An Exposition of a problem in Multiparty Communication Complexity and its Application to Lower Bounds on Branching Program

By William Gasarch

1 Introduction

Multiparty communication complexity was first defined by Chandra, Furst, and Lipton [6] and used to obtain lower bounds on branching programs. Since then it has been used to get additional lower bounds and tradeoffs for branching programs [1, 3], lower bounds on problems in data structures [3], time-space tradeoffs for restricted Turing machines [1], and unconditional pseudorandom generators for logspace [1].

Def 1.1 Let $f : \{\{0,1\}^n\}^k \rightarrow X$. Assume, for $1 \le i \le k$, P_i has all of the inputs *except* x_i. Let $d(f)$ be the total number of bits broadcast in the optimal deterministic protocol for f. At the end of the protocol all parties must know the answer. This is called the *multiparty communication complexity* of f. The scenario is called the *forehead model.*

Note 1.2 Note that there is always the $n+1$-bit protocol of (1) P_1 broadcasts x_2, (2) P_2 computes and broadcasts $f(x_1,\ldots,x_k)$. The cases of interest are when $d(f) \ll n$.

We will need the following lemmas about multiparty protocols. The first one is the $k=3$ case of the second one. We leave it for an exercise.

Lemma 1.3 *Let P be a multiparty protocol for a function $f : \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n \rightarrow X$.*

1. *Let $TRAN$ be a possible transcript of the protocol P. There exists $A_1, A_2, A_3 \subseteq \{0,1\}^n$ such that, for all $x_1, x_2, x_3 \in \{0,1\}^n$ the following holds: The protocol P on input (x_1,x_2,x_3) produces transcript $TRAN$ iff $(x_1,x_2,x_3) \in A_1 \times A_2 \times A_3$.*

2. *Let $x_1, x_2, x_3 \in \{0,1\}^n$, $\sigma_1, \sigma_2, \sigma_3 \in \{\{0,1\}^n\}^3$, $TRAN$ be a transcript. Assume that σ_1 has x_1 as its first element, σ_2 has x_2 as its second element, σ_3 has x_3 as its third element. (In symbols, if $*$ means we don't care about the element, then*

$$\begin{aligned} \sigma_1 &= (x_1, *, *) \\ \sigma_2 &= (*, x_2, *) \\ \sigma_3 &= (*, *, x_3). \end{aligned}$$

) Further assume that $\sigma_1, \sigma_2, \sigma_3$ all produces transcript $TRAN$. Then (x_1,x_2,x_3) produces transcript $TRAN$.

Lemma 1.4 *Let P be a multiparty protocol for a function $f : \{\{0,1\}^n\}^k \rightarrow X$.*

1. *Let $TRAN$ be a possible transcript of the protocol P. There exists $A_1, \ldots, A_k \subseteq \{0,1\}^n$ such that, for all $x_1, \ldots, x_k \in \{0,1\}^n$ the following holds: The protocol P on input $(x_1, \ldots, x_k)$ produces transcript $TRAN$ iff $(x_1, \ldots, x_k) \in A_1 \times \cdots \times A_k$.*

2. *Let $x_1, \ldots, x_k \in \{0,1\}^n$, $\sigma_1, \ldots, \sigma_k \in \{\{0,1\}^n\}^k$, $TRAN$ be a transcript. Assume that σ_i has x_i as its ith element. Further assume that each σ_i produces transcript $TRAN$. Then $(x_1, \ldots, x_k)$ produces transcript $TRAN$.*

We will study the following function.

Def 1.5 Let $n \in \mathbb{N}$. Let $N(n) : \mathbb{N} \to \mathbb{N}$. We define $\mathrm{MOD}_n^{N(n)}$ as follows.

$$\mathrm{MOD}_n^{N(n)}(x,y,z) = \begin{cases} YES & \text{if } x+y+z \equiv 0 \pmod{N(n)} \\ NO & \text{if } x+y+z \not\equiv 0 \pmod{N(n)} \end{cases} \tag{1}$$

Note 1.6 Chandra, Furst, Lipton actually examined the function $\mathrm{EQ}_n^{N(n)}$ which is defined as

$$\mathrm{EQ}_n^{N(n)}(x,y,z) = \begin{cases} YES & \text{if } x+y+z = N(n) \\ NO & \text{if } x+y+z \neq N(n) \end{cases} \tag{2}$$

However, everything we do here is an easy modification of what they have done (unless otherwise noted).

We will first establish a connection between $d(\mathrm{MOD}_n^{N(n)})$ and some concepts in Ramsey Theory. We will then use results from Ramsey Theory to obtain upper and lower bounds on $d(\mathrm{MOD}_n^{N(n)})$. The lower bounds will be applied to obtain lower bounds on branching programs.

Here is what we will show.

1. If $\lim_{n\to\infty} N(n) = \infty$ then $d(\mathrm{MOD}_n^{N(n)}) \geq \omega(1)$ (First proven by Chandra, Furst, Lipton [6].)

2. $d(\mathrm{MOD}_n^{N(n)}) \geq \log\log\log N(n) + \Omega(1)$ (First proven by Beigel, Gasarch, Glenn [5].)

3. $d(\mathrm{MOD}_n^{N(n)}) \leq \sqrt{\log(N(n))}$ (First proven by Chandra, Furst, Lipton [6].)

4. Let $M(m)$ be a function such that $M(m) << m$. Let $\mathrm{mod}_m^{M(m)} : \{0,1\}^m \to \{0,1\}$ be defined by

$$\mathrm{mod}_m^{M(m)}(b_1, \ldots, b_m) = \begin{cases} 1 \text{ if } \sum_{i=1}^m \equiv 0 \pmod{M(m)} \\ 0 \text{ otherwise} \end{cases} \tag{3}$$

   There is no $O(m)$ length, $O(1)$ width Branching Program for $\mathrm{mod}_m^{M(m)}$. (First proven by Chandra, Furst, Lipton [6].)

2 Connections Between Multiparty Comm. Comp. and Ramsey Theory

In this section we review the connections between the multiparty communication complexity of f and Ramsey Theory that was first established in [6].

Def 2.1 Let $c, T \in \mathbb{N}$. We think of $[T]$ as being $\{1, \ldots, T\}$ mod T.

1. A *proper c-coloring of* $[T] \times [T]$ is a function $\text{COL} : [T] \times [T] \rightarrow [c]$ such that there do not exist $x, y \in [T]$ and $\lambda \in [T-1]$ such that

$$\text{COL}(x,y) = \text{COL}(x+\lambda, y) = \text{COL}(x, y+\lambda)$$

   (all of the additions are mod T). Another way to look at this: In a proper coloring there cannot be three vertices that (a) are the same color, and (b) are the corners of a right isosceles triangle with legs parallel to the axes and hypotenuse parallel to the line $y = -x$.)

2. Let $\chi(T)$ be the least c such that there is a proper c-coloring of $[T] \times [T]$.

Theorem 2.2 *Let* $N(n) : \mathbb{N} \rightarrow \mathbb{N}$.

1. $d(\text{MOD}_n^{N(n)}) \leq \lg(\chi(N(n))) + O(1)$.
2. $d(\text{MOD}_n^{N(n)}) \geq \lg(\chi(N(n)) + \Omega(1)$.

Proof:
1) Let COL be a proper c-coloring of $[N(n)] \times [N(n)]$. We represent elements of $[c]$ by $\lg(\chi(N(n))) + O(1)$ bit strings. P_1, P_2, P_3 will all know COL ahead of time. We present a protocol for this problem for which the communication is $2\lg(\chi(N(n))) + O(1)$. We will then show that it is correct.

