equiv 0$ if all $x_i = y_i \Rightarrow x = y$
$\neq 0$ o/w

$$a_{xy} = \tilde{Q}(\delta(x_1, y_1), \delta(x_2, y_2), \dots)$$



$$B_{i_1 \dots i_\ell} = B^{i_1, \dots, i_\ell}_{xy} = \delta(x_{i_1}, y_{i_1}) \cdot \delta(x_{i_2}, y_{i_2}) \cdot \dots \cdot \delta(x_{i_\ell}, x_{i_\ell})$$

do the indices agree $\forall$ combos?

Build Hypergraph $\Longrightarrow$ Dual (gives set system of 3.2)

$\Longrightarrow$ 1.2

start with an n, we "find" an n via 3.2 to give a close enough universe $\approx n$.

$|\mathcal{H}| = n^n$, to prove $n^n \geq \exp\left(\frac{c(\log n)}{\log\log n}\right)$ we use Stirling approx

Verification

$r = 2$

lets say we want $h = \frac{n^{2\sqrt{n}}}{(\sqrt{n})!}$

Via 3.2, for n, $\exists$ $\mathcal{H}$ over $2(m-1)n^{2d}/d!$ elements ($d = O(n^{1/r}) \Rightarrow O(n^{1/2})$ due to degree of OR_n representation of a m w/ 2 prime factors)

$\Rightarrow$ n gives us $\mathcal{H}$ over $\underbrace{2(m-1)}_{\text{ignore b/c constant}} \cdot \frac{n^{2(\sqrt{n})}}{(\sqrt{n})!}$

3.2 says $|\mathcal{H}| = n^n$.

By stirling's approximation: $\ln(n!) = n\ln(n) - n + O(\ln n)$

Prove $n^n \geq \exp\left(\frac{c \log^2 h}{\log\log h}\right) \Rightarrow \underbrace{n \log(n) \geq \frac{c \log^2 h}{\log\log h}}_{\text{WTS}}$

$\log h = 2\sqrt{n}\log(n) - \log(\sqrt{n}!)$

$= 2\sqrt{n}\log(n) - [\sqrt{n}\log(\sqrt{n}) - \sqrt{n}]$

$= 2\sqrt{n}\log(n) - \frac{1}{2}\sqrt{n}\log(n) - \sqrt{n}$

$= 1.5\sqrt{n}\log(n)$

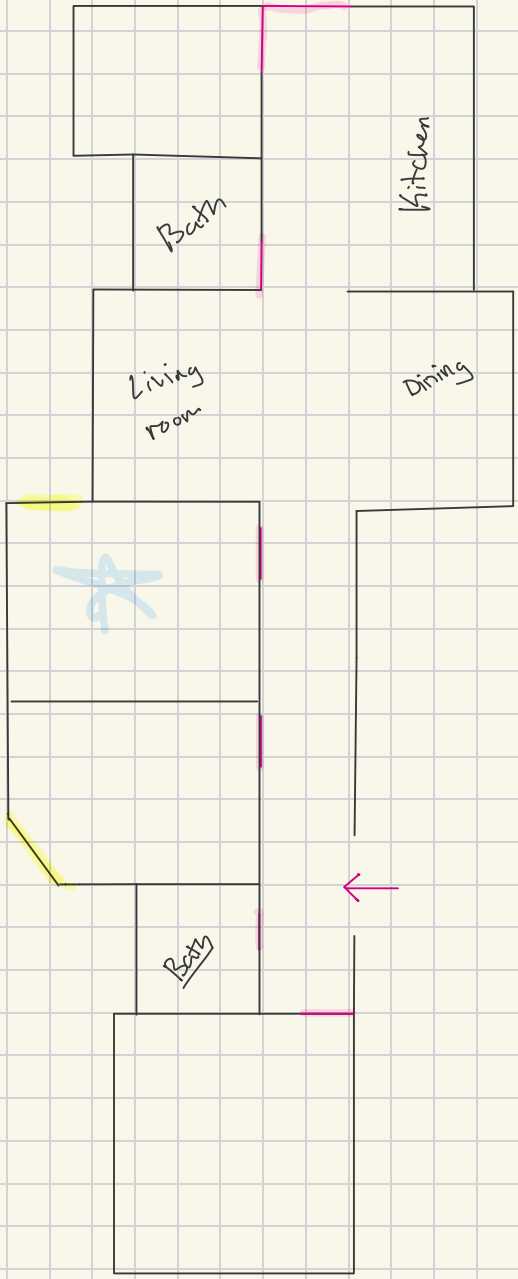
$\log\log h = \log(1.5\sqrt{n}) + \log\log(n)$

$= \underbrace{\log(1.5)}_{\text{constant}} + \frac{1}{2}\log(n) + \log\log(n)$

$\frac{c \cdot [1.5\sqrt{n}\log(n)]^2}{(\log(1.5) + \frac{1}{2}\log(n) + \log\log(n))}$ $\log(n) > \log\log(n)$

$\geq \frac{c \cdot 1.5^2 \cdot n \cdot (\log^2(n))}{1.5\log(n)}$

$= c \cdot n\log(n)$ ∎ c is < 1


Bath
Kitchen
Living room
Dining
Bath

3-Query Locally Decodable Codes of Subexponential Length

Klim Efremenko *

November 13, 2008

Abstract

Locally Decodable Codes (LDC) allow one to decode any particular symbol of the input message by making a constant number of queries to a codeword, even if a constant fraction of the codeword is damaged. In a recent work [Yek08] Yekhanin constructs a 3-query LDC with sub-exponential length of size $\exp(\exp(O(\frac{\log n}{\log\log n})))$. However, this construction requires a conjecture that there are infinitely many Mersenne primes. In this paper we give an unconditional 3-query LDC construction with a shorter codeword length of $\exp(\exp(O(\sqrt{\log n \log\log n})))$. We also give a 2^r-query LDC with length of $\exp(\exp(O(\sqrt[r]{\log n \log\log^{r-1} n})))$. The main ingredient in our construction is the existence of super-polynomial size set-systems with restricted intersections by [Gro00] which hold only over composite numbers.

1 Introduction

Locally decodable codes (LDCs) are codes that allow to retrieve any symbol of the original message by reading only a constant number of symbols from the codeword. Formally a code C is said to be locally decodable with parameters (q, δ, ε) if it is possible to recover any bit x_i of message x by making at most q queries to $C(x)$. Such that if up to a δ fraction of $C(x)$ is corrupted then the decoding algorithm will return the correct answer with probability at least $1 - \varepsilon$.

Locally decodable codes have many applications in cryptography and complexity theory, see surveys in [Tre04] and [Gas04]. The first formal definition of locally decodable codes was given by Katz and Trevisan in [KT00]. The Hadamard code is the most famous 2-query locally decodable code of length 2^n. For a two-query LDC tight lower bounds of $2^{\theta(n)}$ were given for linear codes in [GKST02] and [KdW03] proved tight lower bounds for two queries for arbitrary codes. For an arbitrary number of queries Katz and Trevisan [KT00] established super-linear lower bounds of $\Omega(n^{q/(q-1)})$ for LDCs with q queries. This lower bound was later improved in [KdW03] to $\Omega\left((\frac{n}{\log n})^{1+1/(\lceil q/2\rceil-1)}\right)$ and in [Woo07] to $\Omega\left(\frac{n^{1+1/(\lceil q/2\rceil-1)}}{\log n}\right)$.

*Weizmann Institute of Science, Rehovot 76100, Israel, Bar-Ilan University, 52900 Ramat-Gan, Israel; klimefrem@gmail.com

For many years it was conjectured that LDCs should have an exponential dependence on n for any constant number of queries, until Yekhanin's recent breakthrough [Yek08]. Yekhanin obtained 3-query LDCs with sub-exponential length of $\exp(\exp(O(\frac{\log n}{\log\log n})))$ under a highly believable conjecture that there are infinitely many Mersenne primes. Using the known Mersenne primes, Yekhanin also obtained unconditional results which significantly improved the previous results on LDCs(i.e. length of $\exp(n^{10^{-7}})$). In [KY08] Kedlaya and Yekhanin proved that infinitely many Mersenne numbers with large prime factors are essential for Yekhanin's construction.

Our Results In this paper we give an unconditional construction of 3-query LDC with sub-exponential codeword length. The length that we achieve for 3 queries is:

$$\exp\exp(O(\sqrt{\log n \log\log n})).$$

We also give a 2^r-query LDC with a codeword length $\exp\exp(O(\sqrt[r]{\log n \log\log^{r-1} n}))$.

