LOWER BOUNDS ON HILBERT’S NULLSTELLENSATZ AND PROPOSITIONAL PROOFS

PAUL BEAME, RUSSELL IMPAGLIAZZO, JAN KRAJÍČEK, TONIANN PITASSI, and PAVEL PUDLÁK

[Received 25 November 1994—Revised 18 July 1995]

ABSTRACT

The so-called weak form of Hilbert’s Nullstellensatz says that a system of algebraic equations over a field, \( Q_i(\bar{x}) = 0 \), does not have a solution in the algebraic closure if and only if 1 is in the ideal generated by the polynomials \( Q_i(\bar{x}) \). We shall prove a lower bound on the degrees of polynomials \( P_\ell(\bar{x}) \) such that \( \sum P_\ell(\bar{x})Q_i(\bar{x}) = 1 \).

This result has the following application. The modular counting principle states that no finite set whose cardinality is not divisible by \( q \) can be partitioned into \( q \)-element classes. For each fixed cardinality \( N \), this principle can be expressed as a propositional formula \( \text{Count}^N_{x_1,...,x_r} \) with underlying variables \( x_\ell \), where \( r \) ranges over \( q \)-element subsets of \( N \). Ajtai [4] proved recently that, whenever \( p,q \) are two different primes, the propositional formulas \( \text{Count}^N_{x_1,...,x_r} \) do not have polynomial size, constant-depth Frege proofs from substitution instances of \( \text{Count}^m_q \), where \( m \neq 0 \pmod{p} \). We give a new proof of this result based on the lower bound for Hilbert’s Nullstellensatz. Furthermore our technique enables us to extend the independence results for counting principles to composite numbers \( p \) and \( q \). This improved lower bound together with new upper bounds yield an exact characterization of when \( \text{Count}_q \) can be proved efficiently from \( \text{Count}_p \), for all values of \( p \) and \( q \).

Introduction

The problem of solvability of a system of algebraic equations

\[
Q_i(\bar{x}) = 0, \quad \text{for } i \in I,
\]

in a fixed finite field \( F \) is one of the most important \( \mathcal{NP} \)-complete problems. If we look for solutions in the algebraic closure of \( F \), or if we just know \textit{a priori} that the solutions must be in \( F \), the solvability is characterized by the basic result in algebraic geometry known as Hilbert’s Nullstellensatz. Namely, the equations (1) do not have a solution in the algebraic closure of \( F \), if and only if there exist polynomials \( P_\ell(\bar{x}) \) from \( F[\bar{x}] \) such that

\[
\sum P_\ell(\bar{x})Q_i(\bar{x}) = 1.
\]

The solvability problem of (1) in \( F \) is known to be \( \mathcal{NP} \)-complete, even if the degrees of the \( Q_i \) are bounded by a constant. This is equivalent to the solvability of \( \{Q_i(\bar{x}) = 0\} \cup \{x_i^p - x_i = 0\} \) in an algebraic closure of \( F \). Now suppose there existed polynomials \( P_\ell \) of constant degree satisfying (2) for the extended system whenever (1) did not have a solution. Then, by solving linear equations which determine the coefficients of the monomials in the \( P_\ell \), we could construct these polynomials in polynomial time. Thus \( \mathcal{P} \neq \mathcal{NP} \) implies that there are instances

The work of P. Beame was supported by NSF grant CCR-9303017, that of R. Impagliazzo, J. Krajíček and P. Pudlák was partially supported by grant #93025 of the joint US-Czechoslovak Science and Technology Program, and that of T. Pitassi was supported by an NSF postdoctoral fellowship and UC presidential fellowship.

1991 Mathematics Subject Classification: primary 03F20, 03B05, 03F30; secondary 68Q25.

such that the degree of the $P_i$ cannot be bounded by a constant. This raises a very natural problem of proving such a lower bound. Furthermore, the assumption $P \neq NP$ does not point to any concrete system of equations which requires polynomials $P_i$ of non-constant degree.

The question of the degree of the polynomials $P_i$ in (2) has been studied in the context of the 'effective Nullstellensatz' of Brownawell and others [8, 13, 9] and upper bounds on the degrees of the $P_i$ are shown that are exponential in the number of variables. In the general case where the question of interest is the solution of a family of polynomials in an algebraic closure of $F$ there is a matching lower bound example as well [8]. However, when using the Nullstellensatz to determine if solutions exist in $F$ by extending the system of polynomials as described above, the degree of the $P_i$ for this example becomes constant.

In this paper we prove a non-constant lower bound on the degree of polynomials $P_i$ in (2) for such an extended system with particular polynomials $Q_i$. We prove a non-constant lower bound for equations which represent the counting principles mentioned in the abstract. For applications we need to consider counting modulo a fixed number, which can be composite (thus we prove it also for some rings which are not fields). There is a much simpler proof for the case of the real field which gives a much larger, linear lower bound, while for counting modulo a number, the bound is an extremely slowly growing function. This is due to the fact that we use Ramsey's theorem repeatedly. Recently Jeff Edmonds proved an $\Omega(\sqrt{n})$ lower bound for a system of equations for a related principle PHP (see below). However his result does not have an application in propositional calculus so far.

Our main motivation for the lower bound comes from propositional calculus where we would like to prove lower bounds on the size of proofs in some natural proof systems. We can think of the polynomials $P_i$ in the equation (2) as a proof of unsolvability of the system (1). Thus, if the underlying field is GF$_2$, then we actually have a propositional proof system based on Hilbert's Nullstellensatz and our lower bound can be interpreted as a lower bound for this proof system.

The interest in lower bounds on the complexity of propositional proof systems stems from two sources. First, any complete and sound proof system can be thought of as a non-deterministic acceptor of the set of propositional tautologies. In fact, Cook and Reckhow [10] define propositional proof systems to be just non-deterministic acceptors of the set of propositional tautologies. Hence the problem of whether there is a proof system in which all tautologies have proofs of size polynomial in the size of the formula is equivalent to the well-known question in computational complexity theory of whether the class of sets accepted by non-deterministic Turing machines is closed under the complementation. A research program of proving lower bounds for particular proof systems somehow parallels the approach taken by boolean complexity of trying to prove lower bounds on the circuit size for restricted classes of circuits.

The second source of interest in lower bounds on the size of propositional proofs are relations between subsystems of first-order arithmetic called systems of bounded arithmetic, and propositional proof systems. These theories formalize effective reasoning about finite structures, meaning that predicates, relations, functions and algorithms used in arguments have bounded computational complexity (most often $\mathcal{AC}^0$, $\mathcal{NC}^1$, $\mathcal{P}$, $\mathcal{NP}$, ..., depending on the system). Each first-order system is associated with a particular propositional proof system with
the property that whenever a finite combinatorial principal is provable in the
theory then the sequence of tautologies formalizing the principle has polynomial
size proofs in the associated proof system. Moreover, these systems are typically
natural systems which have been considered prior to bounded arithmetic, like
Extended Resolution, Hilbert-style calculi (called Frege systems), or quantified
propositional calculus.

It follows that to demonstrate the unprovability of a principle in a theory (and
hence, in a sense, its computational difficulty) it is sufficient to prove that it has
no polynomial size proofs in the particular proof system.

The theories of interest here are the bounded arithmetic system \( I\Delta_0(R) \) and its
extensions by \( \Pi^0_2 \) axioms; cf. [17]. The corresponding propositional proof systems
are constant-depth Frege systems.

The first strong lower bound for a natural proof system was obtained by Haken
[11] who proved an exponential lower bound to the resolution proofs of the
pigeonhole principle (PHP). Then Ajtai [1] showed that constant-depth Frege
systems (which are stronger than resolution) do not admit polynomial size proofs
of PHP. The first exponential lower bound for such systems was obtained in [14]
(for different formulas), followed by an exponential lower bound for PHP
established in [6, 16, 18]. Constant-depth Frege systems are the strongest proof
systems for which a non-trivial lower bound is known. The problem of proving
superpolynomial lower bounds for a general Frege system appears to be
analogous to important open problems in complexity theory (for example,
whether \( \text{NP} \neq \text{NEXP} \)).

In [2] Ajtai showed that there are no polynomial size constant-depth Frege
proofs of \( \text{Count}_2 \) (the propositional mod 2 counting principle) from the pigeon-
hole principle. (We use the notation \( \text{Count}_q \) for the generic version of \( \text{Count}_m \)
where \( m \not\equiv 0 \pmod{q} \).) This was subsequently improved to an exponential lower
bound in [7, 19]. In [4] Ajtai studied the relation of principles \( \text{Count}_p \) and \( \text{Count}_q \),
which had been considered earlier in [17]. For example, assume that \( q \mid p \) and that
\( R \) is a \( q \)-partition of \( N \). We may expand each point of \( N \) into \( p/q \) copies creating
thus a \( p \)-partition of \( N' = pN/q \). Moreover, if \( q \) does not divide \( N \) then \( p \) does not
divide \( N' \). This demonstrates that \( \text{Count}_q \) follows from a simple instance of
\( \text{Count}_p \) if \( q \mid p \); more precisely, there are polynomial size, constant-depth Frege
proofs of \( \text{Count}_q \) from \( \text{Count}_p \).

Ajtai [3, 4] showed that if \( p, q \) are different primes then \( \text{Count}_q \) does not admit
polynomial size, constant-depth Frege proofs from instances of \( \text{Count}_p \). This
implies by the above that whenever \( p \) is a prime and \( q \) has a prime divisor
different from \( p \) then \( \text{Count}_q \) cannot be reduced to \( \text{Count}_p \), that is, proved by
polynomial size, constant-depth Frege proofs from its instances. (In [4] he states
that his methods can be extended to mutually prime \( \text{squarefree} \) numbers; in fact,
using the relations between counting principles proved below one can reduce the
general problem to this special case.) The same problem was studied in [19] and
reduced there to a purely combinatorial question.†

of the problem of the independence of the counting principles which was based on a partial solution of
the combinatorial problem from [19]. However, the manuscript contained serious gaps as well as did
some later versions available in Summer and Fall of 1993. Only recently (Summer 1994) a substantial
revision occurred, while our manuscript was already available in Spring 1994.
In this paper we give a new proof of this lower bound which also applies when $p$ and $q$ are not primes. We need only that there is a prime that divides $q$ but not $p$. The proof consists of two parts. The first part is similar to the proof strategy of [6, 16, 18] and it is a universal method allowing us to reduce the lower bound problem to a combinatorial question about the existence of certain finite structures. This part employs ideas from boolean complexity (partial truth assignments, switching lemmas). The second part is an application of the lower bound for Hilbert's Nullstellensatz.

We see the main contribution of the paper in the new method. In particular, by proving a better lower bound for the Nullstellensatz our method would yield a better lower bound on proofs while the method of [4] (its boolean complexity part) inherently can give only the non-existence of polynomial upper bounds (see the last section for more discussion on this topic). Furthermore, our technique does not need a priori a reduction to the case of two primes.

The paper is organized as follows. In the first section we define the constant-depth Frege systems and the propositional formulas $\text{Count}_q^a$ with which we are concerned and we formulate the main theorem. In the second section we give a number of upper bound constructions relating the provability of different modular counting axioms.