1. P_1 has y, z. P_2 has x, z. P_3 has x, y.
2. P_1 calculates x' such that $x' + y + z \equiv 0 \pmod{N(n)}$. P_1 broadcasts $\sigma_1 = \text{COL}(x', y)$.
3. P_2 calculates y' such that $x + y' + z \equiv 0 \pmod{N(n)}$. P_2 broadcasts 1 if $\sigma_2 = \text{COL}(x, y')$, 0 otherwise.
4. P_3 looks up $\sigma_3 = \text{COL}(x, y)$. P_3 broadcasts YES if $\sigma_1 = \sigma_2 = \sigma_3$ and NO otherwise. (We will prove later that these answers are correct.)

Claim 1: If $\mathrm{MOD}_n^{N(n)}(x,y,z)=YES$ then P_1, P_2, P_3 will all think $\mathrm{MOD}_n^{N(n)}(x,y,z)=YES$.

Proof: If $\mathrm{MOD}_n^{N(n)}(x,y,z)=YES$ then $x_1'=x_1$, $x_2'=x_2$, and $x_3'=x_3$. Hence $\sigma_1=\sigma_2=\sigma_3$ Therefore P_1, P_2, P_3 all think $\mathrm{MOD}_n^{N(n)}(x,y,z)=YES$.
End of proof of Claim 1.

Claim 2: If P_1, P_2, P_3 all think that $\mathrm{MOD}_n^{N(n)}(x,y,z)=YES$ then $\mathrm{MOD}_n^{N(n)}(x,y,z)=YES$.

Proof: Assume that P_1, P_2, P_3 all think $\mathrm{MOD}_n^{N(n)}(x,y,z)=YES$.

Hence

$$\mathrm{COL}\ (x_1,x_2) = \ \mathrm{COL}\ (x_1',x_2) = \ \mathrm{COL}\ (x_1,x_2').$$

We call this **The Coloring Equation.**

Assume

$$x_1+x_2+x_3 \equiv \lambda \pmod{N(n)}.$$

We show that $\lambda \equiv N(n) \equiv 0 \pmod{N(n)}$.

By the definition of x_1'

$$x_1'+x_2+x_3 \equiv 0 \pmod{N(n)}.$$

Hence

$$x_1'+x_1+x_2+x_3-x_1 \equiv 0 \pmod{N(n)}.$$

$$x_1'-x_1 \equiv \lambda \pmod{N(n)}$$

$$x_1' \equiv x_1+\lambda \pmod{N(n)}$$

By the same reasoning

$$x_2' \equiv x_2+\lambda \pmod{N(n)}.$$

Hence we can rewrite The Coloring Equation as

$$\mathrm{COL}\ (x_1,x_2) = \ \mathrm{COL}\ (x_1+\lambda,x_2) = \ \mathrm{COL}\ (x_1,x_2+\lambda).$$

Since COL is a proper coloring, $\lambda \equiv 0 \pmod{N(n)}$.
End of proof of Claim 2.

2) Let P be a protocol for $\mathrm{MOD}_n^{N(n)}$. Let d be the maximum number of bits communicated. Note that the number of transcripts is bounded by 2^d. We use this protocol to create a proper 2^d-coloring of $[N(n)] \times [N(n)]$.

We define COL (x, y) as follows. First find z such that $x+y+z \equiv 0 \pmod{N(n)}$. Then run the protocol on (x, y, z). The color is the transcript produced.

Claim 3: COL is a proper coloring of $[N(n)] \times [N(n)]$.
Proof: Let $\lambda \in [N(n)]$ be such that

$$\text{COL }(x, y) = \text{ COL }(x+\lambda, y) = \text{ COL }(x, y+\lambda).$$

We denote this value $TRAN$ (for Transcript). We show that $\lambda \equiv 0 \pmod{N(n)}$.

Let z be such that

$$x+y+z \equiv 0 \pmod{N(n)}.$$

Since

$$\text{COL }(x, y) = \text{ COL }(x+\lambda, y) = \text{ COL }(x, y+\lambda).$$

We know that the following tuples produce the same transcript $TRAN$ (all arithmetic is mod $N(n)$):

- (x, y, z).
- $(x+\lambda, y, z-\lambda)$.
- $(x, y+\lambda, z-\lambda)$.

All of these input produce the same transcript $TRAN$ and this transcript ends with a YES. By Lemma 1.3.2 the tuple $(x, y, z-\lambda)$ also goes to $TRAN$. Hence $x+y+z-\lambda \equiv 0 \pmod{N(n)}$. Since $x+y+z \equiv 0 \pmod{N(n)}$ we have $\lambda \equiv 0 \pmod{N(n)}$.
End of Proof of Claim 3 ▮

3 Lower Bounds

3.1 An $\omega(1)$ Lower Bound for $d(\text{MOD}_n^{N(n)})$

We will need the following theorem from Ramsey Theory.

Theorem 3.1 *For all c there exists T such that, there are no proper c-colorings of* $[T] \times [T]$.

Theorem 3.1 can be proven several ways. We enumerate them:

1. This can be proven from van der Waerden's theorem.
2. This can be proven by the same techniques as van der Waerden's theorem.
3. This follows from the Galai-Witt Theorem. This generalizes to coloring $[T]^k$.

4. We will give a concrete lower bound (rather than $\omega(1)$) and is in Section 3.2. Other ways generalize to k variables.

Theorem 3.2 *If* $\lim_{n\to\infty} N(n) = \infty$ *then* $d(\mathrm{MOD}_n^{N(n)}) = \omega(1)$.

Proof: By Theorem 2.2

$$d(\mathrm{MOD}_n^{N(n)}) \ge \lg(\chi(N(n))) + \Omega(1).$$

Hence we need to show that $\chi(T)$ is not bounded by a constant (as T goes to infinity).

Assume, by way of contradiction, that there exists c such that, for all T, there is a proper c-coloring of $[T] \times [T]$. This contradicts Theorem 3.1. ▮

We will need to look at k-party protocols for the following function.
$\mathrm{MOD}_{n,k}^{N(n)} : (\{0,1\}^n)^k \to \{0,1\}$

$$\mathrm{MOD}_{n,k}^{N(n)}(x_1,\ldots,x_k) = \begin{cases} 1 & \text{if } \sum_{i=1}^k x_i \equiv 0 \pmod{N(n)} \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The following can be proven in a manner similar to the $k=3$ case.

Theorem 3.3 *Fix k. If* $\lim_{n\to\infty} N(n) = \infty$ *then* $d(\mathrm{MOD}_{n,k}^{N(n)}) = \omega(1)$.

3.2 An $\Omega(\log\log\log N(n))$ Lower Bound for $d(\mathrm{MOD}_n^{N(n)})$

The following combinatorial lemma will allow us to prove a lower bound on $d(\mathrm{MOD}_n^{N(n)})$. This lemma is a reworking of a theorem of Graham and Solymosi [9].