Our construction is a kind of a generalization and simplification of [Yek08]. We extend Yekhanin's construction to work not only with primes but also with composite numbers. Raghavendra in [Rag07] gives a nice presentation of Yekhanin's construction using homomorphisms, and we will follow this approach. The main ingredient in our construction is the existence of super-polynomial size set-systems with restricted intersections [Gro00], which hold only over composite numbers.

Private Information Retrieval schemes: The notion of locally decodabale codes is closely related to the notion of private information retrieval(PIR) schemes. PIR schemes with k servers is a protocol which allows for a user to access a database distributed between k servers without yielding any information on the identity of the accessed place to any individual server (we assume that there is no communication between servers). The main parameter of interest in PIR schemes is the total communication complexity between the user and the servers. PIR schemes were first introduced by [CGKS95]. After that there were many works written on this topic, see [CGKS95, Amb97, Man98, Ito99, BIK05, GKST06, KdW03, RY07, WdW05, Yek08]. The best upper bound for 2-server PIR is $O(n^{1/3})$ due to [CGKS95]. The best upper bound of 3 and more server PIR schemes is $\exp\left(O\left(\frac{\log n}{\log^{1-\varepsilon}\log n}\right)\right)$ due to [Yek08] which is based on the construction of LDCs.

Let us define formally perfect PIR schemes:

Definition 1.1. A one-round perfect *private information retrieval* scheme is a randomized algorithm $\mathcal{U}$ (for the user), and k deterministic algorithms $\mathcal{S}_1, \ldots \mathcal{S}_k$ (for the servers), s.t.

1. (a) On input $i \in [n]$ the user $\mathcal{U}$ produces k random queries $q_1 \ldots q_k$ and sends them to respective servers.
   (b) Each server based on his query q_j and database $\mathcal{D}$ produces a response $r_j = \mathcal{S}_i(\mathcal{D}, q_j)$ and sends it back to the user.
   (c) The user based on $i, r_1, \ldots, r_k$ and his randomness calculates $\mathcal{D}[i]$.
2. The distribution of each query q_j is independent of the input i.

The communication complexity of this protocol is a total number of bits exchanged between user and servers.

It is well known that LDCs with perfectly smooth decoder imply PIR schemes. In particular, as in [Yek08], our LDC yields a PIR schemes with communication complexity $\exp(O(\sqrt{\log n \log\log n}))$ for 3-servers and $\exp(O(\sqrt[r]{\log n \log^{r-1} \log n}))$ for 2^r-servers.

2 Definitions and Basic Facts

We will use the following standard mathematical notation:

- $[s] = \{1, \dots s\}$;
- $\mathbb{F}_q = GF(q)$ is a finite field of q elements;
- $\mathbb{F}^*$ is a multiplicative group of the field;
- $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$, the integers modulo m;
- $d_H(x, y)$ denotes the Hamming distance between vectors $x, y \in \Sigma^n$, i.e. number of indices where $x_i \neq y_i$.

Definition 2.1. A code C over a field $\mathbb{F}$, $C : \mathbb{F}^n \mapsto \mathbb{F}^N$ is said to be (q, δ, ε) locally decodable if there exist randomized decoding algorithms d_i for $i = 1, 2, \dots n$ such that for all $i = 1, 2, \dots n$ the following holds:

1. For every message $\vec{x} = (x_1, x_2, \dots x_n) \in \mathbb{F}^n$ and for every $\vec{y} \in \mathbb{F}^N$ such that $d_H(C(\vec{x}), \vec{y}) \leq \delta N$ it holds that $\Pr(d_i(\vec{y}) = x_i) \geq 1 - \varepsilon$; i.e. the decoding algorithm succeeds to recover the relevant symbol even if up to δ fraction of the codeword is damaged.
2. The algorithm d_i makes at most q queries to y.

A code C is called linear if C is a linear transformation over $\mathbb{F}$. A locally decodable code is called nonadaptive if d_i makes all its queries simultaneously. Our constructions of locally decodable codes are linear and nonadaptive.

Definition 2.2. A code C is said to have a *perfectly smooth decoder* if $d_i(C(\vec{x})) = x_i$ for every $\vec{x}$ and each query of d_i is uniformly distributed over $[N]$.

Fact 2.3 (from [Tre04])**.** *Any code with a perfectly smooth decoder which makes q queries is also $(q, \delta, q\delta)$ locally decodable.*

We will use the following fact:

Fact 2.4. *For every odd m there exists a finite field $\mathbb{F} = GF(2^t)$, where $t \leq m$, and an element $\gamma \in \mathbb{F}$ that is a generator of a multiplicative group of size m, i.e. $\gamma^m = 1$ and $\gamma^i \neq 1$ for $i = 1, 2, \dots m - 1$.*

Proof. Since m is odd $2 \in \mathbb{Z}_m^*$. Therefore, there exists $t < m$ such that $2^t \equiv 1 \mod m$. Let us set $\mathbb{F} = GF(2^t)$. The size of the multiplicative group $\mathbb{F}^*$ is $2^t - 1$ and therefore it is divisible by m. Let g be a generator of $\mathbb{F}^*$. Then $\gamma = g^{\frac{2^t-1}{m}}$ is a generator of a multiplicative group of size m. □

In Appendix A for simple construction of S-matching vectors we will need the following definition and fact about tensor product:

Definition 2.5 (Tensor Product)**.** Let R be a ring and let $\vec{x}, \vec{y} \in R^n$. The *tensor product of* $\vec{x}, \vec{y}$ denoted by $\vec{x} \otimes \vec{y} \in R^{n^2}$, is defined by $\vec{x} \otimes \vec{y}(i,j) \triangleq x_i \cdot y_j$, (where we identify $[n^2]$ with $[n] \oplus [n]$.) In the same way we define *the ℓ'th tensor power* $\vec{x}^{\otimes \ell} \in R^{n^\ell}$ by

$$\vec{x}^{\otimes \ell}(i_1, i_2, \ldots i_\ell) \triangleq \prod_{j=1}^{\ell} x_{i_j}. \tag{1}$$

We will use only the following fact about tensor products:

Fact 2.6.

$$\langle u^{\otimes \ell}, v^{\otimes \ell} \rangle = \langle u, v \rangle^\ell$$

Proof.

$$\begin{aligned} \langle u^{\otimes \ell}, v^{\otimes \ell} \rangle = \sum_{1 \le i_1, i_2, \ldots i_l \le m} \left(\prod_{j=1}^{\ell} u_{i_j} \prod_{j=1}^{\ell} v_{i_j} \right) = \\ \left(\sum_{1 \le i_1 \le m} u_{i_1} v_{i_1} \right) \ldots \left(\sum_{1 \le i_\ell \le m} u_{i_\ell} v_{i_\ell} \right) = \langle u, v \rangle^\ell. \end{aligned} \tag{2}$$

□

3 Locally Decodable Codes

In this construction we follow Yekhanin's general framework. Our construction consists of two parts. The first part is a construction of matching sets of vectors that correspond to "combinatorially nice" sets used in [Yek08]. The second part is a construction of an S-decoding polynomial with a small number of monomials, which correspond to "algebraically nice" sets used in [Yek08]. Let us fix some composite number m for our construction. We wiil give a general scheme for construction of LDCs followed by a concrete example of a 3-query LDC.

3.1 Matching sets of vectors

All inner products $\langle x, y \rangle$ in this section are done $\mod m$.

Definition 3.1. The family of vectors $\{u_i\}_{i=1}^n, u_i \in (\mathbb{Z}_m)^h$ is said to be *S-matching* if the following conditions hold:

1. $\langle u_i, u_i \rangle = 0$ for every $i \in [n]$.
2. $\langle u_i, u_j \rangle \in S$ for every $i \neq j$.

holds via Grolmusz's super polynomial sets (mod m)

The goal of this subsection is to construct large S-matching family over a small domain. The main advantage of working with composite numbers comes from the following lemma from [Gro00], which holds only for composite numbers.

