On the degree of ideal membership proofs from uniform families of polynomials over a finite field

Jan Krajíček*

Mathematical Institute, Academy of Sciences
Žitná 25, Prague, 115 67, The Czech Republic
and
Mathematical Institute, Oxford University
24-29 St.Giles', Oxford, OX1 3LB, U.K.

*Partially supported by cooperative research grant INT-9600919/ME-103 from the NSF (USA) and the MŠMT (Czech republic) and by the grant #A1019602 of the Academy of Sciences of the Czech Republic. Written while the author was funded by the EPSRC fellowship number GR/L01176.

Proposed running head:

Degree of ideal membership proofs

Mailing address:

Jan Krajíček
Mathematical Institute
Academy of Sciences
Žitná 25
Prague, 115 67
The Czech republic
`krajicek@math.cas.cz`

Abstract

Let $f_0, f_1, \ldots, f_k$ be n-variable polynomials over a finite prime field $\mathbf{F}_p$. A proof of the ideal membership $f_0 \in \langle f_1, \ldots, f_k \rangle$ in *polynomial calculus* is a sequence of polynomials $h_1, \ldots, h_t$ such that $h_t = f_0$, and such that every h_i is either an f_j, $j \geq 1$, or obtained from $h_1, \ldots, h_{i-1}$ by one of the two inference rules: g_1 and g_2 entail any $\mathbf{F}_p$-linear combination of g_1, g_2, and g entails $g \cdot g'$, for any polynomial g'. The degree of the proof is the maximum degree of h_i's.

We give a condition on families $\{f_{N,0}, \ldots, f_{N,k_N}\}_{N<\omega}$ of n_N-variable polynomials of bounded degree implying that the minimum degree of polynomial calculus proofs of $f_{N,0}$ from $f_{N,1}, \ldots, f_{N,k_N}$ cannot be bounded by an independent constant and, in fact, is $\Omega(\log(\log(N)))$. In particular, we obtain an $\Omega(\log(\log(N)))$ lower bound for the degrees of proofs of 1 (so called *refutations*) of the (N, m) - system (defined in [4]) formalizing a modular counting principle (m fixed and not divisible by p, parameter N not divisible by m), and a similar lower bound for refutations of systems encoding that N is composite (whenever N is a prime). No bounds were previously known for these systems. The same method yields $\Omega(\log(N))$ lower bounds for the degree of coefficient polynomials in Nullstellensatz proofs.

The method is based on a new result about a uniform way of generating all submoduli of tabloid moduli.

Keywords:

Nullstellensatz, polynomial ideals, Young tableaux, algebraic proof complexity, modular counting principles, polynomial calculus.

1991 Mathematics Subject Classification:

Primary 03F20, 12L12; Secondary 12E12, 68Q15, 13L05.

Introduction

We are concerned here with algebraic proof complexity and we shall use the combinatorics of the representation theory of symmetric groups (Young tableaux, in particular) to obtain new degree lower bounds for Nullstellensatz proofs and for polynomial calculus proofs. The application requires a new result describing a uniform combinatorial way of generating all submoduli of tabloid moduli. The connection between the fields is achieved by identifying certain polynomial rings with permutation moduli. I shall discuss first the background for proof complexity and then the combinatorics of Young tableaux used.

Algebraic proof complexity studies time complexity of non-deterministic algorithms accepting exactly systems of polynomial equations over a fixed field that are not solvable in that field (or some specified extension). In our case the field will be a finite prime field so the encoding of the system for algorithms is straightforward.

Proof complexity in a broad sense, as I understand it, studies time complexity of non-deterministic algorithms accepting a set L of finite objects of certain type (natural numbers, polynomials, graphs, formulas, etc.). A non-deterministic acceptor A of L is called a proof system for L, a computation accepting y is an A-proof of "$y \in L$". The main goal of proof complexity is to exhibit an explicit L for which there is no proof system running in polynomial time. The qualification *explicit* means that L should belong to some computational complexity class X like, for example, coNP, polynomial time hierarchy or polynomial space. As the existence of a proof system running in polynomial time just means that $L \in$ NP, proof complexity aims in effect at showing that X $\not\subseteq$ NP. For example, whether coNP $\subseteq$ NP is a major open problem in complexity theory and logic.

The name *proof complexity* for this area comes from its most important and most developed part when L is the set of propositional tautologies TAUT (this part of proof complexity is called propositional proof complexity), cf. [9]. Logical calculi for proving propositional tautologies that are sound and complete are, in particular, non-deterministic acceptors of TAUT. A standard conjecture that NP $\neq$ coNP thus implies that none of these calculi admit polynomial size proofs of all tautologies. However, it is a challenge and a fundamental open problem of logic (cf.[17]) to prove unconditional lower bound for ordinary Hilbert-style propositional calculi that one can find in logic text-books. Such bounds would also have concrete corollaries for bounded arithmetic in terms of consistency of various open complexity-theoretic conjectures, cf. [15]. These open problems about bounded arithmetic are the most important open questions about the logic of first-order theories of arithmetic.

Despite some remarkable achievements no non-trivial lower bounds for the usual calculus based on a finite number of axiom schemes and on modus ponens are known.

In previous research it turned out that some open proof complexity questions about usual propositional calculi working with boolean connectives can be reduced to questions about degree lower bounds in various ideal membership proof systems. In fact, boolean connectives can be represented by polynomials (with 0 and 1 representing the truth values) and thus polynomial rings are as good an environment for propositional logic as is boolean logic itself. We shall not repeat the reductions here and instead concentrate on "algebraic" proof systems from the beginning. The reader is encouraged to consult [4, 6].

There are two main algebraic proof systems studied at present time. One system is the *Nullstellensatz proof system* NS, where a proof of the ideal membership $f_0 \in \langle f_1, \ldots, f_k \rangle$ (we will say just "a proof of f_0 from $f_1, \ldots, f_k$"), with $f_i \in F[\overline{x}]$, is a sequence $g_1, \ldots, g_k \in F[\overline{x}]$ such that

$$g_1 f_1 + \ldots + g_k f_k = f_0$$

By Hilbert's Nullstellensatz, a proof of unsolvability of $f_1 = 0, \ldots, f_k = 0$ in the algebraic closure $\tilde{F}$ of F is a proof of 1 from $f_1, \ldots, f_k$. This is also called *a refutation* of $f_1, \ldots, f_k$.

Note that unsolvability of polynomial systems over $\mathbf{F}_p$ (in $\mathbf{F}_p$) is a coNP-complete language and hence the complexity of proof systems for it are as important as for tautologies.

Another proof system is *polynomial calculus*[1]. A polynomial calculus proof (PS-proof) of f_0 from $f_1, \ldots, f_k$ is a sequence of polynomials $h_1, \ldots, h_t$ such that $h_t = f_0$, and such that every h_i is either an f_j, $j \geq 1$, or obtained from $h_1, \ldots, h_{i-1}$ by one of the two inference rules:

1. Addition rule: g_1 and g_2 entail any $\mathbf{F}_p$-linear combination of g_1, g_2.

2. Multiplication rule: g entails $g \cdot g'$, for any polynomial g'.

The degree of the NS-proof is the maximum degree of $g_i f_i$'s, and the degree of the PC-proof is the maximum degree of h_i's. Our goal is to prove lower bounds on the degree of proofs.

We shall study the case when the underlying field is $\mathbf{F}_p$ and when we are interested in solvability in $\mathbf{F}_p$ rather than in $\tilde{\mathbf{F}}_p$. This is easily achieved by including equations $x^p - x = 0$, for all variables x, among the starting polynomials.

The degree of NS-proofs is obviously related to the so called effective Nullstellensatz of Brownawell [5] and Kollár [14]. However, their results are proved for algebraically closed fields and the methods yield nothing for the case of bounds over finite fields. First non-trivial lower bounds for $\mathbf{F}_p$ were proved in [4, 6]; see those papers for a more detailed discussion of this connection.

We shall consider dense representation of polynomials. A polynomial f with variables $x_1, \ldots, x_n$ and of degree d is represented by a list of coefficients (in $\mathbf{F}_p$)

[1] It was first formally considered as a proof system in [8].

of all monomials of degree up to d, even if these are zero. The list has the length proportional to n^d, hence superpolynomial lower bounds on size correspond to non-constant lower bounds on the degree.[2]

Other studied algebraic systems include various extensions of NS (see [4, 6]), and system $F(MOD_p)$ combining algebraic and propositional reasoning (see [15, Sec.12.6]).

NS-proofs are particular PC-proofs (first derive by the multiplication rule all $g_i f_i$'s and then sum them up using the addition rule) and strong lower bounds are known for them, cf.[6]. PC is a natural "depth 1" subsystem of $F(MOD_p)$. Recently a lower bound for PC was proved in [21]. No lower bounds are known even for constant-depth $F(MOD_p)$, except for a weak fragment (cf. [16]).

Although we put the emphasis on proof complexity, the technical heart of our paper is a new way of generating submoduli of certain permutation moduli. It can be described briefly as follows. Let $\lambda = (\lambda_1, \ldots, \lambda_k)$, $\lambda_1 + \ldots + \lambda_k = N$, be a partition of a natural number N into integers. A decomposition $X_1 \cup \ldots \cup X_k = \{0, \ldots, N-1\}$ of N into disjoint sets of size $|X_i| = \lambda_i$ is called a λ-tabloid. Tabloid module M^λ over a field F is the F-vector space whose basis is the set of λ-tabloids. The symmetric group $Sym(N)$ acts on λ-tabloids, hence on elements of M^λ, and M^λ is an $F[Sym(N)]$-module, where $F[Sym(N)]$ is the appropriate group-algebra.

We give a combinatorial/logical description of all submoduli of M^λ, depending only on $\sum_{i \geq 2} \lambda_i$ but not on N (Theorem 3.3). This is based on analysing uniformity (definability) of various combinatorial manipulations with Young tableaux and Specht moduli, as used in James's characteristic-free representation theory of the symmetric group [13]. The connection of algebraic proof complexity to combinatorics of tabloid moduli is achieved by identifying certain polynomial rings with certain moduli. The present paper is the first one to make this connection and to apply the representation theory to algebraic proof systems.

Our method and results expand upon work of Ajtai [1, 2]. In connection with a combinatorial problem arising in propositional proof complexity he needed to characterize solvability of symmetric systems of linear equations (we describe his work in Section 2). While working on [18] I noticed that a notion of uniformity used there could yield a lower bound for PC proofs of the modular counting principles represented by the (N, m) - systems of [4] if a certain construction with tabloid moduli is made uniform (no lower bound for these principles follows from [21]). This we achieve by exploiting the constructions in [2] and [13, 12] (results on Specht modules there, in particular).

Another corollary of our method is a strengthening of Ajtai's results in the form of effective bounds (Theorem 3.5). As its application to algebraic proof

[2] I remark that some of the lower bounds, most notably for the (N, m)-systems, actually hold also for the size of the sparse representation.

complexity we show that the minimum degree of PC-refutations of the (N, m) - systems cannot be bounded by an independent constant and, in fact, must be at least $\Omega(\log(\log(N)))$ (m not divisible by p and the field $\mathbf{F}_p$ fixed, N parameter) (Corollary 5.3). No lower bounds for degree of PC - refutations of these systems were known previously.[3]

In fact, we prove a general sufficient condition (Theorem 5.1) implying $\Omega(\log(\log(N)))$ lower bound for any family $\{f_{N,1}, \ldots, f_{N,k_N}\}_{N<\omega}$ of polynomial systems, provided the family is given in a *uniform* way, with uniform being defined in terms of first-order logic (or, equivalently by Lemma 1.4, combinatorially). We use this to show also a lower bound for proofs of primality of N as encoded by a uniform system of equations defined here. The same arguments give $\Omega(\log(N))$ lower bounds for NS - proofs (Theorem 5.6).

These explicit non-constant lower bounds are derived from new technical results in Section 3. The reader interested only in the non-existence of constant degree proofs can bypass this section and go to Theorem 5.5. There we prove that families with constant degree PC - refutations admit also constant degree NS - refutations. For this result we need only a bound to the number of submoduli of particular tabloid moduli (Corollary 3.2 that follows already from Ajtai's work). In particular, a non-constant lower bound for the (N, m)-systems follows from Theorem 5.5 combined with the non-constant degree lower bound for NS from [4]. The non-existence of constant degree NS - refutations for general uniform systems can be derived from Theorem 3.1. Note that linear lower bounds would follow from the non-constant lower bounds (that we proved already in the preliminary version of this paper) should the main theorem announced in [22] hold in characteristic p (presently it is claimed only in characteristic zero[4]).

We shall not actually use any results or methods from [18] but only a rudimentary notion of uniformity used there, and the whole presentation is self-contained (in Section 3 we use some definitions and constructions from [13, 12] that are fully described there).

The reader interested in more background information and in connections to some other topics may consult [15] or the introduction to [6].

For the whole paper we fix a finite prime field $\mathbf{F}_p$.

1 Uniform systems of polynomials

We shall define first the (N, m) - system as it nicely motivates the more general definition of uniform polynomial families. N is identified with $\{0, \ldots, N-1\}$ and $[N]^m$ denotes the set of m - element subsets of N.

[3]More than a year after the preliminary version of this paper was available [7] improved the bounds for the particular (N, m)-systems to $\Omega(N)$.