The third section is devoted to the first part of the lower bound proof, reducing the lower bound to a combinatorial problem. Relevant notions from [6, 16, 18] are recalled there and the main theorem is proved from a combinatorial lemma.

In the fourth section we reduce the combinatorial lemma to the main lemma stating a lower bound on the degrees of certain polynomials.

The bound on the degree of polynomials in the Nullstellensatz is proved in the fifth section.

In the sixth section we define the extensions of constant-depth Frege systems by modular counting gates and we conclude with several remarks on the relation of our method to those of [4, 19].

We refer the reader to [15] for information about bounded arithmetic, propositional proof systems and their relations.

1. Proof systems and counting principles

We shall confine ourselves to the following propositional language: atoms $x, y, \ldots$, constants 0 (falsity) and 1 (truth), negation $\neg$ and disjunction $\lor$ (binary). We shall use $\land$ as an abbreviation. The depth of a formula is the maximal number of alternations of $\neg$ and $\lor$ and its size is the number of occurrences of $\lor$. We shall use symbol $\lor_\phi$, denoting the disjunction of unbounded arity as an abbreviation for the disjunction formed from binary $\lor$ with brackets distributed arbitrarily.

A Frege system is a sound and implicationally complete proof system having a finite number of axiom schemes and inference rules. A typical Frege system is the usual calculus based on a finite number of axioms with modus ponens as the only rule of inference. See [10] for details. The size of a proof in a Frege system is the number of distinct subformulas appearing in the proof where we do not distinguish $F$ and $\neg F$. We will also need a notion of the size of the inference rules and axiom schemes in a Frege system. For this we use the same notion of the
number of distinct subformulas appearing in the axiom scheme or inference rule and we again identify $F$ and $\neg F$.

A depth $d$ Frege system is a Frege system allowing only formulas of depth at most $d$ in proofs. It is not complete but there is a constant $c$ such that any depth $d$ tautology has a depth $c + d$ proof in the Frege system.

Now we shall formally define the counting principles discussed in the introduction.

**Definition 1.1.** Let $N \geq r \geq 2$ and let $V$ be a set of cardinality $N$. We use $[V]'$ to denote the set of $r$-element subsets of $V$.

Formula $\text{Count}_N^r$ is formed from atoms $x_e$, where $e \in [V]'$, and it is the formula

$$\bigvee_{e \in V} \bigwedge_{e \in \partial e} \neg x_e \bigvee_{e \in f} (x_e \land x_f)$$

where $e \perp f$ abbreviates the conjunction $e \cap f \neq \emptyset \land e \neq f$.

Denote by $\text{Count}_{r,i}$ the set of formulas $\text{Count}_N^r$ for $N = i \pmod{r}$. For $R$ a set of pairs $(r, i)$ such that $0 < i < r$, $\text{Count}_R$ denotes the union of the sets of formulas $\text{Count}_{r,i}$ such that $(r, i) \in R$. Finally, $\text{Count}$, denotes the set of all formulas $\text{Count}_N^r$ for $N \equiv 0 \pmod{r}$; equivalently, $\text{Count}_r$ is $\text{Count}_R$ for $R = \{(r, i) \mid 0 < i < r\}$.

We want to use the counting principles rather as axiom schemas. Thus we define that an instance of $\text{Count}_N^r$ is obtained by substituting some formulas $\psi_e$ for the variables $x_e$; we shall denote it by $\text{Count}_N^r(\Psi)$, where $\Psi = \{\psi_e\}_{e \in [V]'}$.

In the introduction we observed that if $r$ divides $s$ then there are polynomial size, constant-depth Frege proofs of $\text{Count}$, from instances of $\text{Count}_r$. (Let us note that it is implicit in the restriction that we use only instances that are of bounded depth).

**Problem.** Assume $s, r \geq 2$ and assume $r$ does not divide $s$. Are there polynomial size, constant-depth Frege proofs of $\text{Count}_r$, from instances of formulas in $\text{Count}_r$? If so, under what circumstances do they exist?

This was solved in the negative by Ajtai [3, 4] for the case when $r$, $s$ are two different primes. We give a new proof of a strengthened version of his theorem that gives a complete characterization of this problem.

Let the symbol $(a, b)$ denote the greatest common divisor of $a$ and $b$.

**Theorem 1.2** (main). Let $q \geq 2$ and $0 < i < q$. Let $R$ be a subset of the set of pairs of integers $(p, j)$ such that $0 < j < p$. Then there are constant-depth, polynomial size Frege proofs of formulas of $\text{Count}_q^r$, from instances of $\text{Count}_p$, if and only if there is a $(p, j) \in R$ such that all prime divisors of $pl(p, j)$ also divide $q/(q, i)$.

**Corollary 1.3.** Let $p, q \geq 2$ and assume that there is a prime factor of $q$ which does not divide $p$. Then there are no constant-depth, polynomial size Frege proofs of $\text{Count}_q^r$, for $N \equiv 0 \pmod{q}$, from instances of $\text{Count}_p$. In particular, this holds for all $N$ such that $(p, q/(q, N)) = 1$ and $N \not\equiv 0 \pmod{q}$.

On the other hand, if all prime factors of $q$ also divide $p$ then there are
constant-depth, polynomial size Frege proofs of Count^N, for N ≠ 0 (mod q), from instances of Count_p.

Proof of Corollary 1.3 from Theorem 1.2. For the first part, if r is a prime factor of q which does not divide p then for N=q/r (mod q) we have (p, q/(q, N)) = (p, r) = 1. It follows that for all j, with 0 < j < p, (p/(p, j), q/(q, N)) = 1. Thus for each j, with 0 < j < p, there is some prime factor of p/(p, j) that is not a factor of q/(q, N). Applying Theorem 1.2 we obtain our desired result.

For the second part, assume that all prime factors of q also divide p. For N ≠ 0 (mod q), let s be any prime factor of q/(q, N). By assumption s also divides p. Thus for j = p/s all prime factors of p/(p, j) also divide q/(q, N) and applying Theorem 1.2 we have finished.

2. Upper bounds for proofs of counting principles

The bulk of our arguments are concerned with lower bounds but we deal with the upper bounds first. We begin by observing some straightforward relationships between the principles.

Lemma 2.1. Assume that r ≥ 2 and k is a positive integer. Then for all i ≠ 0 (mod r), with 0 < i < rk, there are polynomial size constant-depth Frege proofs of Count^r^k from instances of Count^r/modr.

Proof. Given an rk-partition (a partition into blocs of size rk) of a set of size N, we can obtain an r-partition of N by splitting each class of size rk into k classes each of size r. If N ≡ i (mod rk), with 0 < i < rk, then there is no such partition; however the construction is local, so we can transform any alleged definition of such a partition into an alleged definition of an r-partition that violates Count^r/modr. The formalization of this argument as a constant-depth Frege proof is easy. It is also possible first to formalize the proof in bounded arithmetic IΔ0(R) and then to refer to a well-known translation onto bounded depth Frege proofs [17, 15].

The following lemma is a rephrasing and extension of the observation from the introduction.

Lemma 2.2. Assume that r ≥ 2, 0 < i < r, and k is a positive integer. Then
(a) there are polynomial size constant-depth Frege proofs of Count^r^k,i from instances of Count^r^k,ik,
(b) if ik ≠ 0 (mod r), there are polynomial size constant-depth Frege proofs of Count^r,i from instances of Count^r,ik mod r.

Proof. Suppose that we have an r-partition of N, N ≡ i (mod r). We can make k copies of each point to create a new set of size N' = Nk. Part (a) follows by creating an rk-partition of N' where each new class contains all k copies of the elements of each class from the partition of N. Part (b) follows instead by creating an r-partition of N' by making each class in the partition of N into k classes in the
new partition. (Or alternatively, apply Lemma 2.1 after applying Part (a).) Again the formalizations are easy.

The following lemma involves a more interesting construction. Some of the ideas used are motivated by similar constructions in [19].

**Lemma 2.3.** Let \( r \geq 2 \) and \( k \) be a positive integer. Then there are polynomial size constant-depth Frege proofs of \( \text{Count}_{r,k} \) from instances of \( \text{Count}_{r,1} \).

**Proof.** Suppose that there is an \( rk \)-partition of \( N = k \) (mod \( rk \)). Then by adding a constant size (at most \( rk \)) set of \( rk \)-classes we obtain an \( rk \)-partition \( E \) of some \( N' > N \) such that \( N' = k \) (mod \( r \cdot k \)). We will define an \( r \)-partition of \( M = \lfloor N' \rfloor k \).

Since \( N' = k \) (mod \( r \cdot k \)), \( |M| = \binom{N'}{k} = 1 \) (mod \( r \)) and thus this \( r \)-partition will violate \( \text{Count}_{r,1} \).

First we define an equivalence on \( M \) with larger blocks. For \( X, Y \in M \) let \( X \sim Y \), if for every block \( Z \) of \( E \), \( X \cap Z = Y \cap Z \). We shall prove that the size of each block is divisible by \( r \). This follows from the following result.

*For every \( a_1, \ldots, a_t \geq 1 \), if \( a_1 + \ldots + a_t = k \), then \( r \) divides \( \binom{kr}{a_1} \ldots \binom{kr}{a_t} \).*

To prove this let \( p \) be a prime, let \( p' \) be the largest power of \( p \) which divides \( r \). Let \( s \geq 0 \) be maximal such that each \( a_i \) is divisible by \( p^s \). Thus \( p' \) divides all \( a_i \), and hence also their sum \( k \). Let \( 1 \leq i \leq l \) be such that \( a_i \) is not divisible by \( p^{s+1} \). Let us write

\[
\binom{kr}{a_i} = kr \binom{kr - 1}{a_i - 1} / a_i .
\]

Since \( p' \) divides \( k \) and no larger power divides \( a_i \), \( \binom{kr}{a_i} \) is divisible by \( p' \). This proves the required result.

Since the size of blocks of \( \sim \) is divisible by \( r \) and is bounded by a constant, we can define a refinement with blocks of size exactly \( r \) by a constant depth formula and the whole argument can be presented as a polynomial size, constant-depth Frege proof.

**Corollary 2.4.** Let \( p \geq 2 \) and \( 0 < j < p \). Then

(a) there are polynomial size constant-depth Frege proofs of \( \text{Count}_{p,j} \) from instances of \( \text{Count}_{p,(p,j),1} \) \( \) and

(b) there are polynomial size constant-depth Frege proofs of \( \text{Count}_{p,(p,j),1} \) from instances of \( \text{Count}_{p,j} \).

**Proof.** For Part (a), we start with instances of \( \text{Count}_{p,(p,j),1} \). By Lemma 2.3 we obtain \( \text{Count}_{p,(p,j)} \). By definition there are integers \( k, l \) such that \( (p, j) = kj + lp \). Applying Lemma 2.2(b) with this value of \( k \), \( r = p \) and \( i = j \), we obtain \( \text{Count}_{p,j} \).

For Part (b), starting with instances of \( \text{Count}_{p,j} \) and applying Lemma 2.2(a)
with \( r = p/(p, j), i = j, \) and \( k = (p, j), \) we obtain \( \text{Count}_{p/(p, j), (p, j)} \). Then, applying Lemma 2.2(b) with \( r = p/(p, j), i = 1, \) and \( k = j/(p, j) \) we obtain \( \text{Count}_{p/(p, j), 1} \).