Lemma 3.2 (Theorems 1.2 and 1.4 from [Gro00]). *Let $m = p_1 p_2 \dots p_r$ be a product of r distinct primes p_i. Then there exists $c = c(m) > 0$, such that for every integer $h > 0$, there exists an explicitly constructible set-system $\mathcal{H}$ over a universe of h elements (i.e $\mathcal{H}$ is a set of subsets of $[h]$) and there is a set $S \subset \mathbb{Z}_m$ such that:*

1. $|\mathcal{H}| \geq \exp(c\frac{(\log h)^r}{\log\log^{r-1} h})$,
2. *Size of every set H in set-system $\mathcal{H}$ is divisible by m i.e.*
3. *Let G, H be any two different sets in set system $\mathcal{H}$. Ther G, H modulo m is restricted to be in S. i.e. $\forall G, H \in \mathcal{H}$ that $|G \cap H| \in S \mod m$*
4. *S is a set of size $2^r - 1$ and $0 \notin S$.*
5. *$\forall s \in S$ for all $i = 1, 2, \dots r$ it holds that $s \mod p_i$ is 0 or 1.*

★ Theorem 1.3

By CRT: $\forall v \in \{0,1\}^r \exists$ a unique $x \in \mathbb{Z}_m$ s.t. $x \equiv v_i \bmod p_i$

Corollary 3.3
Using set-systems (verified by grolmusz) over h elements denote by $v_1, \dots, v_n$
$|\mathcal{H}| = n$ (new notation)
$\forall H \in \mathcal{H}$, $\exists$ an indicator vector, v, of length h. (0 if $v_i \notin H$, 1 if $v_i \in H$)
$\Rightarrow \langle H, G \rangle$ = how many overlaps in elements
$\langle H, H \rangle \equiv 0 \bmod m$
$\langle H, G \rangle \not\equiv 0 \bmod m$, $H \neq G$
Me

For our construction we will only need the following simple corollary:

Corollary 3.3. *For every h, r and integer $m = p_1 p_2 \dots p_r$ there exists a set S of size $2^r - 1$ and a family of S-matching vectors $\{u_i\}_{i=1}^n, u_i \in (\mathbb{Z}_m)^h$ such that $n \geq \exp(c\frac{(\log h)^r}{\log\log^{r-1} h})$.*

u_i = sets in $\mathcal{H}$. $\exists$ n such sets

Proof. Let us take set-system $\mathcal{H}$ as in Lemma 3.2. For each set $H \in \mathcal{H}$ we will have one vector $u_H \in (\mathbb{Z}_m)^h$ which is the indicator vector of H. Then it holds that $\langle u_H, u_H \rangle = |H| \equiv 0 \mod m$ and $\langle u_H, u_G \rangle = |H \cap G| \in S \mod m$. □

The construction of [Gro00] is complicated; therefore, we will not give it here. We will give a simple construction of S-matching set in Appendix A which is less strong but it is more simple.

3.2 S-decoding polynomials

Let us fix any odd number m. Recall from Fact 2.4 that there exists t, $\mathbb{F} = GF(2^t)$ and an element $\gamma \in \mathbb{F}$ such that γ is a generator of a multiplicative group of size m. We will first construct a linear code over the field $\mathbb{F}$. In the next section we will show how to reduce the alphabet size to 2.
We will need the following definition:

$C: \mathbb{F}^n \to \mathbb{F}^{m^h}$ maps inputs of length n to codewords of length m^h, one for each indicator vector in $(\mathbb{Z}_m)^h$

Each codeword is a function $: (\mathbb{Z}_m)^h \to \mathbb{F}$

let $x \in \{\mathbb{F}^3\}^n$, $C(x)$ = Codeword$_x$

let $v \in \{\mathbb{F}^3\}^n$, an indicator vector (i, $v_i \neq 0$?)

$C(x)(v)$ = v^{th} coordinate of Codeword$_x$

$x = (1,0)$ $\quad C(x) = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ $\quad [0,1,1] \leftarrow \mathbb{F}_3, \mathbb{F}_3$ $\quad [1,1,0] \leftarrow \mathbb{F}_2, \mathbb{F}_3$ $\quad [1,0,1] \leftarrow \mathbb{F}_2, \mathbb{F}_3$

$C(x)(C[0,1,1]) = a$ Incorrect technically but general idea

Me

3.4. A polynomial $P \in \mathbb{F}[x]$ is called an S-decoding polynomial if the following hold:

$\frac{(x-\gamma^{s_1})\cdot(x-\gamma^{s_2})\ldots(x-\gamma^{s_{m-1}})}{(1-\gamma^{s_1})\cdot(1-\gamma^{s_2})\ldots(1-\gamma^{s_{m-1}})}$

- $P(\gamma^s) = 0$,
- $= P(1) = 1$.

For any S such that $0 \notin S$ there exists an S-decoding polynomial P with at monomials.

$P(\gamma^s) = 0$ $\forall s \in S$ ensures that $P(1) = 1$

s take $\tilde{P} = \prod_{s \in S}(x - \gamma^s)$. Then $P(x) = \tilde{P}(x)/\tilde{P}(1)$ is an S decoding polynomial. of P is $|S|$. Thus P has at most $|S| + 1$ monomials. □

expand $\tilde{P}$ and you get $|S|+1$ monomials

$\tilde{P}(1)^{-1}$ exists and is a polynomial b/c $\mathbb{F}$ is a field.

3.3 The code and its decoding algorithms

Now we are ready to present the construction of our locally decodable codes.
In order to construct our code we will fix some set S and construct S-matching vectors $\{u_i\}_{i=1}^n$, $u_i \in (\mathbb{Z}_m)^h$ and an S-decoding polynomial P. We define a code $C : \mathbb{F}^n \mapsto \mathbb{F}^{m^h}$ where we think of a codeword as a function from $(\mathbb{Z}_m)^h$ to $\mathbb{F}$. Let $e_i \in \mathbb{F}^n$ be the i'th unit vector. We define C by defining $C(e_i)$ for all i. The general definition will follow by the linearity of C, i.e. $C(\sum c_i e_i) \triangleq \sum c_i C(e_i)$. The encoding of e_i is

$$C(e_i) \triangleq (\gamma^{<u_i,x>})_{x \in (\mathbb{Z}_m)^h}. \tag{3}$$

One can think of $C(e_i)$ as a homomorphism from the additive group $(\mathbb{Z}_m)^h$ to the multiplicative group $\mathbb{F}^*$. Equivalently, we can write

$$C((c_1, c_2, \ldots c_n)) \triangleq \sum_{i=1}^{n} c_i f_i, \tag{4}$$

where $f_i(x) \triangleq \gamma^{<u_i,x>}$.

We will now describe how to retrieve the i'th coordinate of the message.

Since P is an S-decoding polynomial and $\{u_i\}$ are S-matching vectors, $\langle u_j, u_i \rangle \in S$ for $i \neq j$, and therefore it follows that $P(\gamma^{<u_i,u_i>}) = 1$ and $P(\gamma^{<u_j,u_i>}) = 0$ for all $i, j \in [n], i \neq j$.
Write $P(x) = a_0 + a_1 x^{b_1} + a_2 x^{b_2} \ldots a_{q-1} x^{b_{q-1}}$. ($a_0 \in \mathbb{F}$, $b_1 \in \mathbb{Z}_m$)
Let us now define the decoding algorithm $d_i(w)$, where w is a codeword with up to δ fraction damaged coordinates.

- Choose $v \in (\mathbb{Z}_m)^h$ at random.
- Query $w(v), w(v + b_1 u_i), \ldots w(v + b_{q-1} u_i)$.
- Output

$$c_i = \gamma^{-<u_i,v>} \left(a_0 w(v) + a_1 w(v + b_1 u_i) \ldots a_{q-1} w(v + b_{q-1} u_i)\right). \tag{5}$$

Algorithm 1: The Decoding Algorithm

Lemma 3.5. *The decoding algorithm d_i is a Perfectly Smooth Decoder.*

Proof. The algorithm d_i chooses v uniformly at random. Each of the queries $v, v+b_1u_i, \dots v+b_{q-1}u_i$ is uniformly distributed. Therefore, in order to prove that d_i is a Perfectly Smooth Decoder it is enough to prove that $d_i(C(x)) = x_i$. Note that d_i is a linear mapping so it is enough to prove that $d_i(C(e_i)) = 1$ and $d_i(C(e_j)) = 0$ for $j \neq i$.

$$d_i(C(e_i)) = (\gamma^{-<u_i,v>})(a_0\gamma^{<u_i,v>} + a_1\gamma^{<u_i,v+b_1u_i>} + \dots + a_{q-1}\gamma^{<u_i,v+b_{q-1}u_i>}).$$

But $\langle u_i, v + cu_i\rangle = \langle u_i, v\rangle + c\langle u_i, u_i\rangle = \langle u_i, v\rangle$. So we have,

$$d_i(C(e_i)) = \gamma^{-<u_i,v>}(a_0\gamma^{<u_i,v>} + a_1\gamma^{<u_i,v>} + \dots + a_{q-1}\gamma^{<u_i,v>}) = \\ = a_0 + a_1 \dots + a_{q-1} = P(1) = 1.$$