[4]As I was told by S. Riis in March'98.

Definition 1.1 ([4]) *Let $N \geq m \geq 2$. The variables of the (N, m) - system are x_e, where e ranges over $[N]^m$. The system consists of the following polynomials*

1. $Q_e \ := \ x_e^2 - x_e$ *, for each e.*

2. $Q_{e,f} \ := \ x_e \cdot x_f$ *, for every e, f such that $e \cap f \neq \emptyset$ but $e \neq f$.*

3. $Q_i \ := \ 1 - \sum_{e:\ i \in e} x_e$ *, for each $i \in N$.*

We identify a polynomial system $\{F_i\}_i$ with the system of equations $\{F_i = 0\}_i$ and, in particular, the (N, m) - system with the system of equations $Q_e = 0$, $Q_{e,f} = 0$, $Q_i = 0$.

Assume that $x_e \ := \ a_e$ is a solution of the (N, m) - system in some integral domain. Then by equations $Q_e = 0$ all a_e are 0 or 1, and by the remaining equations the set

$$\{e \in [N]^m \mid a_e = 1\}$$

is a partition of N into m - element sets. Thus the system has a solution iff m divides N.

For m not dividing N the ideal in $\mathbf{F}_p[\overline{x}]$ generated by the system is then necessarily trivial (a simple consequence of Nullstellensatz). We shall be interested in the minimum degree of PC-refutations (i.e., proofs of 1) of the system.

The crucial uniformity property (to be defined formally bit later) of the (N, m)-system is the following. Variables are indexed by $e \in [N]^m$ and equations are naturally indexed by e (Case 1.), by pairs (e, f) (Case 2.), and by singletons $\{i\}$ (Case 3.). The permutations of N act on the (indices of) variables, on (indices of) equations, as well as on monomials and polynomials, and the (N, m) - system is invariant under such actions. This means that the coefficient of a monomial, say $x_g x_h$, in the equation indexed by, say, (e, f) depends only on the isomorphism type of finite structure $(U; e, f, g, h)$ with universe $U := e \cup f \cup g \cup h$ and distinguished sets e, f, g, h, but not on N. This invariance property is the key property of families of polynomials systems allowing to apply the representation theory of the symmetric group.

We shall now define a general notion of an index, and of uniformity and definability of sets of indices. After that we treat polynomial systems, in particular.

We define uniformity of sets of indices in two equivalent ways. In a logical way as a definability in a certain first order language, and in a combinatorial way in terms of types of finite structures formed from indices. Although the combinatorial version is somewhat more rudimentary and would suffice for the present paper, the logical version allows easier explanation of Ajtai's work in Section 2, as well as generalizations of the method developed in [18].

Definition 1.2 *Let $L(C)$ be a first order language having an element-sort and a set-sort, and consisting of an equality predicate $=$ in both sorts, of a membership relation $\in$ between elements and sets, and of a finite set C of element constants.*

The set-sort can be used only as parameters; i.e., quantifiers may range only over the element-sort.

An $L(C)$-structure M interprets the element-sort by its elements and the set-sort by its subsets. We shall always assume that all constants from C are different in M.

Let $k \geq 1$.

1. *A k-ary index in M is a k-tuple $i = (i_1, \ldots, i_k)$ of finite subsets of M such that $k \geq |i_j|$, for $j = 1, \ldots, k$.*

2. *$Index(M, k)$ is the set of k-ary indices in M. The support of i is the set $supp(i) := \bigcup_j i_j$. The support-size is the cardinality of the support.*

3. *$M^{(k)}$ is the $\mathbf{F}_p$-vector space whose basis is the set $Index(M, k)$.*

4. *A vector from $M^{(k)}$ is definable (tacitly in $L(C)$) over a set $A \subseteq M$ if for each $a \in \mathbf{F}_p$ there is an $L(C)$-formula $\theta_a(\alpha)$ with parameters from $C \cup A$, $\alpha = (\alpha_1, \ldots, \alpha_k)$, such that the coefficient of the index $j = (j_1, \ldots, j_k)$ is a iff the formula $\theta_a(j)$ holds in M.*

5. *A family $X_i^M \subseteq M^{(k)}$ of sets of vectors from $Index(M, k)$, M ranging over finite $L(C)$-structures and $i \in Index(M, \ell)$, is* uniform *(or* definable*) iff there are $L(C)$-formulas $\theta_a(\gamma, \alpha)$, $a \in \mathbf{F}_p$, with no parameters other than C, with $\gamma = \gamma_1, \ldots, \gamma_\ell$, $\alpha = \alpha_1, \ldots, \alpha_k$ of the set-sort such that for every M and every $i \in Index(M, \ell)$ the vector X_i^M is defined in M by formulas $\theta_a(i, \alpha)$.*

The results proved later on hold for a more general notion of an index (roughly, a finite structure of any order with a bounded support) but we suppress such generality here as we do not have interesting applications for it.

Now we characterize the definability in a more combinatorial way.

Definition 1.3 *Let M be an $L(C)$-structure, $A \subseteq M$, and $i \in Index(M, r)$. $\mathbf{Sym}_C(M/A)$ is the group of permutations of M fixing point-wise the set $C \cup A$.*

The type of i over A, denoted $\mathbf{tp}_C(i/A)$, is the isomorphism type of the $L(C)$-structure

$$\langle C \cup supp(i) \cup A; \{c\}_{c \in C}, \{a\}_{a \in A}, i_1, \ldots, i_r \rangle$$

with the universe $C \cup supp(i) \cup A$.

The support-size of the type is the cardinality of the set $C \cup supp(i) \cup A$.

Note that over any M, types of r-ary indices over A are in a bijective correspondence with orbits of $\mathbf{Sym}_C(M/A)$ acting on $Index(M, r)$. The following lemma is a simple model-theoretic fact. Note that it is important that we allow quantification only over the element-sort.

Lemma 1.4 *Let $L(C)$, set A and r be fixed. Then the following two statements hold:*

1. *For any $L(C)$-formula $\psi(\alpha)$, $\alpha = \alpha_1, \ldots, \alpha_r$, with parameters from A there is a (necessarily finite) set S of types over A of r-ary indices such that for each $M \supseteq A$ large enough (depending on the formula ψ) and every $i \in Index(M, r)$*
$$M \models \psi(i) \quad \textit{iff} \quad \mathbf{tp}_C(i/A) \in S$$

2. *For any (necessarily finite) set S of types over A of r-ary indices there is a quantifier-free formula $\psi(\alpha)$ with parameters from A such that for every M and every $i \in Index(M, r)$*
$$M \models \psi(i) \quad \textit{iff} \quad \mathbf{tp}_C(i/A) \in S$$

In particular, let $X^M \subseteq Index(M, r)$, M ranging over $L(C)$-structures containing A, be a family of sets of r-ary indices. Then X^M is definable by $L(C)$-formula iff it is definable in terms of types (for M large enough).

Now we discuss an embedding of polynomial rings into vector spaces $M^{(r)}$.

Definition 1.5 *Let M be an $L(C)$-structure, and let $k \geq 1$.*

1. *$Var(M, k)$ is the set of variables indexed by elements of $Index(M, k)$ different from $\overline{\emptyset} = (\emptyset, \ldots, \emptyset)$. (The index $\overline{\emptyset}$ will represent* 1 *in monomials.)*

2. *$Mon(M, k, d)$ is the set of all monomials formed from $Var(M, k)$ and of degree at most d. Monomials from $Mon(M, k, d)$ are identified with indices from $Index(M, kd)$, utilizing the index $\overline{\emptyset}$ for monomials of degree $\ell \leq d$. In particular, monomial $x_{i^1} x_{i^2} \ldots x_{i^\ell}$, $i^t = (i^t_1, \ldots, i^t_k)$, is identified with the kd -ary index $(i^1_1, \ldots, i^1_k, \ldots, i^\ell_1, \ldots, i^\ell_k, \emptyset, \ldots, \emptyset)$.*

3. *A polynomial over M is a function from some $Mon(M, k, d)$ to $\mathbf{F}_p$. It is definable over a set $A \subseteq M$ if for each $a \in \mathbf{F}_p$ there is an $L(C)$-formula $\theta_a(\alpha)$ with parameters from $C \cup A$, $\alpha = (\alpha_{1,1}, \ldots, \alpha_{1,k}, \ldots, \alpha_{d,1}, \ldots, \alpha_{d,k})$ such that the coefficient of the monomial $x_{j_{1,1},\ldots,j_{1,k}} \cdot \ldots \cdot x_{j_{d,1},\ldots,j_{d,k}}$ is a iff the formula $\theta_a(j)$ holds in M.*

   $Poly(M, k, d)$ is the set of polynomials of degree at most d with variables from $Var(M, k)$.

4. *A family F^M_i, M ranging over finite $L(C)$-structures and $i \in Index(M, \ell)$, of polynomials with variables from $Var(M, k)$ and of degree at most d is* uniform *(or* definable*) iff there are $L(C)$-formulas $\theta_a(\gamma, \alpha)$, $a \in \mathbf{F}_p$, with no parameters other than C, with $\gamma = \gamma_1, \ldots, \gamma_\ell$, $\alpha = \alpha_{1,1}, \ldots, \alpha_{d,k}$ of the set-sort such that for every M and every $i \in Index(M, \ell)$ the polynomial F^M_i is defined in M by formulas $\theta_a(i, \alpha)$.*

We shall use the notations $F^M(i,j)$ or $F^M_{i,j}$ for the unique $a \in \mathbf{F}_p$ such that $\theta_a(i,j)$ holds in M, and we shall write $F^M_i = \sum_j F^M(i,j)x^j$, where x^j (sometimes written only as j) denotes the monomial corresponding to index j.

Note that, in particular, Lemma 1.4 says that a family F^M is uniform iff the value $F^M(i,j) \in \mathbf{F}_p$ depends only on the type $\mathbf{tp}_C((i,j))$ of the pair of indices (i,j) (for M large enough).

The reason for the somewhat unnatural embedding of monomials of degree $\ell \leq d$ into kd -ary indices is that we need a fixed number of variables α_{ij} for formulas $\theta_a(\alpha)$ defining polynomials.

The representation of monomials by indices does not account for the commutativity of variables. As we wish to work only in the commutative context we remedy the situation by including in all families F a particular uniform family COMM^k.

Definition 1.6 *$COMM^k$ is a uniform family indexed by pairs i,j of elements $Var(M,k)$, with the polynomial $COMM^k_{i,j}$ being $x_i x_j - x_j x_i$.*

We conclude this section with two technical lemmas.

Lemma 1.7 *Let $r \geq 1, s \geq 0$, let $A \subseteq M$, and let ω be a type of an r-index $i = (i_1, \ldots, i_r)$ over A, and let s be the support size of ω.*

Then for $|M| \geq s$ the cardinality of

$$\{i \in Index(M,r) \mid \mathbf{tp}_C(i/A) = \omega\}$$

modulo p depends only on the remainder of $|M|$ modulo p^ν, for any $p^\nu > r^{r^2/2}$.

Proof :

A type is given by $2^r - 1$ numbers $m_1, \ldots, m_{2^r-1}$ specifying the number of elements of $M \setminus (A \cup C)$ in all regions of the Venn diagram of $i_1, \ldots, i_r$ inside *supp*(i), together with a placement of constants from $C \cup A$ in these regions. Clearly $|C \cup A| + \sum_j m_j = s$, so if $|M| \geq s$ the number of r-indices for a given C, A with the type specified by m_j's is

$$K := \binom{M'}{m_1}\binom{M' - m_1}{m_2} \cdot \ldots \cdot \binom{M' - \sum_{j=1}^{2^r-2} m_j}{m_{2^r-1}}$$

$$= \frac{M'(M'-1) \cdot \ldots \cdot (M' + 1 - \sum_{j=1}^{2^r-1} m_j)}{\Pi_{j=1}^{2^r-1}(m_j!)}$$

where $M' := |M| - |A \cup C|$.

To determine $K \pmod p$ it is enough to know

$$M'(M'-1) \cdot \ldots \cdot (M' + 1 - \sum_{j=1}^{2^r-1} m_j)$$

modulo the power of p in the product in the denominator. As $\Pi_{j=1}^{2^r-1}(m_j!) \leq (r!)^{r/2} \leq r^{r^2/2}$, knowing M' modulo any power $p^\nu > r^{r^2/2}$ will do.

q.e.d.

The following lemma is a simple corollary of the previous statement. Recall that for polynomial $f \in Poly(M,k,t)$ and monomial $j \in Mon(M,k,t)$, f_j is the coefficient of j in f.

Lemma 1.8 *Let $A, B \subseteq M$ and let $k, t \geq 1$. Let $f, g \in Poly(M,k,t)$ be polynomials defined over A and B respectively.*

Then for M such that $|M| \geq |A \cup B \cup C| + k^2 t$ the value

$$\sum_{j \in Mon(M,k,t)} f_j \cdot g_j \ \in \ \mathbf{F}_p$$

depends only on $|M|$ (mod p^ν), for any $p^\nu > (kt)^{k^2t^2/2}$.

2 Ajtai's work on definability of solutions of uniform systems of linear equations

Ajtai [2] analysed[5] in a very interesting work the definability (in certain expansions M^* of the original structures M) of solutions of uniform systems of linear equations.