Lemma 3.4

1. *$\chi(N(n)) \ge \Omega(\log\log N(n))$.*
2. *$d(\mathrm{MOD}_n^{N(n)}) \ge \log\log\log N(n) + \Omega(1)$. (This follows from part 1 and Theorem 2.2.)*

Proof: Assume that COL is a proper c-coloring of $[N(n)] \times [N(n)]$. We find sets $X_1, Y_1 \subseteq [N(n)] \times [N(n)]$ such that COL restricted to $X_1 \times Y_1$ uses $c-1$ colors. We will iterate this process to obtain X_c, Y_c such that COL restricted to $X_c \times Y_c$ uses 0 colors. Hence $|X_c| = 0$ which will yield $c = \Omega(\log\log\log N(n)) = \Omega(\log\log n)$.

For $0 \le s \le c$ we define $X_s, Y_s, h_s,$ USED-COL$_s$.

1. $X_0 = Y_0 = [N(n)]$. $h_0 = |X_0| = |Y_0| = N(n)$. USED-COL$_0 = [c]$.

2. Assume X_s, Y_s, h_s are defined and inductively USED-COL$_s = [c-s]$ (we will be renumbering to achieve this). Also assume that Partition $X_s \times Y_s$ (which is of size h_s^2) into sets P_a indexed by $a \in [N(n)]$ defined by

$$P_a = \{(x,y) \in X_s \times Y_s \mid x + y \equiv a \pmod{N(n)}\}.$$

   (P_a is the ath anti-diagonal.) There exists an a such that $|P_a| \ge \lceil h_s^2/N(n) \rceil$. There exists a color, which we will take to be $c-s$ by renumbering, such that at least $\lceil \lceil h_s^2/N(n) \rceil /c \rceil$ of the elements of P_a are colored $c-s$. (We could use $c-s$ in the denominator but we do not need to.) Let $m = \lceil \lceil h_s^2/N(n) \rceil /c \rceil$. Let $\{(x_1,y_1),\ldots,(x_m,y_m)\}$ be m elements of P_a such that, for $1 \le i \le m$, COL $(x_i,y_i) = c-s$. We will later show that, for all $i \ne j$, COL $(x_i,y_j) \ne c-s$.

3. Let

$$\begin{aligned} h_{s+1} &= m' = \lceil m/3 \rceil \\ X_{s+1} &= \{x_1, x_2, \ldots, x_{m'}\} \\ Y_{s+1} &= \{y_{m+1-m'}, \ldots, y_m\} \\ \text{USED-COL}_{s+1} &= [c-(s+1)] \end{aligned}$$

   Note that for all $(x_i,y_j) \in X_{s+1} \times y_j \in Y_{s+1}$, $i < j$ hence $i \ne j$. Since we will show that for all $i \ne j$, COL $(x_i,y_j) \ne c-s$, we will have that, for all $(x,y) \in X_{s+1} \times y_j \in Y_{s+1}$, COL $(x,y) \ne c-s$.

Claim 1: For all $i \ne j$, $x_i \ne x_j$ and $y_i \ne y_j$.

Proof: If $x_i = x_j$ then

$$x_j + y_j = a = x_i + y_i = x_j + y_i.$$

Hence $y_j = y_i$. Therefore $(x_i,y_i) = (x_j,y_j)$. This contradicts P_a having m distinct points. The proof that $y_i \ne y_j$ is similar.

End of Proof of Claim 1

Claim 2: For all $i \ne j$, COL $(x_i,y_j) \ne c-s$.

Proof: Assume, by way of contradiction, that COL $(x_i,y_j) = c-s$. Note that

$$\text{COL } (x_i,y_j) = \text{ COL } (x_i,y_i) = \text{ COL } (x_j,y_j) = c-s.$$

We want a $\lambda \not\equiv 0 \pmod{N(n)}$ such that $y_i = y_j + \lambda$ and $x_j = x_i + \lambda$. Using that $x_i + y_i = x_j + y_j = a$ we can take $\lambda = (x_j + y_i - a)$. The element $\lambda \not\equiv 0 \pmod{N(n)}$: if $\lambda \equiv 0 \pmod{N(n)}$ then one can show $y_i = y_j$, which contradicts Claim 1.

We now have

$$\text{COL } (x_i,y_j) = \text{ COL } (x_i+\lambda,y_j) = \text{ COL } (x_i,y_j+\lambda).$$

This violates COL being a proper coloring.

End of Proof of Claim 2

Note that, by Claim 2 above

$$\{ \text{ COL } (x,y) \mid x \in X_{s+1}, y \in Y_{s+1}\} \subseteq \text{ USED-COL}_{s+1}.$$

Look at what happens at stage c. $|X_c| = |Y_c| = h_c$ and COL restricted to $X_c \times Y_c$ uses 0 colors. The only way this is possible is if $h_c = 0$. We will see that this implies $c = \Omega(\log\log N(n))$.

We have $h_0 = N(n)$ and

$$h_{s+1} = \left\lceil \left\lceil \left\lceil \frac{h_s^2}{N(n)} \right\rceil /c \right\rceil /3 \right\rceil \geq \frac{h_s^2}{3cN(n)}.$$

We show that for $s \in \mathbb{N}$, $h_s \geq \frac{N(n)}{(3c)^{2^s-1}}$.

Claim 3: $(\forall s)[h_s \geq \frac{N(n)}{(3c)^{2^s-1}}]$.

Base Case: $h_0 = N(n) \geq \frac{N(n)}{(3c)^0} = N(n)$.

Induction Step: Assume $h_s \geq \frac{N(n)}{(3c)^{2^s-1}}$. Since $h_{s+1} \geq (h_s)^2/3cN(n)$ we have, by the induction hypothesis

$$h_{s+1} \geq (h_s)^2/3cN(n) \geq \frac{\frac{N(n)^2}{(3c)^{2^{s+1}-2}}}{3cN(n)} \geq \frac{N(n)}{(3c)^{2^{s+1}-1}}.$$

End of proof of Claim 3

Taking $s = c$ we obtain $h_c \geq \frac{N(n)}{(3c)^{2^c-1}}$. Hence there is a set of h_c^2 points that are 0-colored. Therefore $h_c < 1$. This yields $c = \Omega(\log\log N(n))$. ∎

4 Applications to Lower Bounds on Branching Programs

Branching programs are a model of computation that are like decision trees except that nodes can be gotten to by several paths; hence they are 'skinny decision trees'. If a function $h : \{0,1\}^m \rightarrow \{0,1\}$ is computed by a branching program the key questions to ask are (1) what is its length? and (2) what is its width? We think of m as being large.

Note 4.1 It is known, and surprising, that all sets in NC^1 can be decided with poly-length, width 5, branching programs [2]. See [2, 14] or a paper on Branching Programs for a formal definition.

We will get lower bounds on the following function:

Def 4.2 Let $M(m) : \mathbb{N} \rightarrow \mathbb{N}$. Let $\mathrm{mod}_m^{M(m)} : \{0,1\}^m \rightarrow \{0,1\}$ be the function

$$\mathrm{mod}_m^{M(m)}(x_1,\ldots,x_m) = \begin{cases} 1 & \text{if } \sum_{i=1}^{m} x_i \equiv 0 \pmod{M(m)}; \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

We will first need to connect branching programs to multiparty communication complexity.

4.1 Branching Programs and Multiparty Communication Complexity

Def 4.3 A *bipartite graph* is a graph where there are two sets X and Y of vertices such that the only edges are between X and Y. We denote such a graph (X,Y,E) where E is the set of edges. The *complement of bipartite graph* G is bipartite graph $\overline{G} = (X, Y, X \times Y - E)$.

Def 4.4 Let $G = (X,Y,E)$ be a bipartite graph. Assume $|X| \le |Y|$. A *matching of X into Y* is a set of $|X|$ edges that share no vertices. Note that this is also an injection from X to Y.