**Lemma 2.5.** Let \( p, q \geq 2 \). If all prime factors of \( p \) also divide \( q \) then there are polynomial size constant-depth Frege proofs of \( \text{Count}_{q, 1} \) from instances of \( \text{Count}_{p, 1} \).

**Proof.** If all prime factors of \( p \) also divide \( q \) then there is some integer \( k \) such that \( p \mid q^k \). By Lemma 2.1, there are polynomial size constant-depth Frege proofs of \( \text{Count}_{q, 1} \) from instances of \( \text{Count}_{p, 1} \). We will show how to obtain \( \text{Count}_{q, 1} \) from instances of \( \text{Count}_{q^k, 1} \).

Suppose that there is some \( q \)-partition \( E \) of \( N = 1 \pmod{q} \). The Fermat–Euler Theorem implies that

\[
N^m \equiv 1 \pmod{q^k}
\]

where \( m \) is any multiple of \( \phi(q^k) \), the number of residues modulo \( q^k \) relatively prime to \( q^k \). Fix some such multiple \( m \) with \( m \equiv k \). Define a \( q^m \)-partition \( E' \) of \( N^m \) by taking the \( m \)th Cartesian power of \( E \), that is, the classes of \( E' \) have the form

\[
e_1 \times \ldots \times e_m
\]

where \( e_i \) are classes of \( E \). Now \( E' \) is a \( q^m \)-partition of \( N^m = N \times \ldots \times N \). Decompose each class of \( E' \) into subclasses of size \( q^k \). This yields a partition violating \( \text{Count}_{q^k, 1} \).

The formalization of this argument as a polynomial size constant-depth Frege proof is quite straightforward.

3. Reducing the lower bound to a combinatorial problem

In this section we reduce the lower bound that we are after to a purely combinatorial problem concerning generic systems. This section is a modification of very similar arguments that can be found in [7] and [19] (the name generic systems was introduced in [19] for similar objects). For the rest of the paper we shall fix two different numbers \( p, q \geq 2 \) and a set \( V \) of cardinality \( N \) such that \( N \equiv 0 \pmod{q} \). We will also sometimes find it convenient to identify an integer \( N \) with a canonical set of size \( N \). We will first state some important definitions.

**Definition 3.1.** A \( q \)-decision tree \( T \) over \( V \) is a finite directed tree whose vertices other than leaves are labelled by elements \( v \in V \), whose edges are labelled by classes \( e \in [V]^q \), whose leaves are labelled from a fixed set \( L \) of values, and which satisfies the condition inductively defined by the following:

1. if the label of the root of \( T \) is \( v \) then for any \( e \in [V]^q, v \in e \), there is exactly one edge outgoing from the root and labelled by \( e \), and there are no other edges at the root;
2. if \( T^e \) is defined to be the proper subtree of \( T \) whose root is the end-point of an edge outgoing from the root and labelled by \( e \), then \( T^e \) is a \( q \)-decision tree over \( V \setminus \{e\} \).

To get an intuition about \( q \)-decision trees, think of them as describing possible
plays in a two player game. One player pretends that there is a partition of \( V \) into blocks of size \( q \) and the other asks questions 'in which block \( e \) is the vertex \( v \)?'. Thus, in particular, the edge labels on a branch form a partial partition.

The \textit{height} of tree \( T \) is the maximum number of edges on a path from the root to a leaf.

To each branch of \( T \) we assign the partial partition consisting of its edge labels and denote by \( \text{br}(T) \), and \( \text{br}_i(T) \), the sets of the partial partitions assigned to the branches of \( T \), and to the branches with the leaf label \( i \), respectively.

We would like to prove the main theorem by finding an assignment making all formulas from \( \text{Count}_R \) true but \( \text{Count}_q^N \) false, demonstrating thus that \( \text{Count}_q \) cannot be proved from \( \text{Count}_R \). This is, of course, impossible as all formulas in \( \text{Count}_q \) are tautologies. So we must find another way of 'evaluating' formulas which would permit such an argument. A possible evaluation is to assign to each formula \( \phi \) a usual decision tree deciding the value of the formula and consider the set of all branches of the tree for \( \phi \) with leaf label 1. Thus \( \phi \) is a tautology if and only if all branches in this decision tree have leaf label 1 (\( \text{br}(T_\phi) = \text{br}_1(T_\phi) \) in future terminology).

We shall approximate this idea by the following definition. Here to each formula \( \phi \) we assign a \( q \)-decision tree \( T_\phi \). Intuitively, a formula \( \phi \) is approximately true if and only if all branches in \( T_\phi \) have leaf label 1. Furthermore, Condition (1) of Definition 3.1 corresponds to the first conjunct in \( \neg \text{Count}_q^N \) (that is, every \( v \in V \) is in some class \( e \) of the partition violating \( \text{Count}_q^N \)) while Condition (2) corresponds to the second one (that any two classes in the partition are disjoint). Hence we expect that all leaves of \( T_\phi \) have leaf label 1 for \( \phi = \neg \text{Count}_q^N \) and thus get a notion of evaluation in which \( \text{Count}_q^N \) is false. The following is a modification of a similar definition in [6, 16].

\textbf{Definition 3.2.} Let \( \Gamma \) be a set of formulas formed from atoms of \( \text{Count}_q^N \) (we shall not repeat this condition as we do not consider formulas formed from other atoms) and closed under subformulas. A \textit{k-evaluation} of \( \Gamma \) is a mapping which assigns a tree \( T_\phi \) to every formula \( \phi \in \Gamma \) such that

1. the set of leaf labels \( L \) of \( T_\phi \) is \( \{0, 1\} \);
2. \( T_\phi \) is a \( q \)-decision tree over \( V \) of height at most \( k \);
3. \( T_0 \) and \( T_1 \) are \( q \)-decision trees of height 0, \( \text{br}_1(T_0) = \emptyset \) and \( \text{br}_1(T_1) = \text{br}(T_1) \);
4. \( T_x \) is a \( q \)-decision tree having the property that every non-leaf vertex of \( T_x \) is labelled by some \( v \in e \), for each branch of \( T_x \), the edge labels cover \( e \), and \( \text{br}_1(T_{x'} \phi) \) is the unique branch of length 1 consisting of the edge labelled by \( e \);
5. \( \text{br}(T_{\neg \phi}) = \text{br}(T_\phi) \) and \( \text{br}_1(T_{\neg \phi}) = \text{br}(T_\phi) \setminus \text{br}_1(T_\phi) \);
6. if \( \phi = \bigvee_i \phi_i \) then \( T_\phi \) refines and represents \( \bigvee H \) with \( H := \bigcup_i \text{br}_1(T_{\phi_i}) \), where
   a. \( T_\phi \) refines \( \bigvee H \) means:
   \[
   \forall E \in \text{br}(T_\phi), \ (\exists F \in H, F \subseteq E) \lor (\forall F \in H, E \perp F)
   \]
   where \( E \perp F \) if and only if \( e \perp f \) for some \( e \in E \) and \( f \in F \),
(b) and $T_\phi$ represents $\vee H$ means:
\[
\text{br}_1(T_\phi) = \{ E \in \text{br}(T_\phi) \mid \exists F \in H, F \subseteq E \}.
\]

The following easy lemma is completely analogous to lemmas F3 and F4 from [6, 16] treating the case of PHP, in place of Count$^N_q$ and we shall not re-prove it here.

**Lemma 3.3.** Assume $\Gamma$ is a set of formulas closed under subformulas. Let $T$ be its $k$-evaluation and assume $kqs < N$. Then

1. if $\phi \in \Gamma$ is an axiom scheme of size at most $s$ then $\text{br}_1(T_\phi) = \text{br}(T_\phi)$;
2. equality $\text{br}_1(T_\alpha) = \text{br}(T_\alpha)$ is preserved by any sound inference rule of size at most $s$; for example, if $3kq < N$, then $\text{br}_1(T_\alpha) = \text{br}(T_\alpha)$ and $\text{br}_1(T_{\alpha \lor \beta}) = \text{br}(T_{\alpha \lor \beta})$ implies $\text{br}_1(T_\beta) = \text{br}(T_\beta)$;
3. if $\phi = \text{Count}_q^N \in \Gamma$ then $\text{br}_1(T_\phi) = \emptyset$ (≠ $\text{br}(T_\phi)$ in particular).

Assume now that $\Pi$ is a short constant-depth Frege proof of $\text{Count}_q^N$ from some instances of $\text{Count}_q^R$; that is from some formulas
\[
\text{Count}_p^M(y_e/\psi_e),
\]
where $M = j \pmod{p}$, $0 < j < p$ and $(p, j) \in R$, $g \in [M]^p$ and $\psi_e$ are formulas in atoms $x_e$ of $\text{Count}_q^N$. There is a constant $s$ such that all axiom schemes and inference rules used in $\Pi$ other than the $\text{Count}_q^R$ axioms are of size at most $s$. Suppose that there is a $k$-evaluation $T$ of the set of all subformulas in $\Pi$ such that

1. $kqs < N$ and
2. for all instances $\phi$ of a $\text{Count}_q^R$ axiom in $\Pi$, $\text{br}_1(T_\phi) = \text{br}(T_\phi)$.

That would give a contradiction with Lemma 3.3 as $\text{br}_1(T_\phi) = \text{br}(T_\phi)$ would hold for all axioms and all inferences in $\pi$ but not for the final formula.

This motivates the structure of our argument. More precisely, we first show that if $\Pi$ is a short proof of $\text{Count}_q^N$ then there is a $k$-evaluation of all subformulas of $\Pi$ satisfying (1) but not (2), and thus derive in Theorem 3.5 that a particular combinatorial object (a generic system) must exist. We then derive a contradiction in Lemma 3.10 by showing that this combinatorial object cannot exist.

Theorem 3.5 is proved by a more-or-less straightforward combination of the arguments in [16, 7, 19, 5]. For the benefit of the reader we outline its proof. The proof of the combinatorial statement (Lemma 3.10) occupies the rest of the body of the paper.

**Definition 3.4.** A $(p, q, l, M)$-generic system over $V$ is a collection of $q$-decision trees over $V$, $T_i$, for $i \leq M$, with leaf labels from $[M]^p$ such that:

1. each $T_i$ has height at most $l$;
2. each leaf label $g$ of $T_i$ contains $i$;
3. for all $g \in [M]^p$, for all $i, j \in g$, $\text{br}_g(T_i) = \text{br}_g(T_j)$.

Informally, a $(p, q, l, M)$-generic system over $V$ specifies locally consistent pieces of a perfect $p$-partition of $M$ as partial functions of $\{ x_e \}$, where $e \in [V]^q$. (Say that $E$ and $F$ are compatible if $E \not\subseteq F$. By locally consistent, we mean that any mutually compatible set of branches in the trees of the generic system have leaf
labels that are themselves mutually compatible.) When $M$ is congruent to 0 mod $p$, $(p, q, l, M)$-generic systems exist (take any system of height 1 $q$-decision trees $\{T_i\}$ and a $p$-partition $\pi$ of $M$ and label all leaves of $T_i$ by the $p$-class $e \in \pi$ such that $i \in e$). Even when $M$ is not congruent to 0 mod $p$, the existence of a $(p, q, l, M)$-generic system is not inconceivable since there need not be any mutually compatible set of branches that contains a branch from each $T_i$.