Definition 2.1 *Let $\nu \geq 1$ and $s \geq 0$ be fixed.*

1. *$L^{\nu,s}$ is the language expanding $L(C)$ by the following (element-sort) constants and predicates: constants e_0, e_1, unary predicates $R_{\nu,j}(x)$ for $j = 0, \ldots, p^\nu - 1$, and binary predicates $x \leq y$ and $D_i(x,y)$, for $i = 0, \ldots, s$.*

   $L^{\nu,s}$-formulas are assumed to maintain the property that no quantifier ranges over the set-sort.

2. *$T^{\nu,s}$ is $L^{\nu,s}$-theory with axioms:*

   (a) *$\leq$ is a linear ordering with e_0 and e_1 the first and the last elements respectively.*

   (b) *$R_{\nu,1}(e_0)$ holds and $R_{\nu,j}(a)$ implies $R_{\nu,j'}(a')$, whenever a' is the successor of a and $j' = j + 1$(mod p^ν).*

   (c) *$D_i(a,b)$ holds iff the distance between a and b is i.*

[5]The presentation of [2] is rather unfriendly. However, readers willing to suffer are rewarded.

It is clear that $T^{\nu,s}$ admits a form of quantifier elimination (verified by induction on the size of a formula, cf. [2, L. 8]).

Lemma 2.2 *Let ϕ be an $L^{\nu,0}$-formula. Then there is $s \geq 0$ depending only on ν and on the size of ϕ, and a quantifier-free $L^{\nu,s}$-formula ψ such that ϕ and ψ are equivalent in any model M^* of $T^{\nu,s}$ of sufficiently large cardinality.*

The lemma explains why $L^{\nu,0}$-definability of a property of structures M implies that the property depends only on the cardinality of M modulo p^ν. This is because if ϕ is a sentence then it is equivalent to a boolean combination of conditions $R_{\nu,j}(e_1)$, some $j < p^\nu$ (for M^* large enough).

Ajtai [2] proved a remarkable theorem that can be stated using our notion of uniformity as follows. We derive later (see Theorem 3.5) the equivalence of conditions 1 and 3 in the theorem from Lemma 1.8 and Theorem 3.3 (with effective bounds on *sufficiently large*).

Theorem 2.3 (Ajtai [2]) *Let $L(C)$ and $k, \ell \geq 1$ be given. Assume that a family F_i^M, $i \in Index(M, \ell)$, is a uniform family of linear polynomials with variables from $Var(M, k)$.*

Then there exists $\nu \geq 1$ (depending only on $|C|$, k, ℓ and tacitly on p) and a set $Q \subseteq \{0, \ldots, p^\nu - 1\}$ such that for every sufficiently large finite $L(C)$ - structure M the following three statements are equivalent:

1. *The system $\{F_i^M = 0\}_{i \in Index(M,\ell)}$ is solvable in $\mathbf{F}_p$.*
2. *The system $\{F_i^M = 0\}_{i \in Index(M,\ell)}$ has a solution in $\mathbf{F}_p$ that is $L^{\nu,0}$ - definable in any expansion M^* of M to a model of $T^{\nu,0}$. The definition of the solution is common for all M.*
3. *$|M| \equiv r \pmod{p^\nu}$ for some $r \in Q$.*

Ajtai states the theorem using the notion of a family induced by quadruples. For the benefit of a reader familiar with [1, 2] we now include a proof that our notion of uniformity (definability) coincides with that notion, and hence that our formulation is equivalent to the original one. (The theorem, however, is not needed anywhere else in the paper, and its part that we shall use will be actually deduced in full from our results.)

By Theorems 1, 2 (using Lemma 1.4 here) and Corollary 6 of [2] (and its extension discussed in the last section of [2], see also [1, Thm. 4]) it is enough to show that if a family is uniform then it is induced by a quadruple (as defined on pp.4-5 of [2]). In fact, we show that these two notions coincide.

A quadruple is in our setting equivalent to a finite set of the following data that we shall call a *pattern* (we give a definition for any $d \geq 1$ as it will be handy later on). A pattern is an ℓ-ary index i and a set $X \supseteq C \cup supp(i)$ such that $|X \setminus (C \cup supp(i))| \geq k^2 d$, and a mapping $u : Mon(X, k, d) \to \mathbf{F}_p$ such that $u(j) = u(j')$ if $\mathbf{tp}_C(j/supp(i)) = \mathbf{tp}_C(j'/supp(i))$.

The pattern uniquely determines a degree at most d polynomial with variables from $Var(M,k)$

$$\sum_j u(\pi_j(j))x^j$$

where $\pi_j \in \mathbf{Sym}_C(M/supp(i))$ is any permutation moving j into $Mon(X,k,d)$. Hence the coefficient of x^j depends only on $\mathbf{tp}_C(j/supp(i))$ and so, by Lemma 1.4, the polynomial is definable over $supp(i)$.

The pattern (X,i,u) corresponds to a uniform system of polynomials F as follows. The polynomials in F are indexed by $\pi(i)$ for $\pi \in \mathbf{Sym}_C(M)$, and the coefficient of x^j in the polynomial $F_{\pi(i)}$ is the coefficient of $x^{\pi^{-1}(j)}$ in the polynomial determined by the pattern. The system is uniform by Lemma 1.4 again, as $F(i,j)$ depends only on $\mathbf{tp}_C((i,j))$. Hence a quadruple (a finite collection of patterns) corresponds to a uniform family too.

On the other hand, a uniform system F is described by a quadruple as follows. For every type ω of an index i (there is finitely many of them) pick one i_ω such that $\mathbf{tp}_C(i_\omega) = \omega$. For X_ω take any set extending $C \cup supp(i_\omega)$ by k^2d elements, and define u_ω as follows:

$$u_\omega(j) := F(i_\omega, j)$$

The quadruple consists of all patterns $(X_\omega, i_\omega, u_\omega)$. Hence we have shown that every uniform system corresponds to a quadruple of patterns of bounded size (at most $|C| + \ell^2 + k^2$ for $d = 1$). Thus Ajtai's theorem applies.

3 Definable generators of permutation moduli

The ambient space for moduli will be the $\mathbf{F}_p$-vector space $M^{(r)}$ defined in Definition 1.2. Recall that $Poly(M,k,t)$ is a subspace of $M^{(kt)}$. Tabloid moduli (see below) will provide another examples of subspaces of $M^{(r)}$.

Let W be a subset of $M^{(r)}$. The group $\mathbf{Sym}_C(M)$ acts on $M^{(r)}$ and we say that W is *symmetric* iff it is closed under this action. A symmetric W that is also vector subspace of $M^{(r)}$ is thus an $\mathbf{F}_p[\mathbf{Sym}_C(M)]$-module, where $\mathbf{F}_p[\mathbf{Sym}_C(M)]$ is the group algebra (corresponding to the group algebra denoted $\mathbf{Z}_p\mathbf{S}_n$ in [2]; we use rather the notation of [20]).

We shall say that a set G *generates* W iff it generates W as an $\mathbf{F}_p[\mathbf{Sym}_C(M)]$ - module, and that it *generates* W *as a vector space* iff W is the $\mathbf{F}_p$-linear span of G.

We recall now several definitions and facts from [13, 12]. Let $\mu = (\mu_1, \ldots, \mu_s)$ be a *partition* of $N := |M|$, i.e., μ_i's are non-negative integers such that $\sum_i \mu_i = N$. A partition is *proper* iff $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_s$.

A μ-*tabloid* is an index that is an ordered s-tuple $T = (T_1, \ldots, T_s)$ of disjoint subsets of M such that $|T_i| = \mu_i$. A μ-*tableaux* t is a μ-*tabloid* T together with linear orderings of all T_i's. The underlying tabloid is denoted $\{t\}$ (in accordance with [13, 12]). Tabloid T has rows $T_1, \ldots, T_s$, and tableaux t has moreover also columns consisting of the first, second, etc. elements of the rows. The *support-size* of T is $\sum_{j\geq 2} |T_j|$.

The tabloid module M^μ is an $\mathbf{F}_p$-vector space whose basis is the set of all μ-tabloids. $\mathbf{Sym}_C(M)$ acts on tabloids and hence on elements of M^μ, and M^μ is $\mathbf{F}_p[\mathbf{Sym}_C(M)]$-module. Note that tabloid T is determined by the rows $T_2, \ldots, T_s$ and hence M^μ is $\mathbf{F}_p[\mathbf{Sym}_C(M)]$-isomorphic to a submodule of $M^{(r)}$, where $s-1 \leq r$ and $|T_i| \leq r$, all $i \geq 2$. The embedding sends a tabloid T to r-index $ind(T) := (T_2, \ldots, T_s)$.

On the other hand, $M^{(r)}$ is isomorphic to a submodule of a direct sum of tabloid moduli. To see this note that all possible 2^r intersections $Y_1 \cap \ldots \cap Y_r$, where Y_j is either i_j or its complement $M \setminus i_j$, form a partition of M that determines the r-index i (not all partitions of M correspond to some so we get a submodule). This encoding of r-indices by tabloids is formalized as follows.

For $X \subseteq M$ put $X^1 := X$ and $X^0 := M \setminus X$, and for $i \in Index(M, r)$ and $\epsilon \in \{0,1\}^r \setminus \{(0, \ldots, 0)\}$ put

$$i^\epsilon := \bigcap_{t \leq r} i_t^{\epsilon_t}$$

Note that i^ϵ are disjoint for varying ϵ, have size at most r, and that $i = i'$ iff $i^\epsilon = i'^\epsilon$ for all $\epsilon \in \{0,1\}^r \setminus \{(0, \ldots, 0)\}$. Define tabloid (not necessarily corresponding to a proper partition)

$$tab(i) := (M \setminus supp(j), j_1, \ldots, j_{2^r-1})$$

where $j \in Index(M, 2^r - 1)$ is an index whose coordinates are numbered by $\epsilon \in \{0,1\}^r \setminus \{(0, \ldots, 0)\}$ in the lexico-graphic order and such that the ϵ^{th} coordinate is the set i^ϵ.

Let μ^j range over partitions corresponding to all $tab(i)$, $i \in Index(M, r)$. As $|supp(i)| \leq r^2$, there are only finitely many such μ^j's (depending on r). The tab is an embedding of $M^{(r)}$ into $\bigoplus_j M^{\mu^j}$. The map ind is not its inverse but clearly definitions of $X \subseteq M^{(r)}$ and of $ind(tab(X))$ are inter-translatable (as rows in $ind(tab(i))$ correspond to the inner regions of the Venn diagram of i). Note that the support-sizes of i and $tab(i)$, and of T and $ind(T)$ are the same.

We use this correspondence to extend the notion of definability and uniformity from Section 1 to tabloid moduli: by definability from A of a subset W of a tabloid module $\bigoplus_j M^{\mu^j}$ is meant the $L(C)$-definability of $ind(W)$ from constants A. More generally, $v \in \bigoplus_\lambda M^\lambda$ is definable from A if for every λ the set $ind(pr_\lambda(v)) \subseteq ind(M^\lambda)$ is definable from A, where pr_λ is the projection on M^λ. In particular, each $ind(pr\lambda(v))$ has its own $L(C)$-formula defining it. Note that in the particular case $\bigoplus_\lambda M^\lambda = \bigoplus_j M^{\mu^j} = tab(Index(M, r))$ the notions are the same, as each μ^j corresponds to a different type of r-indices and hence all the (possibly different) definitions for $ind(pr_{\mu^j}(v))$ can be incorporated into one definition using a definition by cases distinguishing the type of μ^j.

In the more general direct sum some λ's may occur several times. Note also that $Poly(M,k,t) \subseteq M^{(kt)}$ and that definability results for $M^{(kt)}$ pull-back to $Poly(M,k,t)$ as it is itself definable in $M^{(kt)}$ (without parameters).

In our terminology Ajtai's theorem [2, Thms. 7 and 7'] says the following.

Theorem 3.1 (Ajtai [2]) *Let M be a finite $L(C)$-structure and let W be a submodule of $M^{(r)}$. Then there is $\nu \geq 1$ depending only on r, and a set $G \subseteq W$ generating W such that*

1. *The cardinality of G depends only on r but not on M.*
2. *Every element of G is $L^{\nu,0}$-definable in any expansion M^* of M to a model of $T^{\nu,0}$ by formulas of size bounded in terms of r only.*

We shall not use this theorem but we note its immediate corollary (that was in a special case also obtained by D.G.D.Gray [10, 11]). Explicit bounds on c, can be computed from our Theorem 3.3.

Corollary 3.2 *Given r there is c such that for any M, the module $M^{(r)}$ has at most c different submoduli.*

Our main technical result is the following theorem describing another form of defining sets of generators for submoduli of $M^{(r)}$. The construction is partly based on Ajtai's construction underlying Theorem 3.1 and partly on constructions from [13, 12] describing the Specht modules and their orthogonal complements as kernels and ranges of suitable linear maps.

We give a proof utilizing some notions related to tabloid moduli (and facts about them) from [13, 12] (these are also explained in the proof of Lemma 3.4).