Now let us prove that

$$\forall i \neq j \quad d_i(C(e_j)) = 0.$$

Math of decoder

We need to show that

$$a_0\gamma^{<u_i,v>} + a_1\gamma^{<u_i,v+b_1u_j>} + \dots + a_{q-1}\gamma^{<u_i,v+b_{q-1}u_j>} = 0.$$

Recall that $P(\gamma^{<u_i,u_j>}) = 0$. Therefore,

$$\gamma^{<u_i,v>}(a_0 + a_1\gamma^{b_1<u_i,u_j>} + \dots + a_{q-1}\gamma^{b_{q-1}<u_i,u_j>}) = \gamma^{<u_i,v>}P(\gamma^{<u_i,u_j>}) = 0.$$

□

$h = \exp(O(\sqrt[r]{\log n \log\log^{r-1}(n)}))$

$\exists |S| = 2^r - 1$

$|\mathcal{H}| = \exp\left(c \cdot \frac{(\log h)^r}{\log\log^{r-1} h}\right)$

S matching vectors

$C : \mathbb{F}^n \to \mathbb{F}^{m^h}$

$|C(n)| = \exp(\exp(\sqrt[r]{\log n \log\log^{r-1} n}))$

codeword length

Decoding Polynomial w/ (at most) $|S|+1$ terms

$\Rightarrow q \leq |S| + 1 = 2^r$

Me

f the code is n-the number of S-matching vectors. The codeword length ^{h} and the number of queries is equal to the number of monomials of te corollary from Corollary 3.3 and Claim 3.1 is that we can choose ^{r}) and an S-decoding polynomial with less than 2^r monomials. Thus wing theorem.

we never talk about δ

or any r there exists a $(q, \delta, q\delta)$ locally decodable code $C : F^n \mapsto F^N$, with $N = \exp(\exp(O(\sqrt[r]{\log n \log\log^{r-1} n})))$ and $q \leq 2^r$. Furthermore, q is the of monomials of S-decoding polynomial.

$_1 \dots p_r$ be the product of r primes. Fix $h = \exp\left(\left(O(\sqrt[r]{\log n \log\log^{r-1} n})\right)\right)$. .3 there exists a set S of size $2^r - 1$ and $n = \exp(c\frac{(\log h)^r}{\log\log^{r-1} h})$ S-matching he construction above we get a code C with codeword length m^h and . Fix m to be a constant. Then $m^h = \exp(O(h))$. Therefore,

$$^h = \exp(O(h)) = \exp\left(\exp\left(O\left(\sqrt[r]{\log n \log\log^{r-1} n}\right)\right)\right).$$

From Claim 3.1 there exists an S-decoding polynomial with $q \leq 2^r$ monomials. Using this polynomial for our decoding algorithm we get from Lemma 3.5 that C has a Perfectly Smooth Decoder which makes q queries. Thus from Fact 2.3 we have that the code C is a $(q, \delta, q\delta)$-LDC. □

The Claim 3.1 gives a trivial polynomial with 2^r monomials. The natural question is: "Do polynomials exist with less monomials?" The answer is **Yes**. Let us give a concrete example of an S-decoding polynomial with 3 monomials. We found this example by an exhaustive search.

Example 3.7. *Let $m = 511 = 7 \cdot 73$ and let $S = \{1, 365, 147\}$. By Corollary 3.3 there exists S-matching vectors $\{u_i\}_{i=1}^n$, $u_i \in (\mathbb{Z}_m)^h$, where $n \geq \exp(c\frac{(\log h)^2}{\log\log h})$. Set*

$$\mathbb{F} = GF(2^9) = \mathbb{F}_2[\gamma]/(\gamma^9 + \gamma^4 + 1).$$

It can be verified that γ is a generator of $\mathbb{F}^*$ and that the polynomial $P(x) := \gamma^{423} \cdot x^{65} + \gamma^{257} \cdot x^{12} + \gamma^{342}$ is an S decoding polynomial with 3 monomials.

An interesting question is what is the best S-decoding polynomial we can choose for $r > 2$? An immediate corollary from this example and Theorem 3.6 is 3-query LDC.

Theorem 3.8. *There exists a $(3, \delta, 3\delta)$ locally decodable code of length $\exp(\exp(O(\sqrt{\log n \log\log n})))$.*

Proof?

4 Binary Locally Decodable Codes

In this section we will think of $\mathbb{F}_{2^t}$ as a vector space $\mathbb{F}_2^t$ over $\mathbb{F}_2$. We will
rmation i.e. for every $a \in \mathbb{F}_{2^t}$ there exists an n by n
ax.

we have message $(c_1, c_2, \ldots, c_n) \in \mathbb{F}_2^n$. First we will
$w = C(c_1, c_2, \ldots c_n)$, $w \in (\mathbb{F}_{2^t})^{m^h}$ be an encoding
on. Next let us extend our codeword to be a concate
$\circ w_{q-1} = w \circ w \circ \ldots \circ w$. Now we will ask the first
w_1 and so on. Note that this does not harm the pr
decreases the rate by a factor q (which is negligible
orithm from the previous section uses some linear co
combination to be over $\mathbb{F}_2$. Let $P(x) = a_0 + a_1 x^{b_1}$ -
polynomial. Next let us now set our codeword to be

$$\tilde{w}_0 \circ \tilde{w}_1 \circ \ldots \circ \tilde{w}_{q-1} \triangleq a_0 w \circ a_1 w \ldots \circ a_{q-1} w,$$

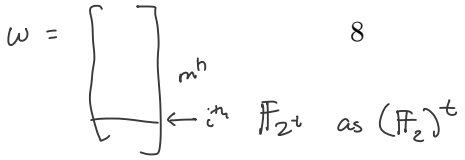
and $a_i w$ is a coordinate wise scalar multiplication
decode the i-th symbol c_i using the identity:

$$c_i \gamma^{<u_i, v>} = \tilde{w}_0(v) + \tilde{w}_1(v + b_1 u_i) + \ldots \tilde{w}_{q-1}(v + b_{q-1} u_i).$$

Now let us take some linear functional $L : \mathbb{F}_{2^t} \mapsto \mathbb{F}_2$ and apply it on every coordinate of our codeword. Then

$$L(c_i \gamma^{<u_i, v>}) = L(\tilde{w}_0(v)) + L(\tilde{w}_1(v + b_1 u_i)) + \ldots L(\tilde{w}_{q-1}(v + b_{q-1} u_i)).$$

We want that $L(c_i \gamma^{<u_i, v>}) = c_i$. If $c_i = 0$ then always $L(c_i \gamma^{<u_i, v>}) = L(0) = 0$ but the problem is that if $c_i = 1$ then it may happen that $L(c_i \gamma^{<u_i, v>}) = L(\gamma^{<u_i, v>}) = 0$. In order



to solve this problem we will not choose v completely at random; we will choose v at random conditioned on $L(\gamma^{<u_i,v>}) = 1$, but this will hurt the smoothness of the code which in turn affects the probability of correct decoding. In order that it will not hurt this probability too much we need to choose L such that for every $i = 1 \dots n$ $\Pr_v(L(\gamma^{<u_i,v>}) = 1) \geq 1/2$.

Lemma 4.1. *There exists a linear functional $L : \mathbb{F}_{2^t} \mapsto \mathbb{F}_2$ such that*

$$\forall i \in [n] \quad \Pr_{v \in (\mathbb{Z}_m)^h}(L(\gamma^{<u_i,v>}) = 1) \geq 1/2.$$

serve that for random v, $\langle u_i, v \rangle$ is a random number in $\mathbb{Z}_m$,since the gcd of u_i's s is 1. Thus it is enough to find L such that

$$\Pr_{j \in \mathbb{Z}_m}(L(\gamma^j) = 1) \geq 1/2.$$

ant j and a random L, $\Pr(L(\gamma^j) = 1) = 1/2$ thus, the expectation of $\Pr_{j \in \mathbb{Z}_m}(L(\gamma^j) =$ e.

$$\mathbf{E}_L(\Pr_{j \in \mathbb{Z}_m}(L(\gamma^j) = 1)) = 1/2.$$

there exists an L such that

$$\Pr_{j \in \mathbb{Z}_m}(L(\gamma^j) = 1) \geq 1/2.$$

□

Let us describe the reduction formally.
Choose L such that $\Pr_{j \in \mathbb{Z}_m}(L(\gamma^j) = 1) \geq 1/2$. Since m is constant we can find it by exhaustive search in constant time.