Notation 4.5 Let $G = (X,Y,E)$ be a bipartite graph. If $A \subseteq X$ then $E(A) \subseteq Y$ is defined by

$$\{y \mid (\exists x \in A)[(x,y) \in E].$$

The following is Hall's theorem [10] and a corollary to it that we will be using.

Lemma 4.6 *Let $G = (X,Y,E)$ be a bipartite graph with $|X| \le |Y|$.*

1. *If for all $A \subseteq X$ $|E(A)| \ge |A|$ then G has a matching of X into Y.*

2. *Let $0 < \alpha < 1$. Let $k \in \mathbb{N}$. There exists m_0 such that, for all $m \ge m_0$, for all bipartite graphs $G = (X,Y,E)$ with $|X| = k$, $|Y| = m$, and $(\forall x \in X)[deg(x) \ge \alpha|Y|]$, there is a matching of X into Y.*

Lemma 4.7 *Let $k \in \mathbb{N}$ and $0 < \alpha < 1$. Let $g : \mathbb{N} \rightarrow \mathbb{N}$ be a monotone increasing function such that $g(m) << m$. There exists m_0 such that, for all $m \ge m_0$ the following holds. If $G = (X,Y,E)$ is a bipartite graph such that the following holds.*

1. *$X = \{v_1,\ldots,v_k\}$,*

2. *$Y = \{b_1,\ldots,b_m\}$, and*

3. *for all i, $1 \le i \le k$, $deg(v_i) \le \alpha m$,*

then there exists sets $R_1, \ldots, R_k \subseteq \{b_1, \ldots, b_m\}$ such that the following hold.

1. *For all $1 \le i \le k$, $|R_i| = g(m)$.*
2. *For all $1 \le i < j \le k$, $R_i \cap R_j = \emptyset$.*
3. *For all i, $E(\{v_i\}) \cap R_i = \emptyset$.*

Proof: We construct $R_1, \ldots, R_k$ in $g(m)$ stages. During the construction we will comment on how large m has to be to make it work. Let $H = \overline{G}$. Since $(\forall x \in X)[deg_G(x) \le \alpha m]$ we have that, $(\forall x \in X)[deg_H(x) \ge (1-\alpha)m]$. Let $\beta = 1 - \alpha$.
CONSTRUCTION

1. $R_1^0 = R_2^0 = \cdots = R_k^0 = \emptyset$. $\beta_0 = \beta$, $X_0 = X$, $Y_0 = Y$, $E_0 = \overline{E}$, $H_0 = (X_0, Y_0, E_0)$. Note that $deg_{H_0}(x) \ge \beta_0 |Y_0|$.
2. Assume that $R_1^s, R_2^s, \cdots, R_k^s$, m_s, β_s, X_s, Y_s, E_s, H_s are defined. Apply Hall's theorem to obtain a matching of X_s into Y_s. Let the matching be

$$\{(v_1, b_{i_1}), (v_2, b_{i_2}), \ldots, (v_k, b_{i_k})\}.$$

Let

$$\begin{array}{rl} (\forall 1 \le j \le m)[R_j^{s+1} & = R_j^s \cup \{b_{i_j}\}] \\ X_{s+1} & = X_s \\ Y_{s+1} & = Y_s - \{b_{i_1}, \ldots, b_{i_k}\} \\ \beta_{s+1} & = \beta_s - \frac{\beta_0}{2^{s+1}} \\ H_{s+1} & \text{is the induced bipartite graph of } H_s \text{ on } (X_{s+1}, Y_{s+1}) \end{array}$$

END OF CONSTRUCTION
Claim 1: If m is large enough then, for all $0 \le s \le g(m)$, the following hold.

1. $R_1^s, R_2^s, \cdots, R_k^s$, m_s, β_s, X_s, Y_s, E_s, H_s are defined.
2. For all $1 \le i < j \le k$, $R_i^s \cap R_j^s = \emptyset$
3. $|R_i^s| = s$,
4. $X_s = X$,
5. $|Y_s| = m - sk$,
6. $\beta_s = (1 - \frac{1}{2} - \frac{1}{2^2} - \cdots - \frac{1}{2^s})\beta_0$, and
7. $\beta_s \ge \beta_0/2$.

8. $(\forall x \in X_s)[deg_{H_s}(x) \geq \beta_s |Y_s|]$.

Proof of Claim 1:

All but the last item are proven by an easy induction. For the last item we need to show that,

$$(\forall x \in X_{s+1})[deg_{H_{s+1}}(x) \geq \beta_{s+1}|Y_{s+1}|].$$

Clearly

$$\begin{array}{rl} deg_{H_{s+1}}(x) & \geq deg_{H_s}(x) - k \geq \beta_s|Y_s| - k \geq \beta_s(m - sk) - k \geq \beta_s(m - g(m)k) - k \\ |Y_{s+1}| & = m - (s+1)k \leq m - (g(m)+1)k \end{array}$$

Hence we need

$$\beta_s(m - g(m)k) - k \geq \beta_{s+1}(m - (g(m)+1)k)$$

$$\frac{\beta_s(m - g(m)k) - k}{m - (g(m)+1)k)} \geq \beta_{s+1}$$

Since $g(m) << m$ we can take m large enough so that

$$\frac{\beta_s(m - g(m)k) - k}{m - (g(m)+1)k} \geq \beta_s - \frac{\beta_0}{2^{s+1}} = \beta_{s+1}.$$

End of Proof of Claim 1

For $1 \leq i \leq k$ let $R_i = R_i^{g(m)}$. These sets clearly satisfy the lemma.

▮

Note 4.8 In the above lemma the size of the R_i does not depend on k. This is important for applications.

Lemma 4.9 *Let $m \in \mathbb{N}$ and $f : \{0,1\}^m \rightarrow \{0,1\}$. Assume there is a BP for f of length $L(m) = cm$ and width $W(m) = d$, where d is a power of 2 (this avoids ceiling-floor problems). Let $k = 2cd$. Let $g : \mathbb{N} \rightarrow \mathbb{N}$ be such that $g(m) << m$. Then there exists a multiparty protocol for $f(b_1, \ldots, b_{kg(m)}, 0, \ldots, 0)$ with the following properties.*

1. *Player P_i has all of the bits except $b_{1+(i-1)g(m)}, \ldots, b_{ig(m)}$.*

2. *The protocol takes $2cd \log d = O(1)$ bits.*

Proof: Divide the BP into $k = 2cd$ segments of $\frac{L(m)}{2cd} = \frac{cm}{2cd} = \frac{m}{2d}$ levels each. Since each level has at most d variables, each segment has at most $\frac{m}{2}$ variables. Let the set of variables in the ith segment be denoted S_i. Let $X = \{S_1, \ldots, S_k\}$ and $Y = \{b_1, \ldots, b_m\}$. Form a bipartite graph (X, Y, E) with

$$E = \{(S_i, b_j) \mid b_j \in S_i\}.$$

Note that $deg(S_i) \leq \frac{m}{2}$. Hence this graph satisfies the premise of Lemma 4.7. Apply Lemma 4.7 to this graph to obtain sets $R_1, \ldots, R_k$ such that the following happens:

1. For all $1 \leq i \leq k$, $|R_i| = g(m)$.

2. For all $1 \leq i < j \leq k$, $R_i \cap R_j = \emptyset$.

3. For all i, $X_i \cap R_i = \emptyset$.

Note that X_i, the variables in the ith segment, contains *none* of the variables in R_i. Hence, if I gave you the values of the variables in $R_1 \cup \ldots R_{i-1}, R_{i+1}, \ldots, \mathrm{R}_k$ and what node was reached at the beginning of the ith segment, you could compute what happens in the ith segment.