**Theorem 3.5.** Let $0 < i < q$ and $R$ be a subset of the set of pairs of integers $(p, j)$ such that $0 < j < p$. Let $d$ be a constant and $k(N)$ a function of $N$ with $k(N) = N^{o(1)}$, and assume that for infinitely many $N$, with $N \equiv i$ (mod $q$), there are depth $d$ size $N^{k(N)}$ Frege proofs of Count$_q^N$ from instances of Count$_R$. Then there is $(p, j) \in R$ such that there are infinitely many $N'$, with $N' \equiv i$ (mod $q$), and $l = O(k(N')^{O(1)})$, with $M = j$ (mod $p$), such that there exists a $(p, q, l, M)$-generic system over a set $V$ of size $N'$.

In the application of Theorem 3.5 we shall use it only for $k(N)$ a constant because we are able to prove Lemma 3.10 only for $l$ a constant independent of $N$.

Now we shall sketch the proof of Theorem 3.5 following closely [16, 7, 19, 5]. In the proof we shall need to apply partial restrictions to simplify formulas. The restrictions that are needed are the following.

**Definition 3.6.** Define the set of restrictions $\mathcal{M}'_m$ to be the set of all partial $q$-partitions $\rho$ of $V$ which cover all but $qm + j$ nodes of $V$ where $j = |V| \mod q$. Every $\rho$ in $\mathcal{M}'_m$ determines a unique assignment to the variables in the formulas:

$$
(x_e)^\rho = \begin{cases} 
1 & \text{if } e \in \rho, \\
0 & \text{if } e \notin \rho \text{ but } e \cap f \neq \emptyset \text{ for some } f \in \rho, \\
x_e & \text{otherwise.}
\end{cases}
$$

We shall denote by $V^\rho$ the set of nodes of $V$ not covered by $\rho$.

Assume that $\Pi$ is a depth of $d$, size $N^{k(N)}$ Frege proof of Count$_q^N$ from instances of some Count$_p^M$ (and thus $M_i = N^{O(k(N))}$ automatically.) We begin by applying a restriction $\rho$ to each formula in $\Pi$ to get a new proof, $\Pi'$ over a smaller universe $N' < N$ with the property that $\Pi'$ has a $k$-evaluation—this is the content of the following lemma.

**Lemma 3.7.** Let $\Pi$ be a depth $d$, size $N^{k(N)}$ Frege proof of Count$_q^N$ from instances of Count$_R$. For some $c_d \leq 5(2q^2)^d$, if $k(N) \leq N^{1/c_d}$, then there exists a restriction $\rho$ such that:

1. $N' = |V^\rho| \geq N^{3c_d}$,
2. $N' \equiv N$ (mod $q$),
3. $\Pi$ restricted by $\rho$ is a proof (over $N'$) of Count$_q^{N'}$ from instances of Count$_R$, and
4. there exists a $k'$-evaluation, $T$, of the subformulas in $\Pi'$ for $k' \leq c_d k(N)$.

The above lemma is proved by inductively generating $k_i$-evaluations (for some
appropriate sequence of values \(k_i\) for the set of subformulas appearing at the \(l\) bottom levels of every formula in \(\Pi\). It is trivial to do this for the literals and constants on the leaves of the formulas. This \(k_i\)-evaluation is extended to a \(k_i+1\)-evaluation of the set of subformulas in \(\Pi\) one level higher by applying a Höstad-style switching lemma \([12]\) on an appropriate class of restrictions. The corresponding switching lemma is stated below. A complete proof of this switching lemma will appear in \([5]\) and \([15]\) where it is also explained in detail how to apply it to obtain the above lemma.

We can apply a restriction \(\rho\) to a \(q\)-decision tree \(T\) over \(V\) in the obvious way to obtain a \(q\)-decision tree \(T^\rho\) over \(V^\rho\).

**Lemma 3.8.** Fix some set \(V\) of vertices and integers \(m\), \(r\), and \(s \geq 0\). Let \(T_i\) be any set of \(q\)-decision trees over \(V\) of height at most \(r\), let \(n = |V|^r\) and let \(u = m/n\). If \(\rho\) is a restriction chosen uniformly at random from \(M^\rho\), then with probability at least \(1 - (4e^{r/2}u^{s-1})^q\), there is a \(q\)-decision tree over \(V^\rho\) of depth at most \(s\) refining and representing \(\bigvee_i b r_i(T^\rho_i)\).

If \(S\) is a set of Boolean formulas closed under subformulas and \(T\) is a \(k\)-evaluation of \(S\) over \(V\) then it is easy to check that the map \(T'\) that sends \(\phi^\rho\) to \(T^\rho\), with the exception when \(\phi\) is an atom such that \(\phi^\rho = 0\) in which case \(\phi^\rho\) is mapped to the height 0 decision tree, is a \(k\)-evaluation of \(S^\rho\) over \(V^\rho\). The exceptional case is needed as for some atoms \(p_e\) it might be that \(p_e^\rho = 0\) while \(T^\rho_e\) need not be the height 0 decision tree, contradicting clause (3) of Definition 3.2. With this observation and the appropriate choice of parameters we obtain Lemma 3.7 from repeated applications of Lemma 3.8.

After applying Lemma 3.7, we are left with a new proof of \(\text{Count}_q^N\) from instances of \(\text{Count}^N\): \(\text{Count}^M_p(\Psi)\) such that \(M_i = j_i \pmod{p_i}\) where \((p_i, j_i) \in R\), together with a \(k'\)-evaluation, \(T\), for all subformulas in the proof, where \(N' = N^e\), \(e > 0\), and \(M_i \leq (N')^{k/e}\) for \(k' = O(k(N))\).

Lemma 3.3 implies (as \(b r(T_\phi) \neq \emptyset\) holds for the final formula \(\eta\) of \(\Pi'\)) that \(b r_0(T_\phi) \neq \emptyset\) for at least one \(\text{Count}_p\) axiom \(\phi\) in \(\Pi'\). Take one such \(\phi\), and any \(G \in b r_0(T_\phi)\) and restrict \(\Pi'\) further by \(G\). In particular, \(\phi\) is reduced to some instance \(\phi^G = \text{Count}_p^M(\psi)\) for which \(b r(T_\phi^G) = b r_0(T_\phi^G)\). To simplify further notation we shall assume that already \(G \subseteq \rho\); hence \(M, N', k', T, \ldots\) remain the same after applying \(G\) and \(b r(T_\phi^G) = b r_0(T_\phi^G)\) for the axiom \(\phi = \text{Count}_p^M(\Psi)\), where \(\Psi = \{\psi_g\}_{g \in \[\mathbb{M}\]}\). Note that the evaluation \(T\) is, in particular, defined on the subformulas \(\psi_g\) of \(\text{Count}_p^M(\psi)\).

**Lemma 3.9.** For all incompatible \(g, h \in \[\mathbb{M}\]\), if \(E\) is the set of edge labels of a branch of \(T_\psi^g\) with the leaf label 1, and \(F\) is the set of edge labels of a branch of \(T_\psi^h\) with the leaf label 1, then \(E \perp F\).

**Proof:** Suppose the claim fails, and let \(g, h \in \[\mathbb{M}\]\) be incompatible, and \(E, F\) be branches labelled 1 in \(T_\psi^g\) and \(T_\psi^h\), respectively, where \(E\) is compatible with \(F\). Then it follows from the definition of evaluation that \(T_{\psi_r \psi_h}\) has a leaf with label 1. But this implies that \(T_\phi\) also has such a leaf, which is a contradiction.

Let \(i \in \[\mathbb{M}\]\). Consider the subformula \(\bigwedge_{i \in \mathcal{G}} \neg \psi_g\) of \(\phi = \text{Count}_p^M(\psi)\). Since \(b r(T_\phi) = b r_0(T_\phi)\), the same is true for \(T_{\bigwedge_{i \in \mathcal{G}} \neg \psi_g}\). Note that \(\bigwedge_{i \in \mathcal{G}} \neg \psi_g\) is just an abbreviation for \(\neg \bigvee_{i \in \mathcal{G}} \psi_g\). Thus we get that \(T_{\bigwedge_{i \in \mathcal{G}} \psi_g}\) has all leaf labels 1. Thus
(again by definition of evaluation) for each branch of this tree, the partial partition extends the partial partition of some branch of \( T_{\phi_k} \) for some \( i \in g \). Furthermore, by the above lemma, this is for exactly one such \( g \). Thus we construct a tree \( T_i \) by putting the corresponding labels \( g \) instead of labels 1 on the leaves.

For all \( i \in [M] \), we can extend the trees \( T_i \) to obtain new trees \( T_i' \), such that for all \( i, j \in g \), the branches of \( T_i' \) with label \( g \) are equal to the branches of \( T_j' \) with label \( g \). The height of the new trees will be equal to \( l = pk' \), where \( k' \) was the height of the original trees \( T_i \). This can be done, for example, as follows. Take some \( i, j \in g \). Take a leaf \( L \) in \( T_i \) labelled by \( g \) and attach \( T_j \) to it. Then prune all incompatible or repeated parts of the attached copy of \( T_j \). There will remain at least one branch of \( T_i \) with the leaf label \( g \) where only edges with repeated labels were cut off. This is because there is a branch in \( T_g \) whose set of edge labels \( E \) is contained in the partial partition of the branch going to \( L \) and also in the partial partition of some branch in \( T_j \). Furthermore, all paths from the root of \( T_j \) with labels consistent with \( E \) can be extended to branches consistent with \( E \). Thus the leaves of \( T_j \) after pruning are a subset of the leaves of the original \( T_j \) with labels \( g \).

If we do it for all leaves of \( T_i \) with the label \( g \), in the new tree we get all possible partial partitions \( E \cup F \) where \( E \) belongs to a leaf labelled by \( q \) in \( T_i \) and \( F \) belongs to a leaf labelled by \( g \) in \( T_j \). By applying the same procedure with \( T_i \) replaced by \( T_j \), we unify the two trees with respect to \( g \). Repeating it for all such pairs (and extending the leaves arbitrarily up to the length \( l \)) we get the desired system of trees.

We are now ready to complete the proof of Theorem 3.5. We will show that the set of \( g \)-partition decision trees \( T_i' \), for \( i \in [M] \), form a \((p, q, l, M)\)-generic system for \( l = O(k(N)) \). By construction, the trees \( T_i' \) have height \( O(pk'/\epsilon) = O(k(N)) \) and the \( T_i' \) are defined over a set of size \( N' = N^e = N^{O(1)} \). Also by construction, each branch in \( T_i' \) with leaf label \( g \) has \( i \in g \). Finally, by the above argument, we have shown that for every \( g \), and every \( i, j \in g \), the set of branches in \( \text{br}(T_i) \) with leaf label \( g \), is equal to the set of branches in \( \text{br}(T_j) \) with leaf label \( g \). Thus, the decision trees \( T_i' \), with \( i \in [M] \), form a \((p, q, l, M)\)-generic system as required by Theorem 3.5.