1. Given a message $(c_1, c_2, \dots c_n)$ encode it , by code from previous section $w = C(c_1, c_2, \dots, c_n)$.
2. Extend it to
$$\tilde{w} \triangleq \tilde{w}_0 \circ \tilde{w}_1 \circ \dots \circ \tilde{w}_{q-1} \triangleq a_0 w \circ a_1 w \dots \circ a_{q-1} w.$$
3. Reduce the alphabet by applying L on every symbol of $\tilde{w}$ and return
$$w_0 \circ w_1 \circ \dots \circ w_{q-1} \triangleq L(\tilde{w}_0) \circ L(\tilde{w}_1) \circ \dots \circ L(\tilde{w}_{q-1}).$$

Let us define the decoding algorithm $d_i(w)$:

- Choose $v \in (\mathbb{Z}_m)^h$ at random conditioned on $L(\gamma^{<u_i,v>}) = 1$.
- Query $w_0(v), w_1(v + b_1 u_i), \dots, w_{q-1}(v + b_{q-1} u_i)$.
- Output $c_i = w_0(v) \oplus w_1(v + b_1 u_i) \dots \oplus w_{q-1}(v + b_{q-1} u_i)$.

Algorithm 2: Decoding Algorithm

Theorem 4.2. *The binary code C defined above is $(q, \delta, 2q\delta)$ locally decodable.*

δ is a fraction

we probe q times, so diverging a dmg place with probability 2qδ.

Proof. We will prove it in two steps.

First let us prove that if at most δ fraction of the codeword $w = w_0 \circ w_1 \ldots \circ w_{q-1}$ is damaged then we query a damaged place with probability at most $2q\delta$. Let δ_i be a fraction of damaged bits in w_i so $\frac{1}{q}\sum \delta_i = \delta$. We chose L such that v is distributed uniformly among half of all possible values. Therefore, the probability that query i will be damaged is at most $2\delta_i$. So the probability that one of the queries will be damaged is at most $\sum 2\delta_i = 2q\delta$.

Next let us prove that if we query only non-damaged places then we will return a correct answer. As before, by linearity it is enough to prove that $d_i(C(e_i)) = 1$ and $d_i(C(e_j)) = 0$ for $i \neq j$.

$$\begin{gathered} d_i(C(e_i)) = L(a_0\gamma^{<u_i,v>}) \oplus L(a_1\gamma^{<u_i,v+b_1u_i>}) \ldots \oplus L(a_{q-1}\gamma^{<u_i,v+b_{q-1}u_i>}) = \\ = L\left(\sum_{j=0}^{q-1} a_j\gamma^{<u_i,v+b_ju_i>}\right) = L\left(\sum_{j=0}^{q-1} a_j\gamma^{<u_i,v>}\right) = \\ L(P(1)\gamma^{<u_i,v>}) = L(\gamma^{<u_i,v>}) \end{gathered}$$

But we choose v such that $L(\gamma^{<u_i,v>}) = 1$. In the same way we can prove that if $C = C(e_j)$ then $c_i = 0$.

$$\begin{gathered} c_i = L(a_0\gamma^{<u_j,v>}) \oplus L(a_1\gamma^{<u_j,v+b_1u_i>}) \ldots \oplus L(a_{q-1}\gamma^{<u_j,v+b_{q-1}u_i>}) = \\ L\left(\gamma^{<u_j,v>}\sum_{t=0}^{q-1} a_t\gamma^{b_t<u_j,u_i>}\right) = L\left(P(\gamma^{<u_i,u_j>})\gamma^{<u_i,v>}\right) = \\ L(0) = 0. \end{gathered}$$

□

We want to mention here that using techniques from [Woo08] Section 5 we can reduce the probability of error to $(q, \delta, q\delta + \varepsilon)$ for any constant $\varepsilon > 0$.

5 Future work

In this paper we give a general construction of LDCs for any S-matching sets and S-decoding polynomials. Any improvement in size of a set-system with restricted intersections will immediately yield improvement in the rate of LDCs. We hope that this paper will give a motivation for future work on set-systems with restricted intersections. We also believe that it is possible to choose an S-decoding polynomial with less monomials as in Example 3.7.

Acknowledgements

I am indebted to Irit Dinur for many helpful in-depth technical discussions and helping me at all stages of this work. I would also like to thank to Venkatesan Guruswami for directing me to [Gro00] and to Ariel Gabizon, Oded Goldreich, Shachar Lovett, Omer Reingold, David Woodruff, and my wife Rivka for their valuable comments.

References

[Amb97] Andris Ambainis. Upper bound on communication complexity of private information retrieval. In Pierpaolo Degano, Roberto Gorrieri, and Alberto Marchetti-Spaccamela, editors, *ICALP*, volume 1256 of *Lecture Notes in Computer Science*, pages 401–407. Springer, 1997.

[BIK05] Amos Beimel, Yuval Ishai, and Eyal Kushilevitz. General constructions for information-theoretic private information retrieval. *J. Comput. Syst. Sci.*, 71(2):213–247, 2005.

[CGKS95] Benny Chor, Oded Goldreich, Eyal Kushilevitz, and Madhu Sudan. Private information retrieval. In *FOCS*, pages 41–50, 1995.

[Gas04] William I. Gasarch. A survey on private information retrieval (column: Computational complexity). *Bulletin of the EATCS*, 82:72–107, 2004.

[GKST02] Oded Goldreich, Howard J. Karloff, Leonard J. Schulman, and Luca Trevisan. Lower bounds for linear locally decodable codes and private information retrieval. In *IEEE Conference on Computational Complexity*, pages 175–183, 2002.

[GKST06] Oded Goldreich, Howard J. Karloff, Leonard J. Schulman, and Luca Trevisan. Lower bounds for linear locally decodable codes and private information retrieval. *Computational Complexity*, 15(3):263–296, 2006.

[Gro00] Vince Grolmusz. Superpolynomial size set-systems with restricted intersections mod 6 and explicit ramsey graphs. *Combinatorica*, 20(1):71–86, 2000.

[Ito99] Toshiya Itoh. Efficient private information retrieval. *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences Vol.E82-A No.1 pp.11-20*, 1999.

[KdW03] Iordanis Kerenidis and Ronald de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. In *STOC*, pages 106–115. ACM, 2003.

[KT00] Jonathan Katz and Luca Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *STOC*, pages 80–86, 2000.

[KY08] Kiran S. Kedlaya and Sergey Yekhanin. Locally decodable codes from nice subsets of finite fields and prime factors of mersenne numbers. In *IEEE Conference on Computational Complexity*, pages 175–186. IEEE Computer Society, 2008.

[Man98] Eran Mann. Private access to distributed information. In *Master's thesis, Technion - Israel Institute of Technology*, 1998.

[Rag07] Prasad Raghavendra. A note on yekhanin's locally decodable codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 2007.

[RY07] Alexander A. Razborov and Sergey Yekhanin. An omega($^{1/3}$) lower bound for bilinear group based private information retrieval. *Theory of Computing*, 3(1):221–238, 2007.

[Tre04] Luca Trevisan. Some applications of coding theory in computational complexity. *Electronic Colloquium on Computational Complexity (ECCC)*, (043), 2004.

[WdW05] Stephanie Wehner and Ronald de Wolf. Improved lower bounds for locally decodable codes and private information retrieval. In *ICALP*, pages 1424–1436, 2005.

[Woo07] David Woodruff. New lower bounds for general locally decodable codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 2007.

[Woo08] David P. Woodruff. Corruption and recovery-efficient locally decodable codes. In *APPROX-RANDOM*, pages 584–595, 2008.

[Yek08] Sergey Yekhanin. Towards 3-query locally decodable codes of subexponential length. *J. ACM*, 55(1), 2008.

A A simple construction of S-matching vectors

Lemma A.1. *Let $p_1 < p_2 \ldots < p_r$ be any r primes and $m = p_1 \cdot p_2 \ldots p_r$. Then for every t, there exists a set S of size $2^r - 1$ and a family of S-matching vectors $\{u_i\}_{i=1}^n, u_i \in (\mathbb{Z}_m)^h$ such that $n = \binom{t}{m-1}$ and $h = O(t^{p_r - 1})$.*

Proof. Let us first construct a family of vectors $\{u'_i\}_{i=1}^n, u'_i \in (\mathbb{Z}_m)^{t+1}$ such that:

1. $\langle u'_i, u'_i \rangle = 0$ for $i \in [n]$.
2. $\langle u'_i, u'_j \rangle \neq 0$ for $i \neq j$.