By renumbering let $R_1 = \{b_1, \ldots, b_{g(m)}\}$, $R_2 = \{b_{1+g(m)}, \ldots, b_{2g(m)}\}$, etc.

We now say which players get which variables. For $1 \leq i \leq k$ player P_i gets all of the variables *except* those in R_i. **Note that he will have all of the variables in segment S_i.**

We now set up the multiparty protocol of k players to compute

$$f(b_1, \ldots, b_{kg(m)}, 0, \ldots, 0).$$

First, take the BP for h and set all of the variables that are not in $R_1 \cup \cdots \cup R_k$ to 0. Because of the renumbering this is setting $b_j = 0$ for $kg(m)+1 \leq j \leq m$. The new branching program computes $f(b_1, \ldots, b_{kg(m)}, 0, \ldots, 0)$. All of the players know this branching program.

We now describe the multiparty protocol. P_1 executes the Branching Program in segment S_1. This does not take any communication. (Note that he has all of the variable to do this. He may have more but this is not important.) The final node that is reached is one of at most $W(m) = d$ nodes. He broadcasts that node. This take $\log d$ bits. P_2 starts at that node and executes S_2. At the end he broadcasts $\log d$ bits to tell P_3 where he ended. This continues until all k players have broadcast $\log d$ bits and have completed the BP. The total number of bits transmitted is $2cd \log d$. ▮

4.2 Lower Bound on Branching Program

Theorem 4.10 *Let $M(m) : \mathbb{N} \rightarrow \mathbb{N}$ be such that $\lim_{m\rightarrow\infty} M(m) = \infty$. Let $\mathrm{mod}_m^{M(m)}$ be defined by*

$$\mathrm{mod}_m^{M(m)}(b_1, \ldots, b_m) \begin{cases} 1 & \text{if } \sum_{i=1}^m b_i \equiv 0 \pmod{M(m)} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

There is no linear-length constant-width BP for $\mathrm{mod}_m^{M(m)}$.

Proof: Assume, by way of contradiction, that there is a linear-length constant-width BP for h. Let $g : \mathbb{N} \rightarrow \mathbb{N}$ be a monotone increasing function such that $g(m) << m$ to be named later. Then, by Lemma 4.9 there is a constant k such that there exists a multiparty protocol for $\mathrm{mod}_m^{M(m)}(b_1, \ldots, b_{kg(m)}, 0, \ldots, 0)$ with the following properties.

1. Player P_i has all of the bits except $b_{1+(i-1)g(m)}, \ldots, b_{ig(m)}$.

2. The protocol takes $O(1)$ bits.

We will define a function $N(n)$ such that $\lim_{n\rightarrow\infty} N(n) = \infty$. We use the multiparty protocol for $\mathrm{mod}_m^{M(m)}$ to obtain a multiparty protocol for $\mathrm{MOD}_n^{N(n)}$ that takes $O(1)$ bits. This contradicts Theorem 3.3.

Let $N(n) : I \rightarrow \mathbb{N}$ be defined as follows: given n, find the least m such that $n = \lceil \lg g(m) \rceil$, then output $M(m)$. Since $\lim_{m\rightarrow\infty} M(m) = \infty$, $\lim_{n\rightarrow\infty} N(n) = \infty$.

In order for $N(n)$ to be well defined we will need that, for all n, there exists m such that $n = \lg g(m)$. We also need that $g(m) << m$. We take $g(m) = 2^{\lg \lg m}$

1. Input is $x_1, \ldots, x_k \in \{0, 1\}^n$. Let m be such that $n = \lg g(m)$. We interpret each x_i as a number in binary. That number is between 0 and $g(m) - 1$.

2. For all $1 \leq i \leq k$, P_i has all x's except x_i.

3. Intuition: For all $1 \leq i \leq k$ P_i will come up with (based on $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_k$ $k - 1$ blocks of $g(m)$ bits. P_i will think of his blocks of bits as being all but the ith block.

4. For all $1 \leq i \leq k$ P_i sets, for all $1 \leq j \leq k$, $j \neq i$, the first x_j of the bits $b_{1+(j-1)g(m)}, \ldots, b_{jg(m)}$ to 1, and the rest to 0. Note that this is possible since $x_i \leq g(m)$.

5. (Note that there are now bits $b_1, \ldots, b_{kg(m)}$ such that Player P_i has all but the ith block of $g(m)$ bits, and that the sum of the ith block is x_i. Hence $\sum_{i=1}^{k} x_i = \sum_{i=1}^{kg(m)} b_i$.) Players $P_1, \ldots, P_k$ execute the $O(1)$ bit protocol for $\mathrm{mod}_m^{M(m)}$.

Since $\sum_{i=1}^{k} x_i = \sum_{i=1}^{kg(m)} b_i$,

$$\mathrm{MOD}_{\lg g(m)}^{M(m)}(x_1, \ldots, x_k) = \mathrm{mod}_m^{M(m)}(b_1, \ldots, b_k g(m), 0, \ldots, 0).$$

▮

5 Upper Bounds: Connection to 3-free Sets

We bound $\chi(N(n))$ and hence, by Theorem 2.2, bound $d(\mathrm{MOD}_n^{N(n)})$.

We first find bounds on $\chi^*(N(n))$ which is the following.

Def 5.1 Let $c, T \in \mathbb{N}$. We think of $[T]$ as being $\{1, \ldots, T\}$ (*not* mod T).

1. A *proper' c-coloring of* $[T] \times [T]$ is a function $\mathrm{COL} : [T] \times [T] \rightarrow [c]$ such that there do not exist $x, y, z \in [T]$ and $\lambda \in [T-1]$ such that

$$\mathrm{COL}\ (x,y,z) = \mathrm{COL}\ (x+\lambda,y,z) = \mathrm{COL}\ (x,y+\lambda,z) = \mathrm{COL}\ (x,y,z+\lambda)$$

   (all of the additions are NOT mod T). Another way to look at this: In a proper' coloring there cannot be three vertices that (a) are the same color, and (b) are the corners of a right isosceles triangle with legs parallel to the axes and hypotenuse parallel to the line $y = -x$.)

2. Let $\chi^*(T)$ be the least c such that there is a proper' c-coloring of $[T] \times [T]$.

We will need the following definition from Ramsey Theory.

Def 5.2

1. A 3-AP is an arithmetic progression of length 3.
2. Let ψ_3^T be the minimum number of colors needed to color $\{1, \ldots, T\}$ such that there are no monochromatic 3-AP's.
3. A set $A \subseteq [T]$ is 3-*free* if there do not exist any 3-AP's in A.
4. Let $\mathrm{r}_3(T)$ be the size of the largest 3-free subset of $[T]$.