This completes the proof of Theorem 3.5.

**Lemma 3.10.** Let \( l > 0 \). For all sufficiently large \( N \) such that \( (p^{(l-1)^2 + 1}, q) \mid N \), there does not exist a \((p, q, l, M)\)-generic system over \( N \), with \( M \equiv 0 \pmod{p} \).

**Proof of main Theorem 1.2.** Let

\[
q' = \frac{q}{(q, i)} \quad \text{and} \quad R' = \left\{ \frac{p}{(p, j)} \middle| (p, j) \in R \right\}.
\]

By Corollary 2.4 it suffices to show that there are polynomial size constant-depth Frege proofs of \( \text{Count}_{q',1} \) from instances of \( \text{Count}_{p',1} \), where \( p' \in R' \), if and only if there is some \( p' \in R' \) such that all prime factors of \( p' \) also divide \( q' \).

In the case that some such \( p' \) exists, the statement follows immediately from Lemma 2.5.

Now suppose instead that for every \( p' \in R' \) there is some prime factor of \( p' \) that does not divide \( q' \). Suppose that there are polynomial size constant depth proofs of \( \text{Count}_{q',1} \) from instances of \( \text{Count}_{p',1} \) for \( p' \in R' \). Let

\[
R'' = \{p'/\{p', q\} \mid p' \in R'\}.
\]
By our assumption \( p'/(p', q) \geq 2 \) for \( p' \in R' \); thus, using Lemma 2.1, we get that \( \text{Count}_{r',1} \) has constant depth polynomial size proofs from instances of \( \text{Count}_{r,1} \) for \( r \in R'' \). By Theorem 3.5, for some \( r \in R'' \) there is a \( (r, q', l, M) \)-generic system over \( N' = 1 \text{ (mod } q') \) for constant \( l, M = 1 \text{ (mod } r) \), and \( N' \) arbitrarily large. Since \( (r, q') = 1 \), this is a contradiction to Lemma 3.10.

4. Counting principles and systems of polynomial equations

It is possible to express the propositional formula \( \text{Count}^N_q \) by a system of polynomial equations, \( Q(x) = 0 \), over the ring \( \mathbb{Z}_p \). We will first describe these polynomial equations and then show that Lemma 3.10 follows from a non-constant lower bound on the degree of any linear combination of the \( Q_i \) that equals 1 modulo \( p \).

**Definition 4.1.** Assume that \( N \not\equiv 0 \text{ (mod } q) \). An \((N, q)\)-polynomial system expressing the modulo \( q \) counting principle is the following system of polynomial equations in variables \( x_v, e \in [V]^q, |V| = N \):

\[
(v) \quad \left( \sum_{u \in e} x_u \right) - 1 = 0
\]

one for each \( v \in V \), and

\[
(e, f) \quad x_e \cdot x_f = 0
\]

one for each \( e, f \in [V]^q, e \perp f \).

Denote by \( Q_v \) the left-hand side of equation \((v)\) and by \( Q_{e,f} \) the left-hand side of equation \((e, f)\).

Assume that \( u_e, e \in [V]^q \), is a solution of the polynomial system in some field. The equations \((e, f)\) imply that for each \( v \) at most one \( u_e \) is non-zero for \( v \in e \) and the equation \((v)\) then implies that the unique non-zero \( u_e \) for \( v \in e \) is equal to 1. Hence the set

\[
\{ e \in [V]^q \mid u_e = 1 \}
\]

is a \( q \)-partition of \( V \) which cannot exist when \( N \) is not congruent to 0 modulo \( q \). Thus the above polynomial system has no solution in any field. *Hilbert's Nullstellensatz* then implies the following lemma. We shall not use it, but we state it here for completeness.

**Lemma 4.2.** Let \( \mathbb{F} \) be any field. There are polynomials \( P_v \), with \( v \in V \), and \( P_{e, f} \), with \( e, f \in [V]^q \) and \( e \perp f \), from the ring \( \mathbb{F}[x_e] \) such that equality

\[
\sum_v P_v \cdot Q_v + \sum_{e \perp f} P_{e, f} \cdot Q_{e, f} = 1
\]

holds in the ring \( \mathbb{F}[x_e] \).

We note that, although \( x_e^{[F]} - x_e \) is not present explicitly in the system of polynomials, \( Q_v, Q_{e,f} \), it is easily derived since \( x_e^2 - x_e \) is obtainable as a linear combination \( x_e \cdot Q_v - \sum_{e' \in e, e' \neq e} Q_{e,e'} \) for any \( v \in e \). Thus a non-constant degree lower bound on the \( P_v \) and the \( P_{e,f} \) in the above linear combination also implies
such a non-constant lower bound for an extended system of the type considered in the introduction.

We shall study linear combinations of polynomials $Q_v, Q_{e,f}$ also for the ring $\mathbb{Z}_p$ of counting modulo $p$, where we do not assume that $p$ is prime. Henceforth a
linear combination $L$ means a polynomial of the form

$$\sum_v P_v \cdot Q_v + \sum_{e \in f} P_{e,f} \cdot Q_{e,f}$$

and the degree of $L$ is the maximum degree of the polynomials $P_v, P_{e,f}$.

For a non-empty $q$-partition $E = \{e_1, \ldots, e_i\}$ of $V$ denote by $x_E$ the monomial $x_{e_1} \cdots x_e$, and put $x_\emptyset := 1$.

**Lemma 4.3.** Let $T$ be a $q$-decision tree of height $l$ and assume $l q < N$. Then the polynomial

$$u_T := \left( \sum_{E \in \text{br}(T)} x_E \right) - 1$$

can be expressed as a linear combination of degree at most $l - 1$.

This means that $u_T$ is equal to 1 modulo the ideal generated by polynomials $Q_v, Q_{e,f}$, but the bound on the degrees is also important for us.

**Proof.** Proceed by induction on $l$. For each $l$ we show that the lemma is true for all $N$. For $l = 1$,

$$u_T = \sum_{v \in e} x_v - 1$$

for some $v \in V$ which is just the polynomial $Q_v$ itself. Hence $u_T$ is a linear combination of degree 0.

Assume $l > 1$ and let $v$ be the label of the root of $T$. Then

$$u_T = \sum_{E \in \text{br}(T)} x_E - 1 = \sum_{v \in e} x_v \left( \sum_{f \in \text{br}(T')} x_f \right) - 1.$$  

All $T'$ are $q$-decision trees over the universe $V \setminus e$ of size $N - q$. By the induction hypothesis there are linear combinations $L_e$ of degree at most $l - 2$ formed from unknowns $x_f$ for $f \subseteq (V \setminus e)$ such that

$$L_e = \sum_{f \in \text{br}(T')} x_f - 1,$$

and so

$$u_T = \sum_{v \in e} x_v (L_e + 1) - 1 = \sum_{v \in e} x_v \cdot L_e + \sum_{v \in e} x_v - 1.$$  

The quantity $\sum_{v \in e} x_v \cdot L_e$ is not yet a linear combination over $V$ as $L_e$ are not over the whole $V$; that is, the polynomials $Q'_f$ in $L_e$ in place of $Q_f$ are only of the form

$$\sum_{f: s_i, f \cap e = \emptyset} x_f - 1,$$

rather than of the form

$$\sum_{f \in s_i} x_f - 1.$$  

However, note that

$$x_e Q'_f = x_e Q_f - \sum_{e \in f} x_e x_f.$$
and hence each term $x_iL_i$ is equivalent to a linear combination of degree at most $l - 1$. The remaining quantity in the expression for $u_T$ is just the polynomial $Q_v$; hence $u_T$ is also a linear combination of degree at most $l - 1$.

The next lemma is an important property of generic systems.

**Lemma 4.4.** Let $T_i, i \in [M]$ be a $(p, q, l, M)$-generic system. Then

$$\sum_{i \in [M]} \sum_{E \in \text{br}(T_i)} x_E = 0$$

in the ring $\mathbb{Z}_p$.

**Proof.** Let $g \in [M]^p$ and $S_g = \bigcup_{i \in [M]} \text{br}_g(T_i)$. By the definition of a generic system, for each $i \in g$, $\text{br}_g(T_i) = S_g$ and for each $i \notin g$, $\text{br}_g(T_i) = \emptyset$. Thus for each $g$, each branch in $S_g$ occurs $p$ times in $\bigcup_{i \in [M]} \text{br}(T_i)$, once for each of the elements $i \in g$, and hence

$$\sum_{i \in [M]} \sum_{E \in \text{br}(T_i)} x_E \equiv 0 \pmod{p}.$$

The lemma below follows from the previous two lemmas.

**Lemma 4.5.** If there is a $(p, q, l, M)$-generic system $T_i, i \in [M]$, such that $l, q < N$, then there is a linear combination $L$ of degree at most $l - 1$ such that $L + M = 0$ in the ring $\mathbb{Z}_p[\bar{x_e}]$.

**Proof.** By Lemma 4.3, we can write the sum

$$\sum_{i \in [M]} \sum_{E \in \text{br}(T_i)} x_E$$

as

$$\sum_{i \in [M]} (L_i + 1) = \sum_{i \in [M]} L_i + M,$$

where the $L_i$ are linear combinations of degree at most $l - 1$. But by Lemma 3.4,

$$\sum_{i \in [M]} \sum_{E \in \text{br}(T_i)} x_E = 0,$$

and thus we have $\sum_{i \in [M]} L_i + M = 0$.

The following is the main technical lemma of this paper.

**Lemma 4.6 (main).** Let $d$ be a constant, let $N$ be sufficiently large and suppose that $N$ satisfies $(p^{d_i+1}, q) \mid N$ and $M \neq 0 \pmod{p}$. Then every linear combination $L$ such that $L = M$ in $\mathbb{Z}_p[\bar{x_e}]$ must have degree larger than $d$.

Put otherwise, linear combinations expressing a constant other than 0 cannot have a constant degree.

We shall prove the main lemma in the next section; now we infer Lemma 3.10 from it.
Proof of Lemma 3.10 from Lemma 4.6. Assume that for some constant $l$ there exists a $(p, q, l, M)$-generic system $T_i, i \in [M], M \neq 0 \pmod{p}$, over some $N$ such that $(p^{(l-1)^2+1}, q) \mid N$.

If $N = 0 \pmod{q}$ then there is some perfect $q$-partition $\pi$ of $N$. For each $q$-decision tree $T_i$, there is some branch $E_i$ in $T_i$ such that $E_i \subset \pi$. By the definition of generic systems, the leaf labels of these branches form a perfect $p$-partition of $M$ which is impossible since $M \neq 0 \pmod{p}$.

Suppose now that $N \neq 0 \pmod{q}$. By Lemma 4.5 the existence of this $(p, q, l, M)$-generic system over $N$ implies the existence of a linear combination, $L$, of degree at most $l - 1$ such that $L = -M$ in the ring of polynomials $\mathbb{Z}_p[\bar{x}_e]$. But this contradicts Lemma 4.6 because $-M$ is not congruent to 0 mod $p$.