Identify the subsets of $[t] = \{1, 2, \ldots t\}$ of size $m-1$ with $\{1, \ldots, \binom{t}{m-1}\}$. For every subset $A \subseteq [t]$ of size $m-1$, let $u'_i \in \mathbb{Z}_m^t$ be the indicator vector of the set, i.e., $u'_i = (a_1, a_2, \ldots a_t)$, where $a_i = 1$ if $i \in A$ and $a_i = 0$ otherwise. In order to simplify the construction let us add an additional coordinate which is always one i.e., $u'_i = (a_1, a_2, \ldots a_t, 1)$. Clearly $\langle u'_i, u'_i \rangle = 0$ since u'_i has exactly m ones and $\langle u'_i, u'_j \rangle = 1 + |A_i \cap A_j| \neq 0$. Since intersection of two different subsets of size $m-1$ is always less than $m-1$.

Now we want to change these vectors such that the inner product of two such vectors will be in some small set S. By the chinese reminder theorem $\mathbb{Z}_m \approx \mathbb{Z}_{p_1} \oplus \mathbb{Z}_{p_2} \ldots \oplus \mathbb{Z}_{p_r}$. Thus any number x in $\mathbb{Z}_m$ we can view as $(x \mod p_1, x \mod p_2, \ldots, x \mod p_r)$. The set S is the set $\{0,1\}^r \backslash (0, 0, \ldots 0)$ i.e. $a \in S$ iff $a \neq 0$ and for every $k = 1, \ldots r$ holds $(a \mod p_k) \in \{0, 1\}$.

By the chinese reminder theorem there exist constants $c_1, c_2 \ldots c_r \in \mathbb{Z}_m$ such that:

1. $c_i \equiv 1 \mod p_i$
2. $c_i \equiv 0 \mod p_j$ for $i \neq j$

Let us define u_i by:

$$u_i = (c_1 u_i'^{\otimes p_1 - 1}, c_2 u_i'^{\otimes p_2 - 1}, \ldots, c_r u_r'^{\otimes p_r - 1}).$$

Now we need to prove that: $\langle u_i, u_i \rangle \equiv 0$:

$$\langle u_i, u_i \rangle = \langle (c_1 u_i'^{\otimes p_1 - 1}, c_2 u_i'^{\otimes p_2 - 1}, \ldots c_r u_r'^{\otimes p_r - 1}), (c_1 u_i'^{\otimes p_1 - 1}, c_2 u_i'^{\otimes p_2 - 1}, \ldots c_r u_r'^{\otimes p_r - 1}) \rangle =$$
$$\sum_{j=1}^{r} c_j^2 \langle u_i'^{\otimes p_j - 1}, u_i'^{\otimes p_j - 1} \rangle = \sum_{j=1}^{r} c_j^2 \langle u_i', u_i' \rangle^{p_j - 1},$$

where the last equation follows from Fact 2.6. Since $\langle u_i', u_i' \rangle = 0$ it follows that $\langle u_i, u_i \rangle = 0$. Now let us prove that $\langle u_i, u_j \rangle \in S$ for any $i \neq j$. In order to prove that $\langle u_i, u_j \rangle \in S$ we will prove that $\langle u_i, u_j \rangle \mod p_k \in \{0, 1\}$ and $\langle u_i, u_j \rangle \neq 0$. Observe that

$$u_i \mod p_k \equiv (0, 0, \ldots, u_i'^{\otimes (p_k - 1)}, 0, \ldots, 0).$$

Thus it follows that:

$$\langle u_i, u_j \rangle \mod p_k \equiv \langle u_i'^{\otimes p_k - 1}, u_j'^{\otimes p_k - 1} \rangle \equiv \langle u_i', u_j' \rangle^{p_k - 1}$$

By Fermat's Little Theorem $x^{p_k - 1} \equiv 0$ or $1 \mod p_k$ for every k. Since $\langle u_i', u_j' \rangle \neq 0 \mod m$ for some k $\langle u_i', u_j' \rangle \neq 0 \mod p_k$. Therefore $\langle u_i, u_j \rangle = \langle u_i', u_j' \rangle^{p_k - 1} \neq 0 \mod p_k$. Therefore $\langle u_i, u_j \rangle \neq 0 \mod m$. □

As a corollary we get:

Corollary A.2. *For every h, r there exists integer $m = p_1 p_2 \ldots p_r$ and a set $S \subset \mathbb{Z}_m$ of size $2^r - 1$ and a family of S-matching vectors $\{u_i\}_{i=1}^n, u_i \in (\mathbb{Z}_m)^h$ such that $n \geq \exp(c \frac{(\log h)^r}{\log \log^{r-1} h})$.*

Note that the only difference between Corollary A.2 and Corollary 3.3 is in order of quantifiers i.e. Corollary 3.3 holds for every m while Corollary A.2 holds for some specific m.

Proof of Corollary A.2. Let us take all primes of the same size (i.e. $p_i = p_j + o(p_i)$) and $t = m^2$ then in Lemma A.1 we will get that $n \geq \binom{m^2}{m-1} \geq m^m = O(m^{p^r})$ and $h = O(m^{2p_r})$. Thus it follows that:

$$n \geq \exp(c \frac{(\log h)^r}{\log \log^{r-1} h}).$$

□

Summary (assume $m = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$)

$S = \{ s \in \mathbb{Z}_m \mid s \bmod p_i = \{0, 1\} \ \forall i \in [r] \}$, by CRT, $\exists$ $2^r - 1$ non-zero such integers $\in \mathbb{Z}_m$.

S-matching Vectors

set of vectors such that $\langle v_i, v_i \rangle = 0 \ \forall i \in [n]$ and $\langle v_i, v_j \rangle = 1 \ \forall i \neq j$, $i, j \in [n]$

$\{v_i\}_{i=1}^n$, $v_i \in (\mathbb{Z}_m)^h$

S-decoding Polynomial (P)

$P(1^s) = 0 \quad \forall s \in S$ — we don't care abt anything else

$P(1^0) = 1$

rmk: for any such S, $\exists$ a S-decoding Polynomial w/ at most $|S| + 1$ terms

Codeword

$C : \mathbb{F}^n \to \mathbb{F}^{m^h}$ (generates code words)

let $a = (a_1, \dots, a_n)$ be a sample input

$$C(a) = C(a_1 \cdot e_1 + a_2 e_2 + \dots + a_n e_n)$$
$$= a_1 C(e_1) + \dots + a_n C(e_n)$$

$$C(e_i) = \begin{bmatrix} \gamma^{\langle v_i, x_1 \rangle} \\ \gamma^{\langle v_i, x_2 \rangle} \\ \vdots \\ \gamma^{\langle v_i, x_{m^h} \rangle} \end{bmatrix}$$

fix v_i and iterate over all possible indicator vectors

$$= \sum_{i=1}^n a_i \cdot C(e_i)$$

$$= \omega_a$$

Decoder $d_i(C(x)) \Rightarrow x_i$

Start at a random vector $(v \in (\mathbb{Z}_m)^h)$

$\omega(v), \omega(v + b_1 v_i), \dots, \omega(v + b_{q-1} v_i)$ — we query uniformly starting at a random spot.

we perform q queries ($q \leq 2^r$ b/c P has $2^r - 1 + 1$ terms)

Output:

$c_i = \gamma^{-\langle v_i, v \rangle} (a_0 \omega(v) + a_1 \omega(v + b_1 v_i) + \dots + a_{q-1} \omega(v + b_{q-1} v_i))$

$d_i(C(x_i)) = 1$ and $d_i(C(x_j)) = 0$ (verified)

therefore $d_i(C(x)) = x_1 \cdot d_i(C(e_1)) + \dots + x_n \cdot d_i(C(e_n))$

$= x_1 \cdot 0 + \dots + x_i \cdot 1 + \dots + x_n \cdot 0$

$= x_i$

$\Rightarrow$ perfectly smooth decoder which makes q queries is also $(q, \delta, \varepsilon\delta)$ locally decodable

(q queries; δ of them we corrupt; $\varepsilon\delta$ = probability of hitting a corrupt spot)

Binary locally decodable codes

Before, we had outputs in $\mathbb{F}^{m^h}$. They could be anything over a huge field.