Lemma 5.3

1. *$\chi^*(T) \le \psi_3^{3T}$.*
2. *There exists a constant c such that $\psi_3^T \le c\frac{T \log T}{\mathrm{r}_3(T)}$.*
3. *There exists a constant c such that $\chi^*(T) \le c\frac{T \log(T)}{\mathrm{r}_3(T)}$. (This follows from 1 and 2.)*

Proof:
1) Let $c = \psi_3^{3T}$. Let COL' be a c-coloring of $[3T]$ with no monochromatic 3-AP's. Let COL be the following c-coloring of $[T] \times [T]$.

$$\text{COL } (x, y) = \text{ COL' } (x + 2y).$$

Assume, by way of contradiction, that COL is not a proper' c-coloring. Hence there exist $x, y, z \in [T]$ and $\lambda \neq 0$ such that

$$\text{COL } (x, y) = \text{ COL } (x + \lambda, y) = \text{ COL } (x, y + \lambda).$$

By the definition of COL the following are equal.

$$\text{COL' } (x + 2y) = \text{ COL' } (x + \lambda + 2y) = \text{ COL' } (x + 2\lambda + 2y)$$

Hence $x + 2y, x + 2y + \lambda, x + 2y + 2\lambda$ form a monochromatic 3-AP. which yields a contradiction.

2) Let $A \subseteq [T]$ be a set of size $r_k(T)$ with no 3-AP's. We use A to obtain a 3-free coloring of $[T]$. The main idea is that we use randomly chosen translations of A to cover all of $[T]$.

Let $x \in [T]$. Pick a translation of A by picking $t \in [T]$. The probability that $x \in A + t$ is $\frac{|A|}{T}$. Hence the probability that $x \notin A + t$ is $1 - \frac{|A|}{T}$. If we pick s translations $t_1, \ldots, t_s$ at random (s to be determined later) then the expected number of x that are not covered by any $A + t_i$ is

$$T\left(1 - \frac{|A|}{T}\right)^s \leq Te^{-s\frac{|A|}{T}}.$$

We need to pick s such that this quantity is < 1 We take $s = 2\frac{T \ln T}{|A|}$ which yields

$$Te^{-s\frac{|A|}{T}} = Te^{-2\ln T} = 1/T < 1.$$

We color T by coloring each of the s translates a different color. If a number is in two translates then we color it by one of them arbitrarily. Clearly this coloring has no monochromatic 3-APs. Note that it uses $\frac{T \ln T}{|A|} = O(\frac{T \log T}{\mathrm{r}_k(T)})$ colors. ▮

6 Three Free Sets

In this section we review several constructions of 3-free sets. Our notation will be to take them to be subsets of $\{1, \ldots, n\}$. In particular, $\mathrm{r}_3(n)$ will be the largest 3-free subset of $\{1, \ldots, n\}$. Do not confuse this n with the n we have used before.

We present constructions in order of how large a 3-free set they give us. This is not the same order they were discovered.

The following are trivial to prove; however, since we use it throughout the paper we need a shorthand way to refer to it:

Fact 6.1 *Let $x \leq y \leq z$. Then x, y, z is a 3-AP iff $x + z = 2y$.*

6.1 $\mathrm{r}_3(n) = \Omega(n^{0.63}$: The Base 3 Method

The following theorem appeared in [7] but they do not take credit for it; hence we can call it folklore.

Theorem 6.2 $\mathrm{r}_3(n) \ge n^{\log_3 2} \approx n^{0.63}$.

Proof:

$A_n = \{m \mid 0 \le m \le n$ and all the digits in the base 3 representation of m are in the set $\{0,1\}$ $\}$.

The following is a (large) subset of A_n: every number in base 3 of length $\lfloor \log_3 n \rfloor$ that only yas 0's and 1's. Hence

$$|A_n| \ge \Omega(2^{\log_3 n}) = \Omega(n^{\log_3 2}) \ge n^{0.63}.$$

We show that A_n is 3-free. Let $x, y, z \in A_n$ form a 3-AP. Let x, y, z in base 3 be $x = x_{k-1} \cdots x_0$, $y = y_{k-1} \cdots y_0$, and $z = z_{k-1} \cdots a_0$, By the definition of A_n, for all i, $x_i, y_i, z_i \in \{0,1\}$. By Fact 6.1 $x+z = 2y$. Since $x_i, y_i, z_i \in \{0,1\}$ the addition is done *without carries*. Hence we have, for all i, $x_i + z_i = 2y_i$. Since $x_i, y_i, z_i \in \{0,1\}$ we have $x_i = y_i = z_i$, so $x = y = z$. ▮

6.2 $\mathrm{r}_3(n) \ge \Omega(n^{0.68-\epsilon})$: The Base 5 Method

According to [7], G. Szekeres conjectured that $\mathrm{r}_3(n) = \Theta(n^{\log_3 2})$. This was disproven by Salem and Spencer [13] (see below); however, in 1999 Ruzsa (Section 13 of [12]) noticed that a minor modification to the proof of the Theorem 6.2 yields the following theorem which also disproves the conjecture. His point was that this is an easy variant of Theorem 6.2 so it is surprising that it was not noticed earlier.

Theorem 6.3 *For every $\epsilon > 0$ there exists n_0 such that, for all $n \ge n_0$,* $\mathrm{r}_3(n) \ge n^{(\log_5 3)-\epsilon} \sim n^{0.68-\epsilon}$.

Proof: Let L be a parameter to be chosen later. Let $k = \lfloor \log_5 n \rfloor - 1$. Let A be the set of positive integers that, when expressed in base 5,

1. use at most k digits,
2. use only 0's, 1's, and 2's, and
3. use *exactly* L 1's.

One can show, using Fact 6.1, that $A \subseteq [n]$ and A is 3-free. If we take $L = \lfloor k/3 \rfloor$ one can show that $|A| \ge n^{(\log_5 3)-\epsilon}$. ▮

Consider the following variant of the Base 5 method. Use Base 5, but use digits $\{-1,0,1\}$ and require that every numbers has exactly L 0's. If $(b_{k-1},\ldots,b_0)$ is a number expressed in Base 5 with digits $\{-1,0,1\}$ and with exactly L digits 0, then $\sum_{i=0}^{k-1} b_i^2 = n-L$. This method, expressed this way, is a our version of the Sphere Method (see Section 6.5) with parameters $d=1$ and $s=n-L$.

6.3 $\Omega(n^{1-\frac{c}{\lg\lg n}})$: The KD Method

The first disproof of Szekeres's conjecture (that $\mathrm{r}_3(n)=\Theta(n^{\log_3 2})$) was due to Salem and Spencer [13].

Theorem 6.4 *There exists c such that* $\mathrm{r}_3(n)\ge\Omega(n^{1-\frac{c}{\lg\lg n}})$.

Proof:
Let $d,n\in\mathbb{N}$. Let $k=\lfloor\log_{2d-1}n\rfloor-1$. Assume that d divides k. $\mathrm{KD}_{d,n}$ is the set of all $x\le n$ such that

1. when expressed in base $2d-1$ only uses the digits $0,\ldots,d-1$, and
2. each digit appears the same number of times, namely k/d.

We leave it to the reader to show that, for all d,n $\mathrm{KD}_{d,n}$ is 3-free.
An easy calculation shows that, for any d,n, $\mathrm{KD}_{d,n}\subseteq[n]$. Clearly

$$|\mathrm{KD}_{d,n}|=\frac{k!}{[(k/d)!]^d}.$$

By picking d such that $(2d)^{d(\lg d)^2}\sim n$ one can show that $|A|\ge\Omega(n^{1-\frac{c}{\lg\lg n}})$ for some c. ▮

Note 6.5 The c in $\frac{c}{\lg\lg n}$ can be replaced by $1+\epsilon$.