Denote by $M^{(r,s)}$ the submodule of $M^{(r)}$ generated by those $i \in Index(M,r)$ whose support-size is at most s. So, in particular, $M^{(r)} = M^{(r,r^2)}$ and $M^{(r,0)} \cong \mathbf{F}_p$. Note that $M^{(r,s)}$ is a submodule of $M^{(r')}$ if $r \leq r'$ with an associated natural projection. For K a uniform family of linear polynomials K_i with variables indexed by r'-indices denote by $V(K^M)$ the vector subspace of $M^{(r')}$ consisting of solutions to the homogeneous system $\{K_i^M \overline{x} = 0\}_i$, and $V_{r,s}(K^M)$ the projection of $V(K^M)$ onto $M^{(r,s)}$.

Theorem 3.3 *Given $r, s \geq 1$ there are $c, \ell \geq 1$ and $r' \geq r$, uniform families $H^1, \dots, H^c$ of vectors from $M^{(r,s)}$ and uniform families $K^1, \dots, K^c$ of vectors from $M^{(r')}$ such that the following hold:*

1. *Vectors H_i^t and K_i^t in all H^t and K^t respectively, $t \leq c$, are indexed by ordered ℓ-tuples i such that the support-size of each i is at most $2^{s+1}s^2$.*

2. *For every M of size at least $|C| + 2^{s+1}s^2$, if $s < p$ then any submodule W of $M^{(r,s)}$ is generated as a vector space by one of the systems $(H^t)^M$.*

3. *For every M of size at least $|C| + 2^{s+1}s^2$, any submodule W of $M^{(r,s)}$ is $V_{r,s}((K^t)^M)$ for a K^t, $t \leq c$.*

Proof :
Assume $|M| \geq |C| + 2^{s+1}s^2$. We shall show that any submodule W of $M^{(r,s)}$ is a span of some H^M, where H is a uniform family indexed by ℓ-tuples of support-size at most $2^{s+1}s^2$. For a fixed ℓ there are a priori finitely many such uniform families as there are finitely many types of pairs (i,j) of indices of $H(i,j)$, and as we are over a finite field (this later condition is, in fact, not necessary but that follows only from the construction).

The existence of a suitable H is proved by an induction argument of structure similar in part to the proof of [2, Thms. 7 and 7'] (we modify the induction and use a construction from [12] in place of a duality argument in [2]). The construction yields families K^t in any characteristic and families H^t if $p > s$; this assumption is used only in Lemma 3.4 (part 2).

We shall first motivate the general set-up for the induction by considering a special case. This allows us also to introduce few necessary concepts.

Let $W \subseteq M^{(r,s)}$ and by the embedding tab we may identify W with a submodule of some sum $\bigoplus_j M^{\mu^j}$, where each μ^j-tabloid has support-size at most s. Consider the special case that, in fact, W is a submodule of one M^μ and that μ is proper.

By the James's submodule theorem [13, Thm.4.8] either $S^\mu \subseteq W$ or $W \subseteq (S^\mu)^\perp$, where S^μ is the Specht module. Let $\psi_{i,v}$ be the maps $M^\mu \to M^{\lambda^{i,v}}$ from [13, Def.17.10] (see the proof of Lemma 3.4 for the definition) and put

$$\eta := \bigoplus_{i=1}^{2^r-1} \bigoplus_{v=0}^{\mu_{i+1}-1} \psi_{i,v}$$

to be the map $M^\mu \to Y^\mu$, where we define $Y^\mu := \bigoplus_{i,v} M^{\lambda^{i,v}}$.

If $S^\mu \subseteq W$ take $U := \eta(W) \subseteq Y^\mu$, and assume that $G \subseteq U$ spans U. Then

$$H := \eta^{(-1)}(G) \cup S^\mu$$

spans W as $S^\mu = Ker(\eta)$ by [13, Cor.17.18]. In fact, in place of $\eta^{(-1)}(G)$ it is enough to have one representative for every fiber of η over G, as $\eta(u) = \eta(v)$ implies $u - v \in S^\mu$.

If on the other hand $W \subseteq (S^\mu)^\perp$, we use the maps $\varphi_{i,v} : M^{\lambda^{i,v}} \to M^\mu$ from [12] (they are denoted $\psi_{i,-v}$ there and their definition is recalled in the proof of Lemma 3.4). Put $\rho := \sum_{i,v} \varphi_{i,v} \circ pr_{\lambda^{i,v}} : Y^\mu \to M^\mu$ to be the sum of these maps (same range for i, v as in η). Take $U := \rho^{(-1)}(W) \subseteq Y^\mu$, and assume that

$G \subseteq U$ spans U. By [12, Cor.3], $\rho(Y^\mu) = (S^\mu)^\perp$, so $H := \rho(G)$ spans W in this case.

Hence in either case we construct a spanning set H for W from a spanning set G for U. To get a good estimate on the parameters of such families we shall proceed slightly differently. We shall construct set $H' \subseteq W$ that generates W (as a module) from a generating set G' for U, and in the construction we shall estimate the growth of the number of parameters needed to define any vector in the generating set. In the construction we also get an estimate for the cardinality of H'. Having such H' we shall construct uniform family H explicitly in Claim 3 (by creating one common definition as a definition by cases working for all vectors in H').

To make this work it will be enough to show:

1. S^μ is spanned by a uniform family indexed by $2s$-tuples.
2. For any $w \in Y^\mu$ definable over A, assuming $p > s$, an element of $\eta^{(-1)}(w)$ of M^μ is definable over some A' such that $|A' \setminus A| \leq 2s^2$. Hence the pull-back $\eta^{(-1)}(G')$ together with S^μ is generated by a set such that every vector in it is defined from the same tuples as those in G' augmented by $2s^2$ - tuples or from $2s$-tuples.
3. The image $\rho(w) \in M^\mu$ of any $w \in Y^\mu$ definable over A is also definable over A. Hence the image $\rho(G')$ is a generating set parametrised by the same indices as G'.
4. Some suitable G' exists.

This will be the basis of an induction argument as the module Y^μ is in a sense simpler than M^μ (it does not contain the irreducible factor $D^\mu := S^\mu/(S^\mu \cap S^{\mu\perp})$, cf. [13]). From these conditions the first three will be guaranteed by Lemma 3.4 and the last condition will amount to the induction hypothesis.

We describe first the inductive construction of H' and at the end estimate a bound ℓ for the number of parameters and construct H, and describe the construction of systems K^t.

For any μ let $\tilde{\mu}$ be the reordering of μ that is proper. So $\mu = \tilde{\mu}$ for proper μ. We assume that $\mu_1 \geq \sum_{j\geq 2} \mu_j$, so $\tilde{\mu}_1 = \mu_1$ ($|M| \geq 2s$ and the support-size of any starting μ^j is at most s, so this holds for all starting μ^j's). For any partition $\lambda^{i,v}$ occurring in the definition of η, $\tilde{\lambda}^{i,v}$ is lexico-graphically bigger than μ, provided μ is proper. Moreover, the support-size of $\lambda^{i,v}$ is at most the support size of μ.

We shall consider sets X of partitions where a partition may occur with repetitions; these will be called multi-sets. We shall study definability of elements of $\bigoplus_{\lambda \in X} M^\lambda$. Recall (from the beginning of the section) that $v \in \bigoplus_{\lambda \in X} M^\lambda$

is definable over A if each projection $pr_{M^\lambda}(v) \in M^\lambda$ (different projections for different occurrences of M^λ in the direct sum) is definable over A by a separate definition $B_\lambda(\overline{\alpha})$. For this purpose we may always replace any λ by $\tilde{\lambda}$ and permute suitably α's in B_λ, and thus assume without a loss of generality that all partitions in X are proper.

For a multi-set X of proper partitions all of which have the support-size at most s we describe the following operation. Let μ be the lexico - graphically minimal partition occurring in X that is different from the maximal $(N, 0, \ldots, 0)$ (if no such exists the operation is undefined). Take any occurrence of it in X and replace it by all $\lambda^{i,v}$ from the definition of η, replacing further those $\lambda^{i,v}$ that are not proper by $\tilde{\lambda}^{i,v}$. The new multi-set is denoted X^{succ}.

Consider the class $\mathcal{X}$ of all such multi-sets X that can be obtained by repeated applications of the operation to the set X_{min} of all partitions corresponding to all $tab(i)$, for all $i \in Index(M, r)$ of the support-size at most s (i.e., to the set of μ^j's at the beginning, with μ^j possibly replaced by $\tilde{\mu}^j$).

The property that Y occurs later than X in the sequential generation (that is unique) of $\mathcal{X}$ from X_{min} defines a strict linear order $Y \succ X$ on $\mathcal{X}$ (by the remark above about the lexico-graphic ordering of proper μ and $\tilde{\lambda}^{i,v}$), the multi-set M_{min} is $\succ$ - minimal, and the $\succ$ - maximal multi-set X_{max} has the form of several copies of the lexico-graphically biggest partition $(N, 0, \ldots, 0)$. Denote by $d(X)$ the number of $Y \succ X$; hence $d(X_{max}) = 0$ and $d(X_{min}) = |\mathcal{X}| - 1$.

Claim 1: $|X_{min}| \leq 2^{(r+2)s}$, $|\mathcal{X}| \leq 2^{(r+2)s} s^{2^s}$, *and any multi-set* $X \in \mathcal{X}$ *has the size (counting multiplicities) at most* $2^{(r+2)s} s^{2^s}$.

We estimate the cardinality of X_{min} by the number of partitions $\mu = (\mu_1, \ldots, \mu_{2^r})$ such that $\sum_{j \geq 2} \mu_j \leq s$. This is

$$\binom{s + 2^r - 1}{s} \leq (2^r + s)^s \leq (2^r + r^2)^s \leq 2^{(r+2)s}$$

There are at most 2^s different proper partitions of support-size $\sum_{j \geq 2} \mu_j \leq s$. Call the number of $\tilde{\mu}$ in the lexico-graphic ordering of such proper partitions *the level* of μ, level 1 corresponding to the lexico-graphically biggest $\tilde{\mu} = (N, 0, \ldots, 0)$ and the largest level h_{max} to the smallest $\tilde{\mu}$. X_{min} has at most $|X_{min}|$ elements in the maximal level $h_{max} \leq 2^s$. So after at most $|X_{min}|$ steps we get Y with all partitions in level $\leq h_{max} - 1$, and of size $|Y| \leq |X_{min}| \cdot (s-1)$. The factor $(s-1)$ comes from the fact that the number of $\lambda^{i,v}$'s is $\leq s - 1$. Getting rid of level $(h_{max} - 1)$ - partitions in Y needs at most $|Y|$ steps, increasing the size of Y to at most $|X_{min}| \cdot (s-1)^2$. So the whole process until reaching X_{max} needs at most

$$|\mathcal{X}| \leq \sum_{t=0}^{2^s} |X_{min}|(s-1)^t \leq 2^{(r+2)s} \cdot s^{2^s}$$

steps, and any Y occurring in it has the cardinality bounded by the same quantity

$$|Y| \leq 2^{(r+2)s} \cdot s^{2^s}$$

This proves the claim.

We shall prove by the $\succ$ - downwards induction for any $X \in \mathcal{X}$ the following statement:

> *for any submodule $U \subseteq \bigoplus_{\lambda \in X} M^\lambda$ there is $H' \subseteq U$ generating U and such that any vector in H' is definable from a $2s^2 d(X)$-tuple. Moreover, the size of H' is at most $2^{(r+2)s} s^{2^s} + d(X)$.*

Consider the initial case $X := X_{max}$. The module M^μ for $\mu = (N, 0, \ldots, 0)$ is just the field $\mathbf{F}_p$, so $\bigoplus_{\lambda \in X} M^\lambda = \mathbf{F}_p \oplus \ldots \oplus \mathbf{F}_p$, with $|X_{max}| \leq 2^{(r+2)s} \cdot s^{2^s}$ copies of $\mathbf{F}_p$.

Any particular vector in this module is definable without parameters (as there is a separate definition for any coordinate picking an element of $\mathbf{F}_p$). It is also enough to take $|X_{max}|$ of them in any generating set H', as that is the maximal dimension. This proves the initial case of the induction.

For the induction step assume that the statement holds for all $Y \succ X$. Let $X_0 \subseteq X$ be the multi-set of all occurrences of the partition μ that is replaced in the operation (i.e., μ is the lexico - graphically minimal one). Let

$$W \subseteq \bigoplus_{\lambda \in X} M^\lambda = \bigoplus_{\mu \in X_0} M^\mu \oplus \bigoplus_{\lambda \in X \setminus X_0} M^\lambda$$

Let $u \in \bigoplus_{\lambda \in X} M^\lambda$ and let t be a μ - tableaux and κ_t its signed column sum. Consider the coordinate $u_\lambda = pr_{M^\lambda}(u)$ of u in M^λ. By [13, L. 4.6 and C. 4.7], $u_\lambda \kappa_t = 0$ for all λ lexico - graphically bigger than μ (as the lexico - graphical ordering refines the partial ordering $\trianglerighteq$ used in [13], see [13, Defs. 3.2 and 3.4]), and $u_\lambda \kappa_t = c \cdot e_t$ for some $c \in \mathbf{F}_p$, if $\lambda = \mu$. Thus $u\kappa_t \in \bigoplus_{\lambda \in X} M^\lambda$ is a vector whose first $|X_0|$ coordinates are $\mathbf{F}_p$ - multiples of the polytabloid e_t and all other coordinates corresponding to $\lambda \in X \setminus X_0$ are zero.