Assume $C = \{0,1\}^n$, pretend $C \in \mathbb{F}_2{}^n$

let $\omega = C(c)$, $\omega \in (\mathbb{F}_2)^{m^h}$ (from previous construction)

$$\omega = \begin{bmatrix} \gamma^{\langle v_i, x_{j_1} \rangle} \\ \gamma^{\langle v_i, x_{j_2} \rangle} \\ \vdots \\ \gamma^{\langle v_i, x_{j_{m^h}} \rangle} \end{bmatrix}$$

$$c_i \cdot \gamma^{\langle v_i, v \rangle} = \omega_0(v) + \omega_1(v + b_1 v_i) + \dots + \omega_{2^t-1}(v + b_{2^t-1} v_i)$$

$$L(c_i \cdot \gamma^{\langle v_i, v \rangle}) = L(\omega_0(v) + \dots + L(\omega_{2^t-1}(v + b_{2^t-1} v_i)))$$

if $c_i = 0$, $L(0) = 0$ b/c linearity

if $c_i = 1$, $L(\gamma^{\langle v_i, v \rangle})$ could be 0 or 1 depending on v. We limit v to the part in which $L(\gamma^{\langle v_i, v \rangle}) = 1$ for any v_i. (larger error)

Presentation

1) Quick introduction to Grolmusz (statement of main theorem)

2) Odd-town explanation (aka: why Grolmusz is cool)

3) Start building Grolmusz
   i) background of OR_n polynomials
   ii) construction of Lemma 3.2
   iii) proof / connection to theorem 1.2 (main)

Odd-town Explanation (Set with prime 2)

Odd-town has 32 inhabitants who want to form clubs. The city passed 2 laws

1) each club needs an odd number of members.
2) each club shares an even number of members

We can form 32 clubs b/c each person can form their own club. But ∄ a way to form 33

Suppose $C_1, \dots, C_m$ are clubs with associated indicator vectors $v_1, \dots, v_m$ s.t. $v_i \in \{0,1\}^{32}$ (does citizen #j belong to club i?)

$$|C_i \cap C_j| \Rightarrow \underbrace{v_i \cdot v_j}_{\text{(in compliance w/ rules)}} = \begin{cases} \text{odd}, & \text{if } i=j \\ \text{even}, & i \neq j \end{cases} \xrightarrow{\text{mod 2}} v_i \cdot v_j = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}$$

WTS: $v_i \ \forall i \in [1,32]$ must be linearly independent over $\mathbb{F}_2$ under these conditions.

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_{32} v_{32} = 0 \quad (\lambda_i \in \mathbb{F}_2) \quad \Big\} \text{ dot } v_1$$
$$= \lambda_1 v_1 \cdot v_1 + \lambda_2 v_2 \cdot v_1 + \dots + \lambda_{32} v_{32} \cdot v_1 = 0 \cdot v_1$$
$$= \lambda_1 1 = 0 \quad (v_i \cdot v_j = 0 \text{ if } i \neq j) \ \square$$

⇒ 32 is an arbitrary number, therefore assuming laws are followed, the indicator vectors of all the clubs are linearly independent over $\mathbb{F}_2$ ☐

WTS: In a town of n citizens, no more than n clubs may be formed under laws

fact: $\text{rank}(AB) \leq \min\{\text{rank } A, \text{rank } B\}$

Let v_i be the rows of M, a $m \times n$ matrix (Incidence Matrix)

Let $A = M \cdot M^T$,
$$A = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}_{m \times n} \cdot \begin{bmatrix} | & | & & | \\ v_1^T & v_2^T & \cdots & v_m^T \\ | & | & & | \end{bmatrix}_{n \times m} = \begin{bmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 & \cdots & v_1 \cdot v_m \\ v_2 \cdot v_1 & v_2 \cdot v_2 & \cdots & v_2 \cdot v_m \\ \vdots & \vdots & \ddots & \vdots \\ v_m \cdot v_1 & v_m \cdot v_2 & \cdots & v_m \cdot v_m \end{bmatrix}_{m \times m} \cdot \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

$\text{rank}(M \cdot M^T) \leq \min\{\text{rank } M, \text{rank } M^T\}$

$\Rightarrow M \leq \text{rank } M = \text{rank } M^T \leq n.$ ☐

OR-Polynomials

A polynomial $P(x_1, \dots, x_n)$ represents OR_n mod m if

1) $p(0,0,\dots,0) = 0 \mod m$
2) $p(x_1, \dots, x_n) \neq 0 \mod m \quad \forall x \neq \vec{0}$

$OR_n(x) = P(x) \ \forall x, \ P(x) \in \{0,1\}$

Strong vs Weak vs One sided → $OR_n(x) = 0$ iff $P(x) = 0 \mod m$; $OR_n(x) = 1$ iff $P(x) \neq 0 \mod m$

↳ $\forall x \neq y, \ P(x) \neq P(y)$

$\Rightarrow \exists \ S \leq 2m$ s.t. $OR_n(x) = 0$ iff $P(x) \in S$

*rmk: we can switch between strong and weak representations without loss/gain of degrees

big O with constant depending on parameter m

Let $m = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$, then $\exists$ a polynomial representation of degree $O_m(n^{1/t})$

Reason why composite numbers have lower degrees:
we can use CRT to design polynomial reps mod $p_1, p_2, \dots, p_r$ and configure them together

Grolmusz overview

From CR_n reps mod m, we get degree $n^{1/r}$ polynomials.

This allows us to form $P(z_1, \dots, z_n)$ s.t. $P(\vec{z}) \equiv 0 \bmod m \iff \vec{z} = \vec{0}$
and therefore $Q(\vec{z}) = P(1 - \vec{z}) \equiv 0 \bmod m$ iff $\vec{z} = \vec{1}$.

Lemma 3.2

$\forall n > 0$, $\exists$ uniform set system $\mathcal{H}$ over $2(m-1)n^{2d}/d!$ elements s.t.

1) $|\mathcal{H}| = n^n$
2) $\forall H \in \mathcal{H}$, $|H| \equiv 0 \bmod m$
3) $\forall G, H \in \mathcal{H}$, $|H \cap G| \not\equiv 0 \bmod m$

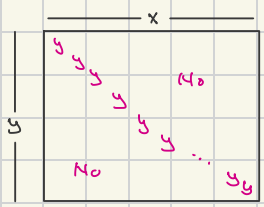
a bunch of monomials of some multiplicity

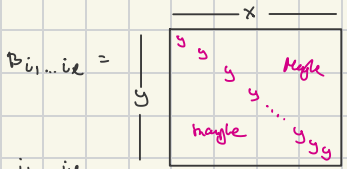
$$\tilde{Q}(z_1, \dots, z_n) = \sum_{i_1, \dots, i_d \in n} \tilde{a}_{i_1, \dots, i_d} \cdot (z_{i_1} \cdot \dots \cdot z_{i_d})$$

$x, y \in \{0, 1, 2, \dots, n-1\}^n$

$\equiv 0$ if all $x_i = y_i \Rightarrow x = y$
$\neq 0$ o/w

$$a_{xy} = \tilde{Q}(\delta(x_1, y_1), \delta(x_2, y_2), \dots)$$





$B_{i_1 \dots i_d} =$

do the indices agree $\forall$ combos?

$$B^{i_1, \dots, i_d}_{xy} = \delta(x_{i_1}, y_{i_1}) \cdot \delta(x_{i_2}, y_{i_2}) \cdot \dots \cdot \delta(x_{i_d}, x_{i_d})$$

Build Hypergraph $\Longrightarrow$ Dual (gives set system of 3.2)

$\Longrightarrow$ 1.2

start with an n, we "find" an n via 3.2 to give a close enough universe $\approx n$.

$|\mathcal{H}| = n^n$, to prove $n^n \geq \exp\left(\frac{c(\log n)^2}{\log\log n}\right)$ we use Stirling approx

Verification ($r = 2$)

lets say we want $h = \frac{n^{2\sqrt{n}}}{(\sqrt{n})!}$

Via 3.2, for n, $\exists$ $\mathcal{H}$ over $2(m-1)n^{2d}/d!$ elements ($d = O(n^{1/r}) \Rightarrow O(n^{1/2})$ due to degree of OR_n representation of a m w/ n prime factors)

$\Rightarrow$ n gives us $\mathcal{H}$ over $\underbrace{2(m-1)}_{\text{ignore b/c constant}} \cdot \frac{n^{2(\sqrt{n})}}{(\sqrt{n})!}$

3.2 says $|\mathcal{H}| = n^n$.

By stirling's approximation: $\ln(n!) = n\ln(n) - n + O(\ln n)$

Prove $n^n \geq \exp\left(\frac{c \log^2 h}{\log\log h}\right) \Rightarrow \underbrace{n \log(n) \geq \frac{c \log^2 h}{\log\log h}}_{\text{WTS}}$

$\log h = 2\sqrt{n}\log(n) - \log(\sqrt{n}!)$

$= 2\sqrt{n}\log(n) - [\sqrt{n}\log(\sqrt{n}) - \sqrt{n}]$

$= 2\sqrt{n}\log(n) - \frac{1}{2}\sqrt{n}\log(n) - \sqrt{n}$

$= 1.5\sqrt{n}\log(n)$

$\log\log h = \log(1.5\sqrt{n}) + \log\log(n)$

$= \underbrace{\log(1.5)}_{\text{constant}} + \frac{1}{2}\log(n) + \log\log(n)$

$\frac{c \cdot [1.5\sqrt{n}\log(n)]^2}{(\log(1.5) + \frac{1}{2}\log(n) + \log\log(n))}$ $\log(n) > \log\log(n)$

$\geq \frac{c \cdot 1.5^2 \cdot n \cdot (\log^2(n))}{1.5\log(n)}$

$= c \cdot n\log(n)$ ∎ c is < 1

Set families (odd-town)

1) how many sets over n of size k s.t. they all intersect?