6.4 $\Omega(n^{1-\frac{c}{\sqrt{\lg n}}})$: The Block Method

Behrend [4] and Moser [11] both proved $\mathrm{r}_3(n)\ge n^{1-\frac{c}{\sqrt{\lg n}}}$, for some value of c. Behrend proved it first and with a smaller (hence better) value of c, but his proof was nonconstructive (i.e, the proof does not indicate how to actually find such a set). Moser's proof was constructive. We present Moser's proof here; Behrend's proof is presented later.

Theorem 6.6 *[11]* $\mathrm{r}_3(n)\ge\Omega(n^{1-\frac{c}{\sqrt{\lg n}}})$.

Proof: Let r be such that $2^{r(r+1)/2}-1 \le n \le 2^{(r+1)(r+2)/2}-1$. Note that $r \ge \sqrt{2\lg n}-1$.

We write the numbers in $[n]$ in base 2. We think of a number as being written in r blocks of bits. The first (rightmost) block is one bit long. The second block is two bits long. The rth block is r bits long. Note that the largest possible number is $r(r+1)/2$ 1's in a row, which is $2^{r(r+1)/2}-1 \le n$. We call these blocks $x_1, \ldots, x_r$. Let B_i be the number represented by the ith block. The concatenation of two blocks will represent a number in the natural way.

Example: We think of $(1001110101)_2$ as $(1001 : 110 : 10 : 1)$ so $x_1 = (1)_2 = 1$, $x_2 = (10)_2 = 2$, $x_3 = (110)_2 = 6$, and $x_4 = (1001)_2 = 9$. We also think of $x_4x_3 = (1001110)_2 = 78$.
End of Example

The set A is the set of all numbers $x_r x_{r-1} \ldots x_1$ such that

1. For $1 \le i \le r-2$ the leftmost bit of x_i is 0. Note that when we add together two numbers in A the first $r-2$ blocks will add with no carries.

2. $\sum_{i=1}^{r-2} x_i^2 = x_r x_{r-1}$

Example: Consider the number $(10110011011000101011010)_2$. We break this into blocks to get $(0000010 : 110011 : 01100 : 0101 : 011 : 01 : 0)_2$. Note that there are $r = 7$ blocks and the rightmost $r-2 = 5$ of them all have a 0 as the leftmost bit. The first 5 blocks, reading from the right, as base 2 numbers, are $0 = 0$, $01 = 1$, $011 = 3$, $0101 = 5$, $01100 = 12$. The leftmost two blocks merged together are $0000010110011 = 179$. Note that $0^2+1^2+3^2+5^2+12^2 = 179$. Hence the number $(10110011011000101011010)_2$ is in A.
End of Example

We omit the proof that A is 3-free, but note that it uses Fact 6.1.

How big is A? Once you fill in the first $r-2$ blocks, the content of the remaining two blocks is determined and will (by an easy calculation) fit in the allocated $r+(r-1)$ bits. Hence we need only determine how many ways the first $r-2$ blocks can be filled in. Let $1 \le i \le r-2$. The ith block has i places in it, but the leftmost bit is 0, so we have $i-1$ places to fill, which we can do 2^{i-1} ways. Hence there are $\prod_{i=1}^{r-2} 2^{i-1} = \prod_{i=0}^{r-3} 2^i = 2^{(r-2)(r-3)/2}$.

$(r-2)(r-3) \ge (\sqrt{2\lg n}-3)(\sqrt{2\lg n}-4) = 2\lg n - 7\sqrt{2\lg n}+12$

So

$(r-2)(r-3)/2 \ge \lg n - 3.5\sqrt{2\lg n}+6$

So

$2^{(r-2)(r-3)/2} \ge 2^{\lg n - 3.5\sqrt{2\lg n}+6} \sim n^{1-\frac{3.5\sqrt{2}}{\sqrt{\lg n}}}$ ∎

6.5 $\mathrm{r}_3(n) \ge \Omega(n^{1-\frac{c}{\sqrt{\lg n}}})$: The Sphere Methods

In Sections 6.1, 6.2, 6.3, and 6.4 we presented constructive methods for finding large 3-free sets of $[n]$ for large n. In this section we present the Sphere Method which is nonconstructive.

The result and proof in this section are a minor variant of what was done by Behrend [4, 8]. We will express the number in a base and put a condition on the representation so that the numbers do not form a 3-AP. It will be helpful to think of the numbers as vectors.

Def 6.7 Let $x, b \in \mathbb{N}$ and $k = \lfloor \log_b x \rfloor$. Let x be expressed in base b as $\sum_{i=0}^{k} x_i b^i$. Let $\vec{x} = (x_0, \ldots, x_k)$ and $|\vec{x}| = \sqrt{\sum_{i=0}^{k} x_i^2}$.

Behrend used digits $\{0, 1, 2 \ldots, d\}$ in base $2d+1$. We use digits $\{-d, -d+1, \ldots, d\}$ in base $4d+1$. This choice gives slightly better results since there are more coefficients to use. Every number can be represented uniquely in base $4d+1$ with these coefficients. There are no carries since if $a, b \in \{-d, \ldots, d\}$ then $-(4d+1) < a+b < (4d+1)$.

We leave the proof of the following lemma to the reader.

Lemma 6.8 *Let* $x = \sum_{i=0}^{k} x_i(4d+1)^i$, $y = \sum_{i=0}^{k} y_i(4d+1)^i$, $z = \sum_{i=0}^{k} z_i(4d+1)^i$, *where* $-d \le x_i, y_i, z_i \le d$. *Then the following hold.*

1. $x = y$ *iff* $(\forall i)[x_i = y_i]$.

2. *If* $x + y = 2z$ *then* $(\forall i)[x_i + z_i = 2y_i]$

The set $A_{d,s,k}$ defined below is the set of all numbers that, when interpreted as vectors, have norm s (norm is the square of the length). These vectors are all on a sphere of radius $\sqrt{s}$. We will later impose a condition on k so that $A_{d,s,k} \subseteq [-n/2, n/2]$.

Def 6.9 Let $d, s, k \in \mathbb{N}$.

$$A_{d,s,k} = \left\{ x : x = \sum_{i=0}^{k-1} x_i(4d+1)^i \wedge (\forall i)[-d \le x_i \le d] \wedge (|\vec{x}|^2 = s) \right\}$$

Def 6.10 Let $d, s, m \in \mathbb{N}$.

$$B_{d,s,k} = \left\{ x : x = \sum_{i=0}^{k-1} x_i(4d+1)^i \wedge (\forall i)[0 < x_i \le d] \wedge (|\vec{x}|^2 = s) \right\}$$

Lemma 6.11 *Let* $n, d, s, k \in \mathbb{N}$.

1. $A_{d,s,k}$ *is 3-free.*

2. *If* $n = (4d+1)^k$ *then* $A_{d,s,k} \subseteq \{-n/2, \ldots, n/2\}$.