Fix t a μ - tabloid and take $X_1 \subseteq X_0$ to be the set of all occurrences of μ's such that for some $u \in W$, $u\kappa_t$ has non - zero coordinate corresponding to the particular μ (and hence it is a non - zero $\mathbf{F}_p$ - multiple of e_t). Putting occurrences from X_1 as first $k := |X_1|$ occurrences of μ in the direct sum, we may assume that we have k vectors $u^1, \ldots, u^k \in W$ such that $u^i\kappa_t$ has first k coordinates $j = 1, \ldots, k$ of the form $c_{i,j} \cdot e_t$ and all other coordinates zero, with all $c_{i,i} \neq 0$.

Let $X_2 \subseteq X_1$ be the maximal set such that there are $u^1, \ldots, u^k \in W$ with $(u^i\kappa_t)_j = c_{i,j}e_t$ and an invertible $k \times k$ matrix E with entries from $\mathbf{F}_p$ such that the following holds; vectors v^i with entries $v^i_j := \sum_{m \leq k} c_{i,m}E_{m,j}$ for $j \leq k$, and $v^i_j := 0$ for $j > k$ are such that $v^i_j = 1$ iff $i = j \in X_2$ and $v^i_j = 0$ otherwise.

Vectors v^i are in $\bigoplus_{\lambda \in X} M^\lambda$ and the map defined by κ_t and E (map u to $u\kappa_t$ and then change the first k coordinates by E) maps W onto a submodule $W' \subseteq \bigoplus_{\lambda \in X} M^\lambda$.

For $i \leq |X_0|$ denote by $pr_i(W')$ the projection of W' on the ith copy of M^μ in the direct sum; we assume that first $k' := |X_2| \leq k$ coordinates corresponds to elements of X_2.

Claim 2: *For $i \leq k'$, $pr_i(W') \supseteq S^\mu$ and, in fact,*

$$0 \oplus \ldots \oplus 0 \oplus S^\mu \oplus 0 \oplus \ldots \oplus 0 \subseteq W'$$

with S^μ on the ith place.

For $k' < i \leq k$ it holds that $pr_i(W') \subseteq (S^\mu)^\perp$.

By the construction of E, for $i \leq k'$ the module W' contains the vector

$$0 \oplus \ldots \oplus 0 \oplus e_t \oplus 0 \oplus \ldots \oplus 0$$

with the polytabloid e_t at the ith place. This proves the first part of the claim. For the second part note that by the James submodule theorem $S^\mu \subseteq pr_i(W')$ or $pr_i(W') \subseteq (S^\mu)^\perp$. If the former happened then a vector $u \in W'$ exists with $pr_i(u) = e_t$, i.e. $pr_i(u\kappa_t) \neq 0$ contradicting the maximality of X_2.

The map determined by E is clearly 1 - to - 1 and both the map and its inverse are definable without parameters. This is because it is a fixed size matrix, so its entries from $\mathbf{F}_p$ are just explicitly listed by the definition. Thus any generator set for W' can be pulled to W without extra parameters. We may thus assume without a loss of generality that W itself has the properties of W'. Namely, there are $k \geq 1$ and $k \geq k' \geq 0$ such that

1. For $1 \leq i \leq k'$; $pr_i(W) \supseteq S^\mu$ and

   $$0 \oplus \ldots \oplus 0 \oplus S^\mu \oplus 0 \oplus \ldots \oplus 0 \subseteq W$$

   with S^μ at the ith place.

2. For $k' < i \leq k$: $pr_i(W) \subseteq (S^\mu)^\perp$.

Now consider two cases, similarly as in the special case earlier. Case 1 is when $k' \geq 1$. Map W to a submodule U of

$$\bigoplus_{\mu \in X_2} \bigoplus_{i,v} M^{\lambda^{i,v}} \oplus \bigoplus_{\lambda \in X \setminus X_2} M^\lambda$$

applying η to the first k' coordinates and the identity to all other coordinates (with i, v ranging as in the definition of η). Call this map η'. Let Y be the multi-set of partitions so obtained. As $Y \succ X$, by induction assumption we have $G' \subseteq U$ generating U, such that any vector in G' is definable from a $2s^2 d(Y)$ - tuple of parameters.

$Ker(\eta') = S^\mu \oplus \ldots \oplus S^\mu$ and so

$$\eta'^{(-1)}(G') \cup \{0 \oplus \ldots \oplus 0 \oplus S^\mu \oplus 0 \oplus \ldots \oplus 0 \mid S^\mu \text{ in the } i\text{th place } \ i \leq k'\}$$

generates W. By parts 1 and 2 of Lemma 3.4 we can take H' consisting of a representative of each fiber of η' over G', and of a generator of each $0 \oplus \ldots \oplus 0 \oplus S^\mu \oplus 0 \oplus \ldots \oplus 0$, such that all elements of H' are definable from a $2s^2 d(Y) + 2s^2 = 2s^2(d(Y)+1) \leq 2s^2 d(X)$ - tuple. H' generates the same set as $\eta'^{(-1)}(G') \cup S^\mu \oplus \ldots \oplus S^\mu$ which is W. Moreover, $|H'| \leq |G'| + k' \leq 2^{(r+2)s} s^{2^s} + d(Y) + k' \leq 2^{(r+2)s} s^{2^s} + d(X)$.

In Case 2 $k' = 0$, i.e., $pr_i(W) \subseteq (S^\mu)^\perp$ for all $i \in X_0$. Let Y be the multiset obtained by removing all μ's from X by the operation. Hence $\bigoplus_{\lambda \in X} M^\lambda$ changes to

$$\bigoplus_{\mu \in X_0} \bigoplus_{i,v} M^{\lambda^{i,v}} \oplus \bigoplus_{\lambda \in X \setminus X_0} M^\lambda$$

Let ρ' be the map

$$\bigoplus_{\mu \in X_0} \bigoplus_{i,v} M^{\lambda^{i,v}} \oplus \bigoplus_{\lambda \in X \setminus X_0} M^\lambda \to \bigoplus_{\lambda \in X} M^\lambda$$

applying ρ to the each of the $|X_0|$ $\bigoplus_{i,v} M^{\lambda^{i,v}}$-parts of $\bigoplus_{\xi \in Y} M^\xi$and the identity to the rest (recall the map ρ from the sixth paragraph of the proof of the theorem). Hence $Rng(\rho') = (S^\mu)^\perp \oplus \ldots \oplus (S^\mu)^\perp \bigoplus_{\lambda \in X \setminus X_0} M^\lambda$ and so if G' generates $\rho'^{(-1)}(W)$, $H' := \rho'(G')$ generates W. By the induction assumption and by part 3 of Lemma 3.4, every vector in H' is definable from $2s^2 d(Y) \leq 2s^2 d(X)$ - tuples.

Note that Lemma 3.4 requires the hypothesis that $|M| \geq |C \cup A| + 2s^2$, i.e., $|M| \geq |C| + 2^{s+1}s^2$ suffices as all $|A| \leq (2^s - 1)2s^2$.

Claim 3: *Let $H' \subseteq W \subseteq \bigoplus_\lambda M^\lambda$ be a set generating W and such that*

1. *Every vector in H' is definable from a $2^{s+1}s^2$ - tuple.*
2. *$|H'| \leq S$.*

Then there is a uniform family H that generates W as a vector space and that is indexed by $(\lceil \log_2(S) \rceil + 1 + 2^{s+1}s^2)$ - tuples of support-size at most $2^{s+1}s^2$.

To prove the claim note first that different vectors in H' might be definable from same index i by different definitions, say by $A_{v,\lambda}(i,j)$ ($v \in H'$, λ a coordinate, j running over indices corresponding to λ - tabloids, and i the parameter index), while we need one definition $B_\lambda(i_v, j)$ (i_v an index of parameters from which v is definable by B). B_λ defines the $\mathbf{F}_p$ - coordinate of λ - tabloid $tab(j)$ in the M^λ - part of the direct sum). This is arranged as follows.

Take ℓ such that $2^{\ell-1} \geq |H'|$. Any ℓ - tuple $i = (i_1, \ldots, i_\ell)$ determines $(\ell - 1)$ - tuple of bits $i^* \in \{0,1\}^{\ell-1}$ by: $i^*_j = 1$ if $i_j = i_{j+1}$ and $i^*_j = 0$ otherwise, for $j < \ell$. To different $v \in H'$ assign different $v^* \in \{0,1\}^{\ell-1}$ and take a definition by cases $B_\lambda(i,j)$, $i = (i', i'')$ an $(\ell + 2^{s+1}s^2)$ - tuple, to be $A_{v,\lambda}(i'', j)$ if the support-size of i' is ≤ 2 and $i'^* = v^*$ and $i'' = i_v$, and identically zero otherwise.

Hence we have one definition $B = (B_\lambda)_\lambda$ such that for any $v \in H'$ there is an $(\ell + 2^{s+1}s^2)$ - tuple i of support-size $\leq 2^{s+1}s^2$ such that $pr_{M^\lambda}(v)$ is definable by B_λ from i. Now take H to be the set of vectors defined by the same definition B but with i running over all possible $(\ell + 2^{s+1}s^2)$ - tuples. Clearly H' generates every element of H, and H is symmetric. So H generates as a vector space the same module that H' generates as a module, i.e. the module W. This proves the claim.

Now we can finish the proof of the theorem. By the inductive construction we have a generating set H' for W satisfying the hypothesis of Claim 3 with $S \leq 2^{(r+2)s}s^{2^s}$. Hence the required H exists by Claim 3.

The systems K^t are obtained as follows. The module W is transformed in the construction (by η or $\rho^{(-1)}$) to submodules W_u of $\bigoplus_{\lambda \in X_u} M^\lambda$, with $X_0, \ldots, X_m$ $\succ$-listing $\mathcal{X}$. Denote by $y^{(u)}$ tuples of variables indexed by tabloid generators of $\bigoplus_{\lambda \in X_u} M^\lambda$. In particular, $x := y^{(0)}$ are variables indexed by μ-tabloids.

Hence, $x \in W$ iff there exists vectors $y^{(u)}$ in all W_u, $u = 0, \ldots, m-1$, such that $\eta(y^{(u)}) = y^{(u+1)}$ or $y^{(u)} = \rho(y^{(u+1)})$ respectively. Any particular sequence of theses equations gives one linear system K^t; the estimate to their number and to the number of parameters follows from Claim 1 analogously as for systems H^t.

q.e.d.

We should remark on the issue of constants C. Namely, the theory in [13, 12] we use is developed for $\mathbf{F}_p[\mathbf{Sym}(M)]$ - moduli while we use $\mathbf{F}_p[\mathbf{Sym}_C(M)]$ - moduli. To see that this is, in fact, equivalent note that an $\mathbf{F}_p[\mathbf{Sym}_C(M)]$ - module M^μ, $\mu = (\mu_1, \ldots, \mu_h)$, is isomorphic to a direct sum of $\mathbf{F}_p[\mathbf{Sym}(M \setminus C)]$ - moduli of the form M^λ, for all partitions λ of $M \setminus C$ into h rows such that $\lambda_i \leq \mu_i$ (all i), having one copy of M^λ in the direct sum for any partition of C into h classes $C_1, \ldots, C_h$ of sizes $\mu_i - \lambda_i$'s. In particular, the copy of M^λ in the sum corresponding to λ and to the partitioning of C into $C_1, \ldots, C_h$ represents the submodule of M^μ generated as a vector space by tabloids $T = (T_1, \ldots, T_h)$ such that $C_i \subseteq T_i$, all i. Moreover, the isomorphism between the $\mathbf{F}_p[\mathbf{Sym}_C(M)]$ - module and the direct sum clearly furnishes a translation between definitions of elements.

The notions considered in the next lemma are defined in [13, 12]. We recall the definitions in the proof of the lemma.

Lemma 3.4 *Let $\mu = (\mu_1, \ldots, \mu_{2^r})$ be a proper partition such that $\sum_{j\geq 2}\mu_j \leq s$. Assume $A \subseteq M$, and $|M| \geq |C \cup A| + 2s^2$.*

1. *Let $e_t \in S^\mu$ be any polytabloid from M^μ. Then $ind(e_t)$ is definable from a $2s$ - tuple.*
2. *Let $\psi_{i,v} : M^\mu \to M^{\lambda^{i,v}}$ be the linear maps from [13, Def. 17.10], and let $\eta := \bigoplus_{i=1}^{2^r-1}\bigoplus_{v=0}^{\mu_{i+1}-1}\psi_{i,v}$. Assume $s < p$. Let $w \in Rng(\eta) \subseteq \bigoplus_{i,v} M^{\lambda^{i,v}}$ be definable from A. Then there is $u \in \eta^{(-1)}(w) \subseteq M^\mu$ definable from some A' such that $|A' \setminus A| \leq 2s^2$.*
3. *Let $\varphi_{i,v} : M^{\lambda^{i,v}} \to M^\mu$ be the linear maps from [12], denoted $\psi_{i,-v}$ there. Let $\rho := \sum_{i,v}\varphi_{i,v} \circ pr_{\lambda^{i,v}}$ (same range for i, v as in part 2), and let $w \in \bigoplus_{i,v} M^{\lambda^{i,v}}$ be definable from A. Then $\rho(w)$ is also definable from A.*

Proof :

Part 1. Assume that $\mu = (\mu_1, \mu_2, \ldots, \mu_{2^r})$. By [13, 4.5] S^μ is generated by any one polytabloid e_t. Here t is some tableau (defining the tabloid $\{t\}$) and $e_t = \{t\}\kappa_t$, where $\kappa_t = \sum_{\pi\in C_t} sgn(\pi)\pi$ is the signed column sum. C_t is the subgroup of $\mathbf{Sym}_C(M)$ that fixes set-wise all columns of t.