$\leq \binom{n-1}{k-1}$, we fix one element and forcefully add it to all sets.

2) what if we require even sized sets with even intersections? (even-town)

Assume 32 residents of even-town. We can put them into pairs giving us 16 pairs

Each club can be represented by an indicator vector $v \in \{0,1\}^{16}$

$\Rightarrow \exists\ 16^2$ indicator vectors, all representing a unique club, very large number

3) What if we require odd sized sets w/ even intersections? (odd-town)

The number decreases dramatically to n (32)

$$m\begin{bmatrix} \overset{n}{\vdots} \\ v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix} \cdot \begin{bmatrix} \overset{m}{\vdots} \\ v_1^T \\ \vdots \\ v_m^T \\ \perp \end{bmatrix}_n = m\begin{bmatrix} v_1 \cdot v_1 & \cdots \\ \vdots & \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ & \ddots \end{bmatrix}$$

rank $\leq n$ rank $\leq n$ rank m

$\Rightarrow m \leq n$ b/c rank $[AB] \leq \min\{\text{rank } A, \text{rank } B\}$

$\Rightarrow$ $v_i \cdot v_j = 0$ $i \neq j$. non-zero orthogonal $\Rightarrow$ linearly independent (m) $\Rightarrow$ $\mathbb{F}_2^n$ has max n linearly independent vectors

3b) even sized sets with odd intersections (similar proof)

$\vec{v}_i \in \{0,1\}^n$, $i \in [1,\dots,m] \to v_i \cdot v_j = 1$ if $i \neq j$

what we want $\longrightarrow$ $\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \ddots & 1 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \cdots & 1 & 0 \end{bmatrix}_m = m\begin{bmatrix} \overset{n}{\vdots} \\ v_1 \\ \vdots \\ v_m \end{bmatrix} \begin{bmatrix} \overset{m}{\vdots} \\ v_1^T \\ \vdots \\ v_m^T \\ \perp \end{bmatrix}_n$

A $\quad M$ $\quad M^T$

Rank $(A) \leq \min\{\text{rank } M, \text{rank } M^T\} \leq n$

Rank $A = m \leq n$

4) we see the same thing with any prime.

This argument breaks down for mod m b/c $\mathbb{Z}_m$ is not a field

you can have non-zero vectors who's dot = 0 mod m without linear independence.

$\Rightarrow$ we can find even more sets that satisfy mod m w/o needing to limit ourselves to linear independence

5) Grolmosz (using OR_n rep mod m we get $n^{1/r}$ polynomial)

Theorem 1.2. Let m be a positive integer, and suppose that m has $r>0$ different prime divisors : $m = p_1^{a_1} p_2^{a_2} \dots p_r^{a_r}$. Then $\exists\ c = c(m) > 0$ s.t. $\forall\ n>0$, $\exists$ uniform set-system $\mathcal{H}$ over n elements s.t.

1) $|\mathcal{H}| \geq \exp\left(c \frac{(\log n)^r}{(\log\log n)^{r-1}}\right)$
2) $\forall H \in \mathcal{H}$, $|H| \equiv 0 \bmod m$
3) $\forall G \neq H \in \mathcal{H}$, $|G \cap H| \not\equiv 0 \bmod m$

* $c \sim p_s^{-r} < 1$

Lemma 3.2

$\forall n>0$, $\exists$ uniform set system $\mathcal{H}$ over $2(m-1)n^{2d}/d!$ elements s.t.

1) $|\mathcal{H}| = n^n$
2) $\forall H \in \mathcal{H}$, $|H| = 0 \bmod m$
3) $\forall H \neq G \in \mathcal{H}$, $|H \cap G| \not\equiv 0 \bmod m$

From OR_n reps mod m, we get degree $n^{1/r}$ polynomials.

This allows us to form $P(z_1, \dots, z_n)$ s.t. $P(\vec{z}) \equiv 0 \bmod m \iff \vec{z} = \vec{0}$ and therefore $Q(\vec{z}) = P(1 - \vec{z}) \equiv 0 \bmod m$ iff $\vec{z} = \vec{1}$.

all coefficients modded down

$$\tilde{Q}(\vec{z}) = \sum_{i_1,\dots,i_L}^{\prime} \tilde{\alpha}_{i_1,\dots,i_L} \cdot z_{i_1} z_{i_2} \dots z_{i_L}$$

AND

bunch of arbitrary monomials

let $\delta(x_1, x_2) = \begin{cases} 1, & x_1 = x_2 \\ 0, & o/w \end{cases}$

let $x, y \in \{0, 1, \dots, n-1\}^n$, n^n such vectors

$A \in M_{n^n \times n^n}(\mathbb{Z})$

$$A = \begin{bmatrix} 0 & & & \neq 0 \\ & 0 & & \\ & & \ddots & \\ \neq 0 & & & 0 \end{bmatrix}_y^x$$

entry $A_{xy} = \tilde{Q}(\delta(x_1, y_1), \delta(x_2, y_2), \dots, \delta(x_n, y_n))$

↳ 0 if $x = y$, $\neq 0$ o/w ($\tilde{Q}(\vec{1}) = 0 \bmod m$)

breach up

$$B_{i_1,\dots,i_L} = \begin{bmatrix} 1 & & & \{0,2\} \\ & 1 & & \\ & & \ddots & \\ \{0,2\} & & & 1 \end{bmatrix}_y^x$$

$B_{xy}^{i_1,\dots,i_L} = \delta(x_{i_1}, y_{i_1}), \dots, \delta(x_{i_L}, y_{i_L})$

↳ 1 if those components of x any y match

All B blocks output 1 for coordinate $A_{xy} \Rightarrow \tilde{Q}(\vec{1}) = 0 \bmod m \Rightarrow x = y$

We build Hypergraph

$$A = \begin{bmatrix} 0 & & x \\ & 0 & \neq 0 \\ \neq 0 & 0 & \ddots \\ & & 0 \end{bmatrix} y$$

Choose all possible vectors of $\{0, 1, \dots, n-1\}^n$ as vertices

Choose all B-Blocks intersections as edges

* This the the absolute max # of edges / terms / B-Blocks

max multiplicity of all possible combinations of all degrees

How many edges / B-Blocks exist?

$$\tilde{G}(n, n, \dots, n) = \sum \underbrace{\alpha_{i_1, \dots, i_k}}_{\text{coefficient of term}} \cdot n^{\overbrace{k}^{\text{degree of term}}}$$

max coefficient

$$\leq (m-1) \sum_{k \leq d} \binom{n}{k} n^k$$

* increase num decrease denom

$$\frac{n! \, n^k}{k! (n-k)!} < \frac{n^k \, n^k}{k!}$$

Universe

$$< (m-1) \sum_{k \leq d} n^{2k} / k!$$

$$= \frac{n^{2(1)}}{1} + \frac{n^{2(2)}}{2!} + \dots + \frac{n^{2(d)}}{d!}$$

$$< \frac{n^{2d}}{d!}$$

$$< (m-1) \frac{2n^{2d}}{d!}$$

How many edges per vertex?

$$\tilde{G}(1, 1, \dots, 1) \leq (m-1) \sum_{k \leq d} \binom{n}{d} \leq 2(m-1) \binom{n}{d}$$

Now we build dual of the hypergraph

Universe: All the B-Blocks $< 2(m-1) n^{2d} / d!$

Sets: diagonals of A: $(x, x) \Rightarrow |\mathcal{U}| = n^n$

↳ each set contains all the B-Blocks which cover it

e.x. $x = (0, 1, 1, 2)$, $y = (2, 3, 1, 2)$

B-Blocks that dictates $e_2 = 2$ & $e_4 = 2$ covers both so it's part of both sets

B-Blocks that dictates $e_1 = 0$ & $e_3 = 1$ covers only x so it's not part of y-set

Size: Each diagonal cell is covered by the same # of B-Blocks (may not be the same ones)

e.x. $Q = x_2 x_3 + x_1$

$x = (0, 1, 2)$, $y = (2, 1, 0)$

both covered by 2 B-Blocks