5. Proof of the lower bound on the degree of the polynomials

In this section we prove the main Lemma 4.6. It is an immediate corollary of the following lemma.

**Lemma 5.1.** Let $d$ be a constant, let $N$ be sufficiently large and suppose that $N$ satisfies $(p^{d^2+1}, q) \mid N$. If $P_v$, for $v \in V$ and $|V| = N$, are of degree at most $d$, then there exists a 0-1 assignment $a$ such that for every $e \perp f$,

$$Q_{e,f}(a) = 0 \pmod{p} \quad (3)$$

and

$$\sum_{v \in V} P_v(a) Q_v(a) = 0 \pmod{p}. \quad (4)$$

The rest of the section is devoted to the proof of this lemma. Before we prove it, we have to do some preliminary work.

A 0-1 assignment corresponds to a set of $q$-element sets (those for which $x_e(a) = 1$). Thus (3) means that we consider a set of disjoint $q$-element sets.

**Convention.** From now on we consider only such assignments.

Since we shall evaluate polynomials on 0-1 inputs, we can replace any $x_e^d$ with $d > 1$ by $x_e$. Thus we shall assume that all polynomials are multilinear. (This assumption could also easily have been justified by the fact that for any $e$, $x_e^2 - x_e$ is obtainable as a linear combination of degree 1.) Let $\Xi = x_{e_1} \ldots x_{e_d}$ be a monomial of $P_v$. If $v \in e_j$, for some $j$, then $\Xi(a)Q_v(a)$ is always 0, since if $Q_v(a) \neq 0$, then $v$ is not covered by a $q$-element set from $a$; thus $x_e(a) = 0$ and $\Xi(a) = 0$. Furthermore, if $e_j \cap e_k \neq 0$, for some $j \neq k$, then $\Xi(a)$ is always 0, since the $q$-element sets in $a$ are disjoint.

**Convention.** From now on we consider only systems $\mathcal{P} = \{P_v\}_{v \in V}$, where $v$ is not contained in any of the $e_1, \ldots, e_d$ and $e_1, \ldots, e_d$ are disjoint for any monomial $\Xi = x_{e_1} \ldots x_{e_d}$ of $P_v$ occurring with non-zero coefficient.

Fix $d$. All systems $\mathcal{P}$ we consider have degree $d$ (that is, all polynomials $P_v$ have degree at most $d$.) For a monomial $\Xi = x_{e_1} \ldots x_{e_d}$, we denote by

$$\text{supp}(\Xi) = e_1 \cup \ldots \cup e_d$$

the support of $\Xi$. 
The system of polynomials \( P = \{ P_v \}_{v \in V} \) is determined by a sequence of coefficients of the form \( a_{\gamma} \in \mathbb{Z}_p \) for \( \gamma = (v, \Xi) \), where \( v \in V \), \( \Xi = x_{e_1} \cdots x_{e_c} \) for \( c \leq d \), \( e_1, \ldots, e_c \in [\mathbb{V}^q] \), and \( \{ v \} \), \( e_1, \ldots, e_c \) are disjoint. We shall assume that the \( q \)-element sets are ordered by their least elements, that is, \( \min e_1 < \cdots < \min e_c \).

Thus we get a 1-1 correspondence between \( (v, \Xi) \) with \( \Xi \) of degree \( c \) and \( q^c + 1 \)-tuples \( (v, e_1, \ldots, e_{c_1}, \ldots, e_{c_q}) \) of distinct elements from \( V \) such that \( e_{i_1} < e_{i,j} \) for \( i < j \) and \( e_{i,j} < e_{i,k} \) for \( j < k \).

**Definition 5.2.** (1) We define \( \text{type}(v, \Xi) \) to denote the isomorphism type of the structure
\[
(\{ v \} \cup \text{supp}(\Xi) ; v, e_1, \ldots, e_c, \leq). \]

(2) We define \( k\)-type \( (v, \Xi) \) over \( V \) to denote the isomorphism type of the structure
\[
(\{ v \} \cup \text{supp}(\Xi) ; v, e_1, \ldots, e_c, \leq, R_0, \ldots, R_{k-1}), \]
where \( R_0, \ldots, R_{k-1} \) are unary predicates defined by
\[
R_i(x) \Leftrightarrow_{af} x = v_i \text{ and } i = j \mod k, \]
where \( V = \{ v_1, \ldots, v_N \} \), \( v_1 < \cdots < v_N \).

**Definition 5.3.** (1) We say that \( P \) is *symmetric*, if for every type \( T \) the \( a_{\gamma} \) are the same for all \( \gamma = (v, \Xi) \) of type \( T \).

(2) We say that \( P \) is \( k\)-*symmetric* over \( V \), if for every \( k\)-type \( T \) over \( V \) the \( a_{\gamma} \) are the same for all \( \gamma = (v, \Xi) \) of \( k\)-type \( T \) over \( V \).

**Lemma 5.4.** If \( P \) is \( p^k\)-symmetric over \( V \), \( |V| = N \), \( (p^{k+1}, q) \mid N \) and \( N \geq p^k + 1 \), then there exists an assignment \( a \) such that
\[
\sum_{v \in V} P_v(a)Q_v(a) = 0 \pmod{p}. \]

**Proof.** Take \( t \geq 0 \) such that
\[
N - qt \geq 0 \quad \text{and} \quad N - qt = 0 \pmod{p^{k+1}}. \]

Let \( a \) consist of a \( q \)-partition of the last \( qt \) elements of \( V \), the rest being uncovered. Then for \( v > N - qt \), \( P_v(a)Q_v(a) = 0 \), since \( Q_v(a) = 0 \) already. For \( 1 \leq v \leq N - qt \) we have \( Q_v(a) = -1 \) and all monomials \( \Xi \) in \( P_v \) that are non-zero under \( a \) have support contained in the last \( qt \) elements of \( V \). Thus, the choice of \( v \leq N - qt \) does not affect the type\( (v, \Xi) \) for any monomials \( \Xi \) that are non-zero under \( a \). By \( p^k\)-symmetry, for any two \( v, v' \leq N - qt \) that agree modulo \( p^k \), the coefficients of all monomials with support among the last \( qt \) elements of \( V \) are the same in \( P_v \) and \( P_{v'} \). Since \( p^{k+1} \) divides \( N - qt \), the number of \( v \leq N - qt \) congruent to \( i \) modulo \( p^k \) is divisible by \( p \) for each fixed \( 0 \leq i < p^k \). Thus \( p^k\)-symmetry implies that for every \( i \),
\[
\sum_{v \leq N - qt, v \equiv i \pmod{p^k}} P_v(a)Q_v(a) = 0 \pmod{p}, \]
whence the lemma follows.
We note that for the argument above we actually used relatively few of the properties of \( p^k \)-symmetry. However, the notion of \( p^k \)-symmetry is more natural for the application of Ramsey's theorem and it facilitates the inductive nature of the argument.

To apply Ramsey's theorem we shall employ restrictions from Definition 3.6. For a restriction \( \rho \), we shall denote by \( V^\rho \) the set of vertices that are not covered by \( q \)-element sets of \( \rho \); \( \mathcal{P}^\rho \) denotes the restricted system of polynomials. Note that after applying such a restriction, we get the same counting principle on \( V^\rho \) which has the same number of vertices as \( V \) modulo \( q \). Also an assignment on \( V^\rho \) combined with the restriction is an assignment on the whole set \( V \) (with the properties that we need).

Now we want to apply Ramsey's theorem to get a \( p^k \)-symmetric system of polynomials after applying some restriction \( \rho \). There are two problems. First, a single application of Ramsey's theorem will only allow us to symmetrize with respect to monomials of some particular degree (since the signatures of monomials of different degrees are different). When we symmetrize with respect to the monomials of some degree \( c \), we apply a restriction and this restriction may also create new monomials of degree \( c \) from monomials of larger degree. Thus it makes sense to symmetrize starting with monomials of large degree first in the hope that the newly created monomials will occur symmetrically. However, the second problem is that if the monomials of degree \( d \) are \( p^k \)-symmetric then it turns out (an example can be given) that it is not possible to achieve \( p^k \)-symmetry for monomials of smaller degree. This is resolved by starting with \( p^r \)-symmetry for \( r < k \) (a stricter notion) for the larger degrees and then relaxing it as the smaller degrees are handled and being careful about the exact details of constructing the restriction.

Suppose that \( |V| = N \), \( V' = V^\rho = \{v_1, \ldots, v_m\} \), \( v_1 < \cdots < v_m \) for some restriction \( \rho \) and that \( \mathcal{P} \) has been made \( p^r \)-symmetric over \( V \) with respect to monomials of degree greater than \( d' \). To argue that the contribution to monomials of degree \( d' \) from monomials in \( \mathcal{P} \) of larger degree is \( p^s \)-symmetric over \( V' \) for some \( s > r \), we will argue that for any \( p^r \)-type \( T' \) over \( V' \) of degree \( d' \), the contribution (modulo \( p \)) to the coefficient of any \( (v, \Xi) \) of \( p^s \)-type \( T' \) over \( V' \) from the restriction of monomials of \( p^r \)-type \( T \) over \( V \) is the same. By the \( p^s \)-symmetry of the larger degree terms over \( V \) we need only count the number of \( (v, \Xi, \Xi_2) \) of \( p^r \)-type \( T \) over \( V \) that contribute to the coefficient of a given \( (v, \Xi) \) of \( p^s \)-type \( T' \) over \( V' \).

Before going into the actual construction, we note that for a given \( (v, \Xi) \), the number of \( \Xi_2 \) such that \( (v, \Xi, \Xi_2) \) is of a given \( p^r \)-type over \( V \) and \( \Xi_2 = 1 \) is affected by the relative order of \( \{v\} \cup \text{supp}(\Xi) \) among the elements of the \( q \)-sets in \( \rho \). Thus, after we choose the set \( V' \) using Ramsey's theorem, we have to be careful when we choose \( \rho \) such that \( V' = V^\rho \) to be very careful how the \( q \)-element sets in \( \rho \) cover the nodes in \( V \setminus V' \) that are between elements of \( V' \). For convenience we say any such \( q \)-element set interleaves the set \( V' \). The following lemma shows that we can construct a \( V' \) and a restriction \( \rho \) with \( q \)-element sets that interleave \( V' \) in a nice way. Afterwards we will argue that these are sufficient to give our desired result.

**Lemma 5.5.** For every \( m, d', k' > 0 \) there exists an \( N_0 \) such that, for every \( N \geq N_0 \) with \( N \equiv m \pmod{q} \) and every \( \mathcal{P} \) over any \( V \), with \( |V| = N \), there exists a restriction \( \rho \) such that
(1) the monomials of degree $d'$ in $\mathcal{P}$ in variables over the set $V^p$ form a symmetric system;

(2) $|V^p|=m$;

(3) for every $v, v' \in V^p$, $v \equiv v' \pmod{k'}$;

(4) if we let $U = \{u_1, \ldots, u_z\} \subseteq V \setminus V^p$ be the set of elements that are between elements of $V^p$ then $\rho$ contains the sets $\{u_i, u_i^{(1)}, \ldots, u_i^{(q-1)}\}$, where for $i = 1, \ldots, z$ we take disjoint sets $\{u_i^{(1)}, \ldots, u_i^{(q-1)}\}$ consecutively, either starting from the largest element of $V^p$ and working downwards, or starting from the smallest element of $V^p$ working upwards.