Consider the elements of the first s columns of t (there are $\leq 2s$ of such elements) as a parameter $2s$ - tuple $i = (i_1, \ldots, i_{2s})$. We think of it as defining a tableau t_i with 2^r rows

$$(i_1, \ldots, i_s, X),\ (i_{s+1}, \ldots, i_{s+\mu_2}),\ \ldots,\ (i_{s+\sum_{2\leq j\leq 2^r-1}\mu_j}, \ldots, i_{s+\sum_{2\leq j\leq 2^r}\mu_j})$$

where X is an arbitrary ordering of $M \setminus supp(i)$. The ordering of X is, in fact, irrelevant as κ_{t_i} depends only on i (only the first s columns of t_i might be non-trivial). The column stabilizer group C_{t_i} is just a finite subgroup of the symmetric group of $supp(i)$ of size depending on s, and hence the signed column sum κ_{t_i} can be listed explicitly (using i as parameters).

Denote by v_i the vector obtained from e_{t_i} by the ind-operation. As any e_{t_i} generates S^μ, v_i generates $ind(S^\mu)$ and the uniform system $\{v_i\}_i$ generates $ind(S^\mu)$ as a vector space.

Part 2. The map η is the direct sum of some maps $\psi_{i,v}$, $1 \leq i < 2^r$, $0 \leq v < \mu_{i+1}$ (cf. [13, Def. 17.10]). The map $\psi_{i,v}$ maps M^μ, $\mu = (\mu_1, \ldots, \mu_{2^r})$, to $M^{\lambda^{i,v}}$, $\lambda^{i,v} = (\mu_1, \mu_2, \ldots, \mu_{i-1}, \mu_i + \mu_{i+1} - v, v, \mu_{i+2}, \ldots, \mu_{2^r})$, by sending μ - tabloid T to $\sum\{T' \mid T' \in X_T\}$, where X_T is the set of all $\lambda^{i,v}$ - tabloids T' that agree with T on all rows except on the ith and the $(i+1)$st ones, and the $(i+1)$st row of T' is a subset of the $(i+1)$st row of T of size v.

We shall utilise moduli $S^{\mu^*,\mu}$ similarly to the proof of [2, L.35]. Let (μ^*, μ) be a pair of partitions with 2^r rows such that $\mu_1^* = \mu_1$, $\mu_i^* \leq \mu_i$ all i, and such that μ^* is also proper (hence μ^* is a partition of an integer not greater than N). Let C_t^*, t a μ - tableaux, be a subgroup of C_t fixing point-wise all elements of t outside μ^*. Let $S^{\mu^*,\mu}$ be the module spanned by all $e_t^{\mu^*,\mu} := \{t\}(\sum_{\pi \in C_t^*} sgn(\pi)\pi)$. Note that $e_t^{0,\mu} = \{t\}$ (0 is the partition with all except the first row empty) and $e_t^{\mu,\mu} = e_t$, so $S^{0,\mu} = M^\mu$ and $S^{\mu,\mu} = S^\mu$. (These modules are defined in [13, Def. 17.2 and 17.4].)

Assume $\mu^* \neq \mu$ and that $i > 1$ is the first row such that $\mu_{i-1}^* = \mu_{i-1}$ but $\mu_i^* < \mu_i$. Define (see [13, Def. 15.10]) two new pairs:

1. Pair $(\mu^*, \mu R_i)$: change μ_i to μ_i^* and μ_{i-1} to $\mu_{i-1} + \mu_i - \mu_i^*$.

2. Pair $(\mu^* A_i, \mu)$: change μ_i^* to $\mu_i^* + 1$ if $\mu_i^* + 1 \leq \mu_{i-1}^*$, or otherwise put the new pair to be $(\mu^* A_i, \mu) := (0, 0)$.

The operations make sense also for μ - tableaux. Here tR_i is t with the $\mu_i - \mu_i^*$ elements of the ith row of t that are outside μ_i^* moved to the end of the $(i-1)$st row.

The crucial facts are [13, Thm. 17.13] that $S^{\mu^*,\mu}\psi_{i-1,\mu_i^*} = S^{\mu^*,\mu R_i}$ and $S^{\mu^*,\mu} \cap Ker(\psi_{i-1,\mu_i^*}) = S^{\mu^* A_i,\mu}$. This will link these moduli with the map η. Map ψ_{i-1,μ_i^*} affects only two rows of a tabloid and thus the next claim can be rephrased (and proved) as a statement about hyper-graphs. We shall follow, however, a more direct approach related to [2, L. 35].

Claim 1: *Let $w' \in S^{\mu^*,\mu R_i}$ be definable over A and assume that $w' = \psi_{i-1,\mu_i^*}(u')$ for some $u' \in S^{\mu^*,\mu}$. Assume $p > s$. Then there is $\omega \in S^{\mu^*,\mu}$ definable over some $A' \supseteq A$ such that $w' = \psi_{i-1,\mu_i^*}(\omega)$ and $|A' \setminus A| \leq 2s$.*

We are going to use the identity (for a μ - tableaux t)

$$\psi_{i-1,\mu_i^*}(e_t^{\mu^*,\mu}) = e_{tR_i}^{\mu^*,\mu R_i}$$

from [13, L. 17.12,p.68]. First we need to discuss encoding of polytabloids $e_t^{\mu^*,\mu}$ by indices. In particular, encode $e_t^{\mu^*,\mu}$ by an index, denoted $I(t, \mu^*, \mu)$

$$\xi = (ind(\{t\}), \xi_1, \ldots, \xi_{\mu_2})$$

where tuple $ind(\{t\})$ is extended by sets ξ_j, the sets of elements of t in the jth column that are inside μ^*. This generalizes the encoding of both e_t $(= e_t^{\mu,\mu})$ and $\{t\}$ $(= e_t^{0,\mu})$ modulo one modification. Namely, all codes ξ as well as codes for e_t's list first μ_2 elements of the first row (they are all in μ^* as $\mu_1^* = \mu_1$) while $j = ind(\{t\})$ does not. To remedy this let $A' \supseteq A$ be a set of $2s$ new constants not in $A \cup C$. To get from j a code ξ for $e_t^{0,\mu}$ pick first (in some fixed ordering of A') μ_2 elements of A' not occurring in $supp(j) \cup A \cup C$ and place them into the first row. As $\mu_2 \leq s$ and $|supp(j)| \leq s$ this is always possible. Clearly, $ind(e_t^{0,\mu}) = j$.

Note that two tableaux t, t' yield same code ξ iff they differ only on ordering of the rows outside μ^* iff $e_t^{\mu^*,\mu} = e_{t'}^{\mu^*,\mu}$. We shall denote such polytabloids $e_\xi^{\mu^*,\mu}$. Also, $I(t, \mu^*, \mu)$ determines $I(tR_i, \mu^*, \mu)$. We denote these codes sometimes ξR_i. Such ξ need not to be unique but their number $\binom{\mu_{i-1}-\mu^*_{i-1}+\mu_i-\mu^*_i}{\mu_i-\mu^*_i}$ is non-zero modulo p (as $s < p$) and so we can take all ξ's with an appropriate weight in $\mathbf{F}_p$.

Let T' be a μR_i - tabloid with $j' = ind(T')$. Let $\xi' := \xi R_i$ be a code of a $(\mu^*, \mu R_i)$ - polytabloid $e_{\xi'}^{\mu^*,\mu R_i}$. The coefficient of T' in $e_{\xi'}^{\mu^*,\mu R_i}$ depends only on $\mathbf{tp}_C(j', \xi')$. By that, if $w' = \sum_{T'} a'_{T'} T'$ is from $S^{\mu^*,\mu R_i}$ with $a'_{T'}$ definable over A, there are $b'_{\xi'}$ definable over A' such that $w' = \sum_{\xi'} b'_{\xi'} e_{\xi'}^{\mu^*,\mu R_i}$.

Using the identity mentioned earlier, we have

$$\psi_{i-1,\mu_i^*}(\sum_\xi b_\xi e_\xi^{\mu^*,\mu}) = \sum_{\xi'} b'_{\xi'} e_{\xi'}^{\mu^*,\mu R_i} = w'$$

with $b_\xi := b'_{\xi R_i}$. It remains to show that a_T, T ranging over μ - tabloids, such that $\sum_T a_t T = \sum_\xi b_\xi e_\xi^{\mu^*,\mu}$ is definable over A'. The value b_ξ depends only on $\mathbf{tp}_C(\xi/A')$ and the coefficient of T in $e_\xi^{\mu^*,\mu}$ only on $\mathbf{tp}_C(ind(T), \xi/A')$. Moreover, by Lemma 1.7, the number of ξ with a given type over A' depends only on $|M| \pmod{p^\nu}$, some $\nu \geq 1$, provided M is at least as big as the support - size of the type $\mathbf{tp}_C(\xi/A')$ (which it is by the hypothesis of the lemma that $|M| \geq |A \cup C| + 2s^2$). Thus a_T are definable over A' as well, with possibly different definitions for each remainder class of $|M| \pmod{p^\nu}$. (This argument is similar to part 3 of the proof of the lemma.) This yields the claim.

Let $\mu^1, \ldots, \mu^k$ be the sequence of proper partitions

$$(\mu_1, 0, \ldots, 0), (\mu_1, 1, 0, \ldots, 0), \ldots, (\mu_1, \mu_2, 0, \ldots, 0),$$

$$(\mu_1, \mu_2, 1, 0, \ldots, 0), \ldots, (\mu_1, \mu_2, \ldots, \mu_{2^r})$$

obtained by filling $(\mu_1, 0, \ldots, 0)$ to μ first in the 2nd row, then in the 3rd row, etc. Note that $k \leq s$.

Let $\psi^1, \psi^2, \ldots$ be the sequence of maps

$$\psi_{1,0}, \psi_{1,1}, \ldots, \psi_{1,\mu_2-1}, \psi_{2,0}, \psi_{2,1}, \ldots, \psi_{2,\mu_3-1}, \ldots, \psi_{2^r-1,\mu_{2^r}-1},$$

up to $\psi_{j_0,0}, \ldots, \psi_{j_0,\mu_{j_0+1}-1}$, where j_0 is the last row j such that $\mu_{j+1} > 0$.

Fix any $u \in M^\mu = S^{0,\mu}$ such that $\eta(u) = w$. Assume that w is defined over A. We are going to modify u suitably.

Claim 2: *For any $j \geq 1$ there is $v_j \in M^\mu$ such that*

1. *v_j is definable over some $A^j \supseteq A$ such that $|A^j \setminus A| \leq j \cdot 2s$.*

2. $\forall i \leq j; \psi^i(u) = \psi^i(v_j)$.

The claim is proved by induction on j. If $j = 1$ take w' to be the projection of w on $M^{\lambda^{1,0}}$. So

$$w' = \psi_{1,0}(u) = \psi^1(u) \in M^{\lambda^{1,0}}$$

is also definable from A. Apply Claim 1 with $u' := u$ and $w' := \psi_{1,0}(u)$ and put v_1 to be the ω provided by the claim.

Assume the claim holds for j. To prove it for $j+1$ take $u_j := u - v_j$ and note that as $\psi^i(u) = \psi^i(v_j)$ for all $i \leq j$, it holds that $u_j \in \bigcap_{i \leq j} Ker(\psi^i)$. By the definition of the sequences $\mu^1, \mu^2, \ldots$ and $\psi^1, \psi^2, \ldots$ and by [13, Thm. 17.13] mentioned earlier, $\bigcap_{i \leq j} Ker(\psi^i) = S^{\mu^{j+1}, \mu}$.

We want $v_{j+1} \in M^\mu$ such that $\delta := v_{j+1} - v_j \in S^{\mu^{j+1}, \mu}$, and $\psi^{j+1}(v_{j+1}) = \psi^{j+1}(u)$ (as then $\psi^i(v_{j+1}) = \psi^i(v_j + \delta) = \psi^i(v_j) = \psi^i(u)$, all $i \leq j$ as well).

Note that $\psi^{j+1}(u), \psi^{j+1}(v_j)$ are defined over A^j (this is analogous to part 3) so $w' := \psi^{j+1}(u - v_j)$ is too. Apply Claim 1 with $u' := u_j \in S^{\mu^{j+1}, \mu}$. The claim gives $\delta \in S^{\mu^{j+1}, \mu}$ definable over some suitable $A' \supseteq A^j$, hence $v_{j+1} := v_j + \delta$ is definable over $A^{j+1} := A'$ as well and has the required properties.

Now take $v_k \in M^\mu$, k the length of the sequence of ψ's. By Claim 2 $\psi^i(v_k) = \psi^i(u)$ for all i, so

$$\eta(v_k) = \eta(u) = w$$

and v_k is defined over A^k with $|A^k \setminus A| \leq k \cdot 2s \leq 2s^2$. This proves part 2 of the lemma.