Proof: a) Assume, by way of contradiction, that $x, y, z \in A_{d,s,k}$ form a 3-AP. By Fact 6.1, $x + z = 2y$. By Lemma 6.8 $(\forall i)[x_i + z_i = 2y_i]$. Therefore $\vec{x} + \vec{z} = 2\vec{y}$, so $|\vec{x} + \vec{z}| = |2\vec{y}| = 2|\vec{y}| = 2\sqrt{s}$. Since $|\vec{x}| = |\vec{z}| = \sqrt{s}$ and $\vec{x}$ and $\vec{z}$ are not in the same direction $|\vec{x} + \vec{z}| < 2\sqrt{s}$. This is a contradiction.

b) The largest element of $A_{d,s,k}$ is at most

$$\sum_{i=0}^{k-1} d(4d+1)^i = d\sum_{i=0}^{k-1}(4d+1)^i = \frac{(4d+1)^k - 1}{2} = \frac{n-1}{2} \le n/2.$$

Similarly, the smallest element is $\ge -n/2$. ∎

Lemma 6.12 *For all d, s, k*

$$|A_{d,s,k}| = \sum_{m=0}^{k} \binom{k}{m} 2^m |B_{d,s,m}|.$$

Proof:

Define

$$A^m_{d,s,k} = \left\{ x : x = \sum_{i=0}^{k-1} x_i (4d+1)^i \wedge (\forall i)[-d \le x_i \le d] \right.$$

$$\left. \wedge (\text{ exactly } m \text{ of the } x_i\text{'s are nonzero }) \wedge (|\vec{x}|^2 = s) \right\}$$

Clearly $|A_{d,s,k}| = \sum_{m=0}^{k} |A^m_{d,s,k}|$.

Note that $|A^m_{d,s,k}|$ can be interpreted as first choosing m places to have non-zero elements (which can be done in $\binom{k}{m}$ ways), then choosing the absolute values of the elements (which can be done in $|B_{d,s,m}|$ ways) and then choosing the signs (which can be done in 2^m ways). Hence $|A^m_{d,s,k}| = \binom{k}{m} 2^m |B_{d,s,m}|$. So

$$|A_{d,s,k}| = \sum_{m=0}^{k} \binom{k}{m} 2^m |B_{d,s,m}|.$$

▮

Theorem 6.13 *There is a c such that* $\mathrm{r}_3(n) \ge \Omega(n^{1-\frac{c}{\sqrt{\lg n}}})$.

Proof:

Let d, s, k be parameters to be specified later. We use the set $A_{d,s,k}$ which, by Lemma 6.11, is 3-free. We seek values of d, k, s such that $|A_{d,s,k}|$ is large and contained in $[-n/2, n/2]$. Note that once k, d are set the only possibly values of s are $\{0, 1, \ldots, kd^2\}$.

A calculation shows that if $k \approx \sqrt{\lg n}$ and d is such that $n = (4d+1)^k$ then $\bigcup_{s=0}^{kd^2} |A_{d,s,k}|$ is so large that *there exists* a value of s such that $|A_{d,s,k}| \ge n^{1-\frac{c}{\sqrt{\lg n}}}$ for some value of c. Note that the proof is nonconstructive in that we do not specify s; we merely show it exists. ▮

7 The Upper Bound

We leave the following lemma to the reader.

Lemma 7.1 *For all $N(n)$ there is a constant c such that $\chi(N(n)) \le c\chi^*(N(n))$.*

Theorem 7.2 $d(\mathrm{MOD}_n^{N(n)}) =$

Proof: By Theorem 2.2

$$d(\mathrm{MOD}_n^{N(n)}) \le 2\lg(\chi(N(n))) + O(1).$$

By Lemma 7.1 there exists a constant c such that $\chi(N(n)) = c\chi^*(N(n))$. Hence

$$d(\mathrm{MOD}_n^{N(n)}) \le 2\lg(\chi^*(N(n))) + O(1).$$

By Lemma 5.3 there exists a constant c such that

$$\chi^*(N(n)) \le c\frac{N(n)\log(N(n))}{\mathrm{r}_3(N(n))} + O(1).$$

Hence

$$d(\mathrm{MOD}_n^{N(n)}) \le 2\lg(\frac{N(n)\log(N(n))}{\mathrm{r}_3(N(n))}) + O(1).$$

By Theorem 6.13 there exists a constant c such that

$$\mathrm{r}_3(N(n)) \ge \Omega(N(n)^{1-\frac{c}{\sqrt{\lg N(n)}}}).$$

Hence

$$d(\mathrm{MOD}_n^{N(n)}) \le 2\lg(\frac{N(n)\log(N(n))}{\mathrm{r}_3(N(n))}) + O(1).$$

■

8 HW

1. k-party case.
2. Prove that if $(x+\lambda, y, z)$ and $(x, y+\lambda, z)$ and $(x, y, z+\lambda)$ produce the same transcript, then (x, y, z) also produces that transcript.
3. Show that non-constant in k-dim case (need Gallai-Witt).
4. What about the case of $f(x, y, z) = T$. For what T is this.... .
5. What happens over a group?
6. Prove that for all c there exists N such that for all c-colorings of $[N] \times [N]$ there exists a monochromatic isos L. A square.
7. Use the result for three parties to get some lower bound on some branching program.
8. Prove Hall's Theorem.
9. Prove that corollary to Hall's theorem.

References

[1] Babai, Nisan, and Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *Journal of Computer and System Sciences*, 45, 1992. Prior version in STOC89.

[2] D. Barrington. Bounded-width polynomial-size branching programs recognize exactly those languages in NC^1. *Journal of Computer and System Sciences*, 38, 1989.

[3] P. Beame and E. Vee. Time-space tradeoffs, multiparty communication complexity and nearest neighbor problems. In *Proceedings of the Thirty-fourth Annual ACM Symposium on the Theory of Computing,* Montreal, Canada, 2002. `http://www.cs.washington.edu/homes/beame/publications.html`.

[4] F. Behrend. On set of integers which contain no three in arithmetic progression. *Proc. of the National Acadamy of Science (USA)*, 23:331–332, 1946.

[5] R. Beigel, W. Gasarch, and J. Glenn. The multiparty communication complexity of exact-t: improved bounds and new problems. In *Proceedings of the 31th International Symposium on Mathematical Foundations of Computer Science 2001,* Stara Lesna, Slovakia, 2006.

[6] A. Chandra, M. Furst, and R. Lipton. Multiparty protocols. In *Proceedings of the Fifteenth Annual ACM Symposium on the Theory of Computing,* Boston MA, pages 94–99, 1983. `http://portal.acm.org/citation.cfm?id=808737`.

[7] P. Erdös and P. Turan. On some sequences of integers. *Journal of the London Mathematical Society*, 11(2):261–264, 1936. `http://jlms.oxfordjournals.org/`.

[8] R. Graham, B. Rothchild, and J. Spencer. *Ramsey Theory.* Wiley, 1990.

[9] R. Graham and J. Solymosi. Monochromatic equilateral right triangles on the integer grid. *Topics in Discrete Mathematics, Algorithms and Combinatorics*, 2006. `www.math.ucsd.edu/~/ron/06\_03\_righttriangles.pdf` or `www.cs.umd.edu/~/vdw/graham-solymosi.pdf`.

[10] P. Hall. On representations of subsets. *Journal of the London Mathematical Society*, 10:26–30, 1935. `http://jlms.oxfordjournals.org/`.

[11] L. Moser. On non-averaging sets of integers. *Canadian Journal of Mathematics*, 5:245–252, 1953.

[12] I. Ruzsa. Erdös and the numbers. *Journal of Number Theory*, pages 115–163, 1999.

[13] R. Salem and D. Spencer. On set of integers which contain no three in arithmetic progression. *Proc. of the National Acadamy of Science (USA)*, 28:561–563, 1942. `http://www.cs.umd.edu/~gasarch/vdw/vdw.html`.

[14] I. Wegener. *Branching Programs and Binary Decision Diagrams: Theory and Application*. SIAM, 2000.