Proof. For any set $V$ and any system of polynomials $\mathcal{P}$ define a colouring of $V$ by assigning to $S \in [V]^{qd'+1}$ the value of the coefficients whose index is supported by $S$. More precisely, let us order arbitrarily the set of types $T$ applicable to monomials of degree $d'$. The colour of $S \in [V]^{qd'+1}$ is defined to be the sequence of $a_T$ for all such types where $a_T$ is the coefficient in $\mathcal{P}$ of the unique $\gamma = (v, \Xi)$ of type $T$ with $\{v\} \cup \text{supp}(\Xi) = S$. By Ramsey's theorem there is an $N_0$ such that if $|V| \geq N_0$ and $\mathcal{P}$ is any system of polynomials over $V$, there exists $V' \subset V$, with $|V'| = 2k'(q - 1)m$, such that $[V']^{qd'+1}$ is monochromatic. Thus, by definition, for any $V'' \subset V'$, the terms of degree $d'$ in $\mathcal{P}$ with variables over the set $V''$ form a symmetric system. By a trivial-averaging argument, there is a subset $V''$ of $V'$ of size $2(q - 1)m$ such that all elements of $V''$ agree modulo $k'$.

We will assume that $|V| = m \pmod{q}$ from now on. Suppose that the elements of $V''$ are partitioned into $(q - 1)$ consecutive segments

$$\{v_1, \ldots, v_m\}, \{v_{m+1}, \ldots, v_{2m}\}, \ldots, \{v_{(2(q-1)-1)m+1}, \ldots, v_{2(q-1)m}\},$$

where the $v_i$ are listed in increasing order. Let $V'''$ be one of these segments with minimal distance between its first and last elements. If $V''' = \{v_{(i-1)m+1}, \ldots, v_{im}\}$ is in the first half of the segments, then there are at least $(q - 1)$ times as many elements in $V$ (in fact in $V'''$) larger than all elements in $V'''$ as there are elements in $[v_{(i-1)m+1}, v_{im}]$, that is, at least $(q - 1)(v_{im} - v_{(i-1)m+1} + 1)$ of them. In this case choose $\rho$ arbitrarily so that $V^p = V'''$ and the elements

$$\{u_1, \ldots, u_z\} = [v_{(i-1)m+1}, v_{im}] \setminus V'''$$

are matched under $\rho$ so that the other elements in the same set as $u_i$ are $v_{im} + (j - 1)(q - 1) + 1, \ldots, v_{im} + j(q - 1)$. If $V'''$ is in the second half of the segments then the other elements in the same set as $u_i$ under $\rho$ are $v_{(i-1)m} - (j - 1)(q - 1) - 1, \ldots, v_{(i-1)m} - j(q - 1)$. Arbitrarily choose the $q$-sets of $\rho$ for the remaining elements of $V \setminus V'$. This is possible since $N = m \pmod{q}$.

It is well known that

$$i \equiv j \pmod{p'} \Rightarrow \binom{i}{j} \equiv \binom{i}{j} \pmod{p}$$

(in fact a weaker assumption suffices). We need the following generalization.

**Lemma 5.6.** Let $0 < i_1, \ldots, i_t < k$ be fixed and consider all $n \geq 0$. Let $C$ be the
number of possible choices of \((b_1, \ldots, b_l)\) with \(0 \leq b_1 < \ldots < b_l < n\) satisfying the condition

\[
\begin{align*}
  b_1 &\equiv i_1 \pmod{k}, \\
  &\vdots \\
  b_l &\equiv i_l \pmod{k}.
\end{align*}
\]

Then the residue class modulo \(p\) of \(C\) is determined by \(n\) modulo \(kp'\).

**Proof.** This is by induction on \(l\). For \(l = 1\) it is trivial, even for \(n \equiv 0 \pmod{kp'}\). Suppose it holds for \(l\). Let \(0 \leq n_1 < n_2\) and \(n_1 \equiv n_2 \pmod{kp'^{l+1}}\). Then for every \(i\), with \(0 \leq i < kp'\), the number of elements \(b_{l+1}\), with \(n_1 \leq b_{l+1} < n_2\) and \(b_{l+1} \equiv i \pmod{kp'}\), is divisible by \(p\). For each such \(b_{l+1}\), we have the same number, modulo \(p\), of sequences \((b_1, \ldots, b_l)\) such that \(0 \leq b_1 < \ldots < b_l < b_{l+1}\) satisfying (5) by the inductive assumption. Thus the number of the sequences \((b_1, \ldots, b_l, b_{l+1})\) with \(n_1 \leq b_{l+1} < n_2\) and \(b_{l+1} \equiv i \pmod{kp'}\) is congruent to 0 modulo \(p\).

Now we state a lemma which essentially formalizes the induction step of the proof of Lemma 5.1.

**Lemma 5.7.** For every \(m > 0\) and \(d' \leq d\), there exists \(N_0\) such that for every \(N \geq N_0\) with \(N \equiv m \pmod{q}\) and every \(P\) of degree \(d\) over \(V\), with \(|V| = N\), if the monomials of \(P\) of degrees \(d' + 1, \ldots, d\) form a \(k\)-symmetric system of polynomials, then there is a restriction \(p\) such that \(|V_p| = m\) and the monomials of degree \(d', d' + 1, \ldots, d\) of \(P_p\) form a \(kp^d\)-symmetric system over \(V_p\).

**Proof.** By Lemma 5.5, there is an \(N_0\) such that for any \(V\), with \(|V| \geq N_0\) and \(|V| = N \equiv m \pmod{q}\), and any \(P\) over \(V\), there is a restriction \(p\) satisfying conclusions (1)–(4) of Lemma 5.5 with \(k' = kp^d\). Choose this \(N_0\) and \(p\). By conclusion (1) of Lemma 5.5, the monomials of degree \(d'\) in \(P\) over \(V_p\) appear symmetrically. Thus, since the monomials of degree greater than \(d'\) in \(P\) form a \(k\)-symmetric system, it is sufficient to show that if \(P'\) is any \(k\)-symmetric system of monomials of degree at most \(d\) over \(V\) then \(P'^p\) is \(kp^d\)-symmetric over \(V_p\).

Let \(V' = V^p = \{v_1, \ldots, v_m\}\) and write \(p = p_1p_2\) where \(p_1\) is the portion of \(p\) that interleaves \(V'\).

Consider first \(P'^{p_1}\) over \(V^{p_1}\). Note that \(V^{p_2}\) is a consecutive sequence of elements of \(V\). Thus, for any two \((v, 2)\) and \((v', 2')\) of \(k\)-type \(T\) over \(V^{p_2}\), if \((v, 2, 2')\) is of \(k\)-type \(T'\) over \(V\) of degree greater than \(d'\) then so is \((v', 2, 2')\). Thus \(P^{p_1}\) is \(k\)-symmetric over \(V^{p_1}\).

It remains to see what happens with monomials after applying \(p_1\). That is, we consider \(P^p = P^{p_2p_2}\). Since \(P^{p_2}\) is \(k\)-symmetric over \(V^{p_2}\) and the elements of \(V^{p_2}\) are consecutive, we can ignore the differences between \(P^{p_2}\) and \(P\) and between \(V^{p_2}\) and \(V\). Thus we want to show that for an arbitrary \(k\)-symmetric system \(P'\) over \(V\) of degree at most \(d\) such that \(V^{p_1} = V'\), \(P^{p_1}\) is \(kp^d\)-symmetric over \(V'\).

**Lemma.** Suppose that \(T\) is a \(k\)-type of degree at most \(d\) over \(V\) and \(T'\) is a \(kp^d\)-type over \(V'\) of degree \(c < d\). For any \((v, 2)\) of \(kp^d\)-type \(T'\) over \(V'\), the number (modulo \(p\)) of \(2\) such that \((v, 2, 2')\) has \(k\)-type \(T\) over \(V\) and \(2' = 1\) is the same.
We will prove the lemma by considering some fixed \((v, \Xi_1)\) and see that the number of monomials \(\Xi_2\) such that \((v, \Xi_1, \Xi_2)\) has some fixed \(k\)-type \(T\) over \(V\) and \(\Xi_2 = 1\) depends only on the \(kp^d\)-type of \((v, \Xi_1)\) over \(V'\).

Let \(\{v\} \cup \text{supp}(\Xi_1) = \{v_{10}, \ldots, v_{1n}\}\) in increasing order (\(c\) is the degree of \(\Xi_1\)). Consider possible monomials \(\Xi_2\) of degree \(b\) such that \(b + c \leq d\) and \(\Xi_2 = 1\). Each \(q\)-set fixed by \(\rho_1\) is determined by a single representative \(u \in [v_1, v_m] \setminus V'\) which determine the monomial \(\Xi_2\). Since \(\Xi_1\) is fixed, by construction of \(\rho_1\) the \(k\)-type of \((v, \Xi_1, \Xi_2)\) over \(V\) depends only on two properties:

(a) the position of \(u_{i_0}, \ldots, u_i\) with respect to \(v_{i_0}, v_{i_1}, \ldots, v_{i_w}\) (which fixes the type of \((v, \Xi_1, \Xi_2)\) since the order of the least elements in the \(q\)-sets containing the \(u_j\) is either always the order of the \(u_h\) or always the reverse);

(b) the residue classes modulo \(k\) of elements

\[\{u_{j_0}, u_{j_1}^{(1)}, u_{j_2}^{(2)}, \ldots, u_{j_{q-1}}^{(q-1)}\}, \ldots, \{u_{j_s}, u_{j_{s+1}}^{(1)}, u_{j_{s+2}}^{(2)}, \ldots, u_{j_{s+q-1}}^{(q-1)}\}.\]

The key observation for (b) is that the residue classes of \(u_{j_1}^{(1)}, u_{j_2}^{(2)}, \ldots, u_{j_{q-1}}^{(q-1)}\) are precisely determined by the residue class modulo \(k\) of \(u_{j_h}\) and the residue class modulo \(k\) of \(k\) of the number of vertices in \(V'\) less than \(u_{j_h}\). This is because the residue classes of \(u_{j_h}^{(1)}, u_{j_h}^{(2)}, \ldots, u_{j_h}^{(q-1)}\) are determined just by \(j_h\) modulo \(k\) and the difference between the residue classes of \(u_{j_h}\) and \(u_{j_l}\) depends on the number of vertices in \(V'\) less than \(u_{j_h}\). (The difference is either positive or negative depending upon whether \(\rho_1\) matches the elements \([v_{i_0}, v_{i_w}] \setminus V'\) above or below \([v_1, v_m]\).)