Part 3.

The maps $\varphi_{i,v} : M^{\lambda^{i,v}} \to M^\mu$ are defined (in [12] they are denoted $\psi_{i,-v}$; we use the other notation to avoid confusion with ψ's from part 2) as follows: $\varphi_{i,v}$ maps a $\lambda^{i,v}$ - tabloid T to $\sum_{T' \in X_T} T'$, where X_T is the set of all μ - tabloids agreeing with T in all except the ith and the $(i+1)$st rows, with the $(i+1)$st row of T' being the $(i+1)$st row of T together with some $\mu_{i+1} - v$ elements of the ith row of T.

Assume now that $w \in \bigoplus_{i,v} M^{\lambda^{i,v}}$ is a vector defined over A. Write w

$$w = \sum_{i,v} w_{i,v}$$

as the sum of its $M^{\lambda^{i,v}}$ - parts, and write a particular $w_{i,v}$

$$w_{i,v} = \sum_T c_T T$$

with T running over $\lambda^{i,v}$ - tabloids. Then

$$\varphi_{i,v}(w_{i,v}) = \sum_T c_T \ \varphi_{i,v}(T) = \sum_T c_T (\sum_{T' \in X_T} T')$$

where X_T are the μ - tabloids occurring in the definition of $\varphi_{i,v}(T)$, and this equals to

$$\sum_{T'}[\sum_{T:T'\in X_T} c_T]\cdot T'$$

Possible sets X_T depend on T' and types of some suitable T (and their multiplicities) depend on the type of T'. So the $\mathbf{F}_p$ - value in $[\ldots]$ depends only on $\mathbf{tp}_C(T'/A)$ as well, by Lemma 1.7, for any particular remainder class of N modulo p^ν. To apply Lemma 1.7 we need $|M|$ at least the support-size of $\mathbf{tp}_C(T/A)$, i.e., at least $|C\cup A|+s$. (Note that this gives different definitions for $\varphi_{i,v}(w_{i,v})$ for different remainder classes.) This proves part 3 of the lemma.

q.e.d.

We now derive an effective version of Ajtai's theorem 2.3 from Theorem 3.3. (It is clear that the generating set H' constructed in the proof of Theorem 3.3 satisfies the conditions posed on G in Theorem 3.1, so we get that theorem as well.) We restate the theorem (the equivalence of its parts 1 and 3) as we incorporate bounds. The proof follows the original proof of [2] except that we use our Theorem 3.3 in place of Theorem 3.1.

Theorem 3.5 *Let $L(C)$ and $k,\ell,s,z\geq 1$ be given. Assume that a family F is a uniform family indexed by ℓ - indices i of support - size $\leq z$ of linear polynomials F_i with variables indexed by k - indices of support size $\leq s$.*

Then there exists $\nu\geq 1$ (depending only on $|C|$, z, s and tacitly on p) and a set $Q\subseteq\{0,\ldots,p^\nu-1\}$ such that for every M such that $|M|\geq |C|+2z+2^{s+1}s^2+s$ the following two statements are equivalent:

1. *The system $\{F_i^M=0\}_i$ is solvable in $\mathbf{F}_p$.*
2. *$|M|\equiv r$ (mod p^ν) for some $r\in Q$.*

Proof :

Let F be a family of linear equations satisfying the hypothesis of the theorem. Let i_0 be an index such that F_{i_0} has a non-zero absolute coefficient c_{i_0} (if all F_i's are homogeneous there is nothing to prove as $F=0$ has the trivial solution). By replacing any F_i by $F_i-c_iF_{i_0}c_{i_0}^{(-1)}$, c_i the absolute coefficient of F_i, we may assume without a loss of generality that all equations of F are homogeneous except F_{i_0}. To maintain the uniformity of the system we add $supp(i_0)$ to C.

Now consider two homogeneous systems F^1 and F^2. System F^1 is obtained from F by deleting the equation $F_{i_0}=0$, system F^2 by changing its absolute coefficient to zero. Denote by $V(F^1)$, $V(F^2)$ the sets of $\mathbf{F}_p$ - solutions of the systems. As both are homogeneous, the sets are vector spaces. As both systems are uniform, the sets are symmetric w.r.t. $\mathbf{Sym}_C(M/supp(i_0))$ and thus they are $\mathbf{F}_p\mathbf{Sym}_C(M/supp(i_0))$ - moduli.

Clearly, F has a solution iff $V(F^2) \not\subseteq V(F^1)$. By Theorem 3.3, for $|M| \geq |C| + z + 2^{s+1}s^2 \geq |C \cup supp(i_0)| + 2^{s+1}s^2$, this is equivalent to the statement that there is a vector $v \in M^{(k)}$ such that

1. v is a solution of $F^1 = 0$.
2. v is not a solution of $F^2 = 0$.
3. v is definable from a $2^{s+1}s^2$ - tuple.

There are only finitely many v satisfying the last condition so the whole statement is finite. It is enough to show that for any particular definition of a vector v satisfying 3., whether or not the vector defined by the definition satisfies also the first two conditions depends only on $|M|$ (mod p^ν), for some suitable p^ν.

This is verified via Lemma 1.8. Let $f \in F^1$ be a linear polynomial defined from an ℓ - index of support - size $\leq z$ (over $supp(i_0)$). By Lemma 1.8, whether or not v satisfies $f = 0$ depends only on p^ν, provided $|M| \geq |C|+z+z+2^{s+1}s^2+s$ ($|C| + z$ is for $|C \cup supp(i_0)|$, z is for parameters of f, $2^{s+1}s^2$ for parameters of v and s for indices of the variables). This proves Theorem 3.5.

q.e.d.

Note that the same argument yields also an upper bound for ν; however, we do not need it.

4 Moduli of polynomials with bounded degree PC-proofs

For the rest of the paper let F be a fixed uniform family of degree at most d polynomials with variables indexed by k-ary indices, $k \geq 1$, that is indexed by indices of support - size at most $2^{k^2d+1}(k^2d)^2$, and that contains the family COMM^k.

Definition 4.1 *For $t \geq d$, $PC_t(M, F)$ is the vector space over $\mathbf{F}_p$ consisting of polynomials from $Poly(M, k, t)$ that have a PC-proof from F^M of degree at most t.*

We are going to show (in the proof of Theorem 5.1) that the property of M that $g^M \in PC_t(M, F)$, g a uniformly defined polynomial, depends only on the cardinality of M modulo some other fixed power $p^{\nu'}$, provided M is large enough.

Lemma 4.2 *Let H, H' be uniform families of vectors indexed by r - indices of support - size at most s and such that both H, H' are indexed by indices of support - size at most z.*

Then there exists $\nu \geq 1$ (depending only on $|C|$, z, s and tacitly on p) and a set $Q \subseteq \{0, \ldots, p^\nu - 1\}$ such that for every M such that $|M| \geq |C| + 2(z + s) + 2^{z+1}z^2 + z$ the following two statements are equivalent:

1. $H'^M \subseteq Span_{\mathbf{F}_p}(H^M)$.
2. $|M| \equiv r \pmod{p^\nu}$ *for some* $r \in Q$.

Proof :
For H, H' uniform families there is an uniform family of linear equations that has a solution in $\mathbf{F}_p$ for a finite M iff $H'^M \subseteq Span_{\mathbf{F}_p}(H^M)$. The system contain for any pair of indices (i', j), index i' of H' and index j of a coordinate of a vector, an equation with variables y_i indexed by indices i of H and expressing the $H'(i', j) = \sum_i y_i H(i, j)$.

By Theorem 3.5, whether or not the system has a solution over M for M of size at least $|C| + 2(z + s) + 2^{z+1}z^2 + z$ (with parameter z in Theorem 3.5 being $z := z + s$ and $s := z$) depends only on $r < p^\nu$ such that $|M| \equiv r \pmod{p^\nu}$.

q.e.d.

The following theorem is the only part of the lower bound proof for PC using the specific definition of $PC_t(M, F)$.

Theorem 4.3 *There are $\nu \geq 1$ (depending only on k and t) and uniform families H^r, $r = 0, 1, \ldots, p^\nu - 1$, of polynomials of degree at most t, with variables indexed by k-ary indices, indexed by indices of support - size $\leq 2^{k^2t+1}(k^2t)^2$, such that the following holds.*

For every finite M, if $|M| \equiv r \pmod{p^\nu}$ and $|M| \geq |C| + 6^{6^{k^2t}}$, the system $(H^r)^M$ generates as a vector space the space $PC_t(M, F)$.

Proof :
A submodule W of $Poly(M, k, t)$ is equal to $PC_t(M, F)$ iff W is the smallest subspace containing F^M and closed under the multiplication rule of PC. In particular, the later condition can be formulated as follows: if $g \in W$ and $deg(g) < t$ then $x \cdot g \in W$, for all variables x.

Due to the inclusion of COMM^k in F this reflects PC in commutative polynomial rings.

Degree $\leq t$ monomials in $Var(M, k)$ are indexed by indices of support - size $\leq k^2t$. By Theorem 3.3 we know that there is $c \geq 1$ such that any submodule of $Poly(M, k, t)$ is generated as a vector space by one of the uniform families $(H^t)^M$, for some $t \leq c$, with H^t indexed by indices of support - size $\leq 2^{k^2t+1}(k^2t)^2$.

For a pair of uniform families H', H from this finite list consider conditions:

1. $F \subseteq Span_{\mathbf{F}_p}(H)$

2. H' consists only of polynomials of degree less than t.

3. $H' \subseteq Span_{\mathbf{F}_p}(H)$

4. H consists only of polynomials of degree at most t.

5. $x_i \cdot H' := \{x_i g \mid g \in H'\} \subseteq Span_{\mathbf{F}_p}(H)$, for all variables x_i.

6. $Span_{\mathbf{F}_p}(H^M) \cap Poly(M,k,t-1) \subseteq Span_{\mathbf{F}_p}(H'^M)$

7. $Span_{\mathbf{F}_p}(H^M)$ is the smallest subspace of $Poly(M,k,t)$ among all spaces $Span_{\mathbf{F}_p}(H''^M)$, for all pairs of uniform families generating sets H''', H'' satisfying the above six conditions.

Denote the condition that $H'^M \subseteq Span(H^M)$ by $\Psi_{H',H}$. By Lemma 4.2 (with $z := 2^{k^2t+1}(k^2t)^2$ and $s := k^2t$) whether or not is $\Psi_{H',H}$ true in M depends only on M (mod p^ν), some fixed $\nu \geq 1$, provided $M \geq |C| + 2(z+s) + 2^{z+1}z^2 + z$ (hence $M \geq |C| + 6^{6^{k^2t}}$ suffices).

Let $\preceq$ be a partial quasi-ordering of uniform families H. There are finitely many such $\preceq$ as there are only finitely many uniform families. For $\preceq$ let $\Lambda_\preceq$ be the conjunction of all conditions $\Psi_{H',H}$ for $H' \preceq H$ together with all $\neg\Psi_{H',H}$ for $H' \not\preceq H$.

Let $\Theta_{H',H}$ be the disjunction of those $\Lambda_\preceq$ where the inclusions of the families according to $\preceq$ satisfy all seven conditions for H', H, and let Φ_H be the disjunction of $\Theta_{H',H}$ over all H'.

Clearly then H^M generates as a vector space $PC_t(M,k,t)$ iff M satisfies Φ_H, because $F^M \subseteq H^M \subseteq Poly(M,k,t)$ by conditions 1. and 4., H^M is closed under the multiplication rule of PC by conditions 2., 3., 5. and 6., and H^M is the smallest such space by condition 7.

For $M \geq |C| + 6^{6^{k^2t}}$, whether or not M satisfies Φ_H depends only on $r < p^\nu$ such that $|M| \equiv r$ (mod p^ν). By Theorem 3.3, for every r there is at least one H^r such that Φ_{H^r} is true in M if $|M| \equiv r$ (mod p^ν) and $|M| \geq |C| + 6^{k^2t} \geq |C| + 2^{k^2t+1}(k^2t)^2$, i.e. $(H^r)^M$ generates as a vector space $PC_t(M,F)$ for all such M.

q.e.d.

Note that we cannot expect to find one uniform family generating as a vector space $PC_t(M,F)$ for all M. Indeed, the example of a linear system F consisting of the equations $\{x_i - 1 = 0\}_{i\in M}$ together with the equation $\sum_{i\in M} x_i = 0$ demonstrates this, as solvability of F^M is equivalent to $1 \notin PC_1(F,M)$.

5 Degree lower bounds

For a uniform system F denote by $F^{(-)}$ the system without the polynomial $F_{\overline{\emptyset}}$. It is also uniform. We continue to assume that F satisfies the conditions assumed at the beginning of Section 4.

Theorem 5.1 *Let F be a uniform family as above. Then for every $t \geq d$ there is $\nu \geq 1$ such that the following holds for every $r < p^\nu$:*

If there is at least one M, $|M| \geq |C| + 6^{6^{k^2 t}}$, such that $|M| \equiv r$ (mod p^ν) and

$$F_{\overline{\emptyset}}^M \notin PC_t(M, F^{(-)})$$

then for no M, $|M| \geq |C| + 6^{6^{k^2 t}}$ and $|M| \equiv r$ (mod p^ν), is there a degree at most t PC - proof of $F_{\overline{\emptyset}}^M$ from $F^{(-),M}$.

Proof :
Let H^r be the uniform families of generating sets for $PC_t(M, F^{(-)})$ (generating it as a vector space) guaranteed by Theorem 4.3, if $|M| \geq |C| + 6^{6^{k^2 t}}$. Systems H^r are indexed by indices i of support - size $\leq 2^{k^2 t+1}(k^2 t)^2$ and let monomials correspond to indices j. Hence we may write

$$H_i^r = \sum_j H^r(i,j) x^j$$

Now take variables y_i indexed by i's and consider the system of linear equations

$$F(\overline{\emptyset}, j) \ - \ \sum_i H^r(i,j) y_i \ = \ 0$$

This system, call it K^r, is obviously also uniform. It is indexed by j's of support - size $z \leq k^2 t$, and variables are indexed by i's of support - size $s \leq 2^{k^2 t+1}(k^2 t)^2$. Moreover, for any M such that $|M| \equiv r$ (mod p^ν), $F_{\overline{\emptyset}} \in PC_t(M, F^{(-)})$ iff K_r^M has a solution in $\mathbf{F}_p$.

So Theorem 3.5 concludes the proof, provided

$$|M| \geq |C| + 6^{6^{k^2 t}} \geq |C| + 2z + 2^{s+1} s^2 + s$$

q.e.d.

The following is immediate.

Corollary 5.2 *Let F be a uniform family as above. Assume that for every $\nu \geq 1$ and every $r < p^\nu$ there is an arbitrarily large M such that $|M| \equiv r$ (mod p^ν) and such that there is a solution in $\mathbf{F}_p$ of the system $F^{(-),M}$ not satisfying $F_{\overline{\emptyset}} = 0$.*

Then $F_{\overline{\emptyset}}^M$ does not admit a degree $t \leq \frac{\log_6(\log_6(|M \setminus C|))}{k^2}$ PC - proof from $F^{(-),M}$.

This yields a lower bound for the (N,m) - systems.

Corollary 5.3 *Let $m \geq 1$ not divisible by p be fixed. Then for N not divisible by m, there is no a PC - refutation of the (N,m) - system of degree $t \leq \frac{\log_6(\log_6(N))}{m}$.*

Proof :

Take any $\nu \geq 1$ and $r < p^\nu$. As p does not divide m, there is arbitrarily large N divisible by m such that $N \equiv r \pmod{p^\nu}$. The (N,m) - system has a solution for such N, however. Hence the ideal generated by it is non-trivial, i.e., 1 is not derivable from the system, and Corollary 5.2 applies.

q.e.d.

We show now that a merely non - constant degree lower bound for PC - proofs of the (N,m) - systems follows already from Corollary 3.2 and the non - constant degree lower bound for NS - proofs of the same system proved in [4]. In fact, such a reduction applies in general.

Definition 5.4 *For $t \geq d$, $NS_t(M,F)$ is the vector space over $\mathbf{F}_p$ consisting of degree $\leq t$ polynomials g with variables from $Var(M,k)$ that have an NS-proof from F^M of degree at most t, i.e. such that there are polynomials G_i with variables from $Var(M,k)$ for which*

$$g = \sum_i G_i F_i^M$$

and $deg(G_i F_i^M) \leq t$, all i.

Theorem 5.5 *For every k,t there is t' such that for any F as in Theorem 5.1 the following holds for all M:*

If $F_{\overline{\emptyset}}^M \in PC_t(M, F^{(-)})$ then $F_{\overline{\emptyset}}^M \in NS_{t'}(M, F^{(-)})$.

In particular, if $F_{\overline{\emptyset}}^M$ (M a parameter) do not admit constant-degree NS-proofs from $(F^{(-)})^M$ they do not admit constant-degree PC-proofs either.

Proof :

Consider the $\mathbf{F}_p[\mathbf{Sym}_C(M)]$-module $Poly(M,k,t)$. It is a submodule of some $M^{(kt)}$. Take $c \geq 1$ guaranteed by Corollary 3.2 for $r := kt$.

Let W_u for $u = t, t+1, \ldots$ be the module of polynomials from $NS_u(M,F)$ that have degrees $\leq t$. As u increases W_u does not decrease. We claim that if for some t'

$$W_{t'} = W_{t'+1}$$

then $PC_t(M,F) \subseteq W_{t'}$. To see this take $f \in W_{t'}$ of degree $< t$. Then its multiple xf by a variable x is in $W_{t'+1} = W_{t'}$. Hence $W_{t'}$ is closed under the PC-rules, and as it contains F^M the claim follows.

Now note that there are at most c different submodules of $Poly(M,k,t)$ hence some $t' < t + c$ has the above property.

q.e.d.

Note that the theorem yields also non - constant PC-lower bounds for the systems encoding the pigeonhole principle (see e.g. one of [3, 21]). A linear lower bound was proved in [21] (another proof of non-constant lower bound is in [18]).

Finally we note that the same method yields $\Omega(\log(N))$ lower bounds for Nullstellensatz proofs. The difference allowing to save one log is the following. The space $PC_t(M, F)$ is a span of one of systems H^r, each indexed by indices of support - size $O(2^t)$. The space $NS_t(M, F)$ is, however, described as a span of one system H indexed by indices of support - size $O(t)$ only. Namely, for every F_i and every monomial x^j such that $deg(x^j F_i) \leq t$ the family H contains polynomial $H_{j,i} := x^j F_i$ indexed by (j, i) of support - size $O(t)$. Hence in the proof of Theorem 5.1 the parameter s is $O(t)$ and the assumption $M \geq |C| + 6^{k^2 t}$ suffices, and we have the following theorem (after Corollary 5.2).

Theorem 5.6 *Let F be a uniform family as above. Assume that for every $\nu \geq 1$ and every $r < p^\nu$ there is an arbitrarily large M such that $|M| \equiv r \pmod{p^\nu}$ and such that there is a solution in $\mathbf{F}_p$ of the system $F^{(-),M}$ not satisfying $F_{\overline{\emptyset}} = 0$.*

Then $F_{\overline{\emptyset}}^M$ does not admit a degree $t \leq \frac{\log_6(|M \setminus C|)}{k^2}$ NS - proof from $F^{(-),M}$.

6 An example with primality

The question about the length of proofs of propositionally encoded primality of a number was raised (in connection with effective interpolation) in [19]. This is an important problem as such tautologies are currently the only reasonable candidates for tautologies hard for the usual Hilbert-style systems (the so called Frege systems, or even for Extended Frege systems). The numbers are encoded in binary there rather than in unary as here. However, no lower bounds (even conditional assuming some unproven complexity-theoretic conjecture of a general nature) for any proof system were known till present for either formulation.

We define a uniform system of polynomials Π such that Π^M is solvable just if the cardinality $|M|$ is composite.

Definition 6.1 *Let M be an $L(C)$-structure, $C = \{c\}$. Variables of the system Π^M are x_i, y_j for $i, j \in M \setminus C$ and z_{ijk} for $i, j \in M \setminus C$ and $k \in M$.*

The polynomials of Π^M are

1. *$w^2 - w$ for all variables $w = x_i, y_j, z_{ijk}$.*
2. *$x_i y_j (1 - \sum_k z_{ijk})$ for all possible i, j.*
3. *$1 - \sum_{i,j} x_i y_j z_{ijk}$ for all possible k.*

4. $x_{i_1}x_{i_2}y_{j_1}y_{j_2}z_{i_1j_1k}z_{i_2j_2k}$ *for all possible* $i, i_1, i_2, j, j_1, j_2, k$ *if* $i_1 \neq i_2$ *or* $j_1 \neq j_2$.

5. $x_iy_jz_{ijk_1}z_{ijk_2}$ *for all possible* i, j, k_1, k_2 *if* $k_1 \neq k_2$.

If $x_i := a_i$, $y_j := b_j$ and $z_{ijk} := c_{ijk}$ is a solution of Π^M then the set $\{((i,j),k) \mid c_{ijk} = 1\}$ defines a bijection between $A \times B$ and M, where $A := \{i \in M \setminus C \mid a_i = 1\}$, $B := \{j \in M \setminus C \mid b_j = 1\}$. As $|A|, |B| < |M|$, such a solution exists just if $|M|$ is a composite number.

Theorem 6.2 *For any prime* N *there is no PC - refutation (NS - refutation, respectively) of the system* Π^M, $|M| = N$, *of degree* $t \leq \frac{\log_6(\log_6(N-1))}{3}$ *(of degree* $t \leq \frac{\log_6(N-1)}{3}$, *respectively).*

Proof :

Take M of prime cardinality N and any p^ν and $r < p^\nu$. Take $N' := np^\nu + r$ for some suitable n large enough such that N' is composite. Then $\Pi^{N'}$ is solvable, i.e. non-refutable, while $N \equiv N' \pmod{p^\nu}$. Hence Corollary 5.2 applies.

q.e.d.

Acknowledgements: This work was done while I was visiting the Mathematical Institute of the University of Oxford.

I thank D. Macpherson for pointing out references [10, 11] and to D. Evans for discussing with me that work. I am indebted to P. Pudlák, J. Sgall and A. Woods for a feedback during a lecture series where I presented this work.

References

[1] M. Ajtai, The independence of the modulo p counting principles, *in*: "Proceedings of the 26th Annual ACM Symposium on Theory of Computing", 402-411. ACM Press, 1994.

[2] M. Ajtai, Symmetric systems of linear equations modulo p, a preprint, 1994. Available as a report number TR94-015 of the *Electronic Colloquium on Computational Complexity*, http://www.eccc.uni-trier.de/eccc

[3] P. Beame, S. A. Cook, J. Edmonds, R. Impagliazzo, and T. Pitassi, The relative complexity of NP search problems. *in*: "Proceedings of the 27th ACM Symposium on Theory of Computing", 303-314. ACM Pres, 1995.

[4] P. Beame, R. Impagliazzo, J. Krajíček, T. Pitassi, and P. Pudlák, Lower bounds on Hilbert's Nullstellensatz and propositional proofs, *Proceedings of the London Mathematical Society* **(3) 73** (1996), 1-26.

[5] D. Brownawell, Bounds for the degrees in the Nullstellensatz, *Annals of Mathematics* (Second Series) **126** (1987), 577-591.

[6] S. Buss, R. Impagliazzo, J. Krajíček, P. Pudlák, J. Sgall, and A. A. Razborov, Proof complexity in algebraic systems and bounded depth Frege systems with modular counting, *Computational Complexity* **6** (1996/1997), 256-298.

[7] S. Buss, D. Grigoriev, R. Impagliazzo, and T. Pitassi, Linear gaps between degrees for the polynomial calculus modulo distinct primes, preprint, 1998.

[8] M. Clegg, J. Edmonds, and R. Impagliazzo, Using the Groebner basis algorithm to find proofs of unsatisfiability, *in*: "Proceedings of the 28th ACM Symposium on Theory of Computing", 174-183. ACM Press, 1996.

[9] A. S. Cook, and A. R. Reckhow, The relative efficiency of propositional proof systems, *J. of Symbolic Logic* **44(1)** (1979), 36-50.

[10] D. G. D. Gray, The submodule structure of some permutation modules, *PhD. Thesis*, University of East Anglia, 1997.

[11] D. G. D. Gray, The structure of some permutation modules for the symmetric group of infinite degree, *J. of Algebra* **193** (1997), 122-143.

[12] G. D. James, The module orthogonal to the Specht module, *J. of Algebra* **46** (1997), 451-456.

[13] G. D. James, (1978) "The representation theory of the symmetric groups", LN in Mathematics, Vol. **682**, Springer-Verlag, 1978.

[14] J. Kollár, Sharp effective Nullstellensatz, *J. Amer. Math. Soc.* **1(4)** (1988), 963-975.

[15] J. Krajíček, "Bounded arithmetic, propositional logic, and complexity theory", Encyclopedia of Mathematics and Its Applications, Vol. **60**, Cambridge University Press, 1995.

[16] J. Krajíček, Lower bounds for a proof system with an exponential speed-up over constant-depth Frege systems and over polynomial calculus, in: Eds. I.Prívara, P. Růžička, 22nd Inter. Symp. *Mathematical Foundations of Computer Science* (Bratislava, August '97), Lecture Notes in Computer Science, Vol. **1295**, 85-90. Springer-Verlag, 1997.

[17] J. Krajíček, A fundamental problem of mathematical logic, *Annals of the Kurt Gödel Society*, Springer-Verlag, Collegium Logicum, Vol.**2** (1996), 56-64.

[18] J. Krajíček, Uniform families of polynomial equations over a finite field and structures admitting an Euler characteristic of definable sets, preprint, 1997.

[19] J. Krajíček, and P. Pudlák, Some consequences of cryptographical conjectures for S_2^1 and EF, *Information and Computation* **140 (1)** (1998), 82-94.

[20] S. Lang, "Algebra" (3rd ed.), Addison-Wesley Publ. Co, 1993.

[21] A. A. Razborov, Lower bounds for the polynomial calculus, *Computational Complexity*, to appear (preprint 1998).

[22] S. Riis, and M. Sitharam, Non-constant degree lower bounds imply linear degree lower bounds, a preprint, 1997. Available as a report number TR97-048 of the *Electronic Colloquium on Computational Complexity*, http://www.eccc.uni-trier.de/eccc