Since we only consider \(\Xi_1\) such that \((v, \Xi_1, \Xi_2)\) has some fixed \(k\)-type \(T\) over \(V\), for each \(\alpha\), with \(0 \leq \alpha \leq qc\), we have fixed the indices \(\beta, \ldots, \gamma\) such that \(v_{i_\beta} < u_{j_\beta} < \ldots < u_{j_\gamma} < v_{i_{\alpha+1}}\), where \(\beta\) and \(\gamma\) depend only on \(\alpha\). We shall handle each such interval \(v_{i_\beta}, v_{i_{\alpha+1}}\) separately and show that (modulo \(p\)) the number of choices of such \(u_{j_\beta}, \ldots, u_{j_\gamma}\), between \(v_{i_\beta}\) and \(v_{i_{\alpha+1}}\) such that

\[\{u_{j_\beta}, u_{j_\beta}^{(1)}, u_{j_\beta}^{(2)}, \ldots, u_{j_\beta}^{(q-1)}\}, \ldots, \{u_{j_\gamma}, u_{j_\gamma}^{(1)}, u_{j_\gamma}^{(2)}, \ldots, u_{j_\gamma}^{(q-1)}\}\]

belong to particular residue classes modulo \(k\) depends only on the residue classes of \(i_\alpha\) and \(i_{\alpha+1}\) modulo \(kp^d\). This will be sufficient since the total number of choices of \(\Xi_2\) is the product of the number of choices in each of these intervals and the \(kp^d\)-type of \((v, \Xi_1)\) over \(V'\) determines these residue classes.

By the key observation, the only further condition that the \(k\)-type \(T\) places on \(u_{j_\beta}, \ldots, u_{j_\gamma}\) is given by a sequence of pairs \((k_\beta, k'_\beta), \ldots, (k_\gamma, k'_\gamma)\) such that, for \(\beta \leq t \leq \gamma\), \(0 \leq k_t, k'_t < k\) and

\[u_{j_t} = k_t \quad (\text{mod } k);\]

and

\[l_t = k'_t \quad (\text{mod } k),\]

where \(l_t\) is the index such that \(v_{i_t} < u_{j_t} < v_{i_{t+1}}\).

We would like to apply Lemma 5.6 to the equations for \(u_{j_t}\) and \(l_t\) above and argue that the number of solutions only depends on \(i_\alpha\) and \(i_{\alpha+1}\) modulo \(kp^d\). We cannot do so immediately since it is possible, if \(k'_t = k'_{t+1}\), that \(l_t = l_{t+1}\) and Lemma 5.6 does not apply in this case. Instead we will break up the cases into the possible partitions \(\pi\) of the interval \([\beta, \gamma]\) into intervals \([\mu_1, \nu_1], \ldots, [\mu_{\ell}, \nu_{\ell}]\) so that for \(t, t'\) in the same interval \(l_t = l_{t'}\) and for \(t, t'\) in different intervals \(l_t \neq l_{t'}\). It is now sufficient to argue two things for each fixed partition \(\pi\):

1. **Residue Classes:**
   - The set of residue classes \(\{u_{j_t}, u_{j_t}^{(1)}, u_{j_t}^{(2)}, \ldots, u_{j_t}^{(q-1)}\}\) determined by \(u_{j_t}\) and \(k_t\) modulo \(k\) depends only on \(\pi\).
   - The set of residue classes \(\{u_{j_t}, u_{j_t}^{(1)}, u_{j_t}^{(2)}, \ldots, u_{j_t}^{(q-1)}\}\) determined by \(u_{j_t}\) and \(k'_t\) modulo \(k\) depends only on \(\pi\).

2. **Independence:**
   - For any two partitions \(\pi_1\) and \(\pi_2\) of \([\beta, \gamma]\), the sets of residue classes determined by \(u_{j_t}\) and the \(k_t, k'_t\) for \(t \in [\beta, \gamma]\) are independent, meaning that the choice of \(u_{j_t}\) and \(k_t, k'_t\) for one partition does not affect the choices for the other partition.

By applying Lemma 5.6 to each possible partition \(\pi\), we can conclude that the number of solutions only depends on \(i_\alpha\) and \(i_{\alpha+1}\) modulo \(kp^d\).
(1) the number of choices (modulo $p$) of the sequence $l_i$ consistent with the $k$-type $T$ and the partition $\pi$ depends only on $i_\alpha$ and $i_{\alpha+1}$ modulo $k^p$;

(2) the number of choices (modulo $p$) of $u_{j_0'}, \ldots, u_{j_\gamma}$, consistent with $k$-type $T$ and a fixed sequence of $l_i$ that is consistent with $T$, is independent of the choice of $(v, \Xi)$.

Given the fixed partition $\pi$, the sequence of $l_i$ for $t \in [\beta, \gamma]$ is precisely determined by $l_{\mu_1}, \ldots, l_{\mu_t}$ where $i_\alpha \leq l_{\mu_1} < \ldots < l_{\mu_t} < i_{\alpha+1}$ and $l_{\mu_t} = k^{i_{\mu_t}}$ (mod $k$) for $i = 1, \ldots, \xi$. By Lemma 5.6, the number of such solutions depends only on $i_{\alpha+1} - i_\alpha$ modulo $k^p$ which is determined by $i_{\alpha+1} - i_\alpha$ modulo $k^p$ since $\xi \leq d$.

Now consider the fixed sequence $l_i$ and its associated partition $\pi$. For each interval $[\mu_i, \nu_i] = [r, s]$ in $\pi$ we count the number of choices of $u_{j_r}, \ldots, u_{j_s}$ consistent with the $k$-type $T$. The solutions $u_{j_r}, \ldots, u_{j_s}$ precisely satisfy

$$v_i + 1 < u_i < \ldots < u_j < v_{i+1}$$

and $u_i = k_i$ (mod $k$) for $t \in [r, s]$. By Lemma 5.6, the number of such choices modulo $p$ depends only $v_{i+1} - (v_i + 1)$ modulo $k^{i-r+1}$. Since all elements of $V'$ are equivalent modulo $k^p$ and $d \geq s - r + 1$, $v_{i+1} - (v_i + 1)$ is always congruent to $-1$ modulo $k^{i-r+1}$ and thus the number of choices modulo $p$ in each interval of $\pi$ is independent of the choice of $(v, \Xi)$. Therefore the number of choices modulo $p$ of $u_{j_0'}, \ldots, u_{j_\gamma}$, consistent with the sequence of $l_i$ and the $k$-type $T$ is independent of the choice of $(v, \Xi)$, as required.

Thus we have proved the lemma, and hence the system $\mathcal{P}^\rho$ is $k^p$-symmetric, which finishes the proof of Lemma 5.7.

\textbf{Proof of Lemma 5.1.} Let $m \geq p^{d^2+1}q$ and $(p^{d^2+1}, q) | m$. (Choose $V$, with $|V| = N \equiv m$ (mod $q$), large enough to apply Lemma 5.7 for $d' = d, d - 1, \ldots, 1$, in order, to a system $\mathcal{P} = \{P_v\}_{v \in V}$ of degree $d$ and still have the combined restriction so constructed have $|V^\rho| = m$. Choose any such system $\mathcal{P}$ and note that Lemma 5.7 implies that there is a restriction $\rho$ with $|V^\rho| = m$ and that $\mathcal{P}^\rho$ is $p^d$-symmetric over $V^\rho$. By Lemma 5.4 there is an assignment $a$ on which $\mathcal{P}^\rho$ vanishes. Combining $\rho$ with $a$, we get an assignment on which the $\mathcal{P}$ vanishes. Finally, recall that all $Q_{ef}$ vanish too, due to the fact that $\rho$ and $a$ are partial partitions.

\section{Frege systems with modular counting}

As remarked in the introduction, the development of lower bounds for propositional proof systems has followed the development of lower bounds in circuit complexity. Thus it might at first seem surprising that we are able to prove lower bounds that separate Count$_5$ from Count$_6$ when it is not known whether counting modulo 5 can be computed by boolean circuits of bounded depth and polynomial size with mod 6 gates.

However, the kind of question we resolve is not the appropriate statement analogous to this circuit complexity question. This circuit complexity question motivates the following extension of a Frege system by \textit{modular counting gates}.

\textbf{Definition 6.1.} Let $r \geq 2$.

1. The MOD$_r$-connectives are countably many connectives

\[ \text{MOD}_{r, \ell}(x_1, \ldots, x_n) \]
where $0 \leq i < r$, $n = 1, 2, \ldots$ The formula $\text{MOD}_{r,i}(x_1, \ldots, x_n)$ is true if and only if

$$|\{j \leq n \mid x_j \text{ true}\}| = i \pmod{r}.$$  

We shall write $x_1 + \ldots + x_n = i$ in place of $\text{MOD}_{r,i}(x_1, \ldots, x_n)$.

2. The $\text{MOD}_r$-rules are the following inference rules:

(a) $\frac{x}{x' = 1}$, $\frac{x = 1'}{x = 0'}$, $\frac{x = 0'}{x'}$;

(b) $\frac{x_1 + \ldots + x_n = i \quad y_1 + \ldots + y_m = j}{x_1 + \ldots + x_n + y_1 + \ldots + y_m = k}$, where $i + j = k \pmod{r}$;

(c) $\frac{x_1 + \ldots + x_n + y_1 + \ldots + y_m = k \quad x_1 + \ldots + x_n = i}{y_1 + \ldots + y_m = j}$, where $k - i = j \pmod{r}$;

(d) $\frac{x_1 + \ldots + x_n = i}{x_{j_1} + \ldots + x_{j_n} = i}$, where $j_1, \ldots, j_n$ is any permutation of $1, \ldots, n$.

3. A $\text{MOD}_r$-Frege system is a Frege system whose language is extended by all the $\text{MOD}_r$-connectives and with all the $\text{MOD}_r$-rules.

4. The depth of a formula in the extended language is the maximum number of alternations of $\to$, $\lor$ and $\text{MOD}_{r,i}$. A depth $d$ $\text{MOD}_r$-Frege system is a $\text{MOD}_r$-Frege system in which only formulas of depth at most $d$ are allowed.

We leave it to the reader to verify that a $\text{MOD}_r$-Frege system is complete, that is, it proves all tautologically valid formulas in the extended language.

The following lemma is not hard.

**Lemma 6.2.** Let $s \geq r \geq 2$ and assume that $r$ divides $s$. Then formulas from Count, have polynomial size, constant-depth $\text{MOD}_r$-Frege proofs.

However we cannot resolve the more general question of whether or not Count has polynomial size, constant-depth Frege proofs in a $\text{MOD}_p$-Frege system, where $p$ and $q$ are relatively prime. It seems that such proofs should not exist but stronger proof techniques than ours seem necessary, since the Count axioms are proved in a $\text{MOD}_p$-Frege system using the $\text{MOD}_p$ connective only in a very trivial way.

7. Concluding remarks

We have introduced a natural approach to proving lower bounds for propositional proof systems that is based on studying the complexity of the Nullstellensatz polynomials witnessing the unsolvability of a system of equations.

One important open question is whether or not the lower bound in our main theorem can be improved. We conjecture that the degree lower bound is nearly
linear, although the techniques of this paper only succeed in proving a non-constant lower bound. Note that an exponential lower bound on the size of constant-depth Frege proofs of $\text{Count}_q^p$ from $\text{Count}_p$ instances would follow from an improvement of the degree lower bound to $n^e$, for some $e > 0$.

It is interesting to compare the methods used in this paper and those of Ajtai [4, 3]. In [4], using switching lemma techniques, the author reduces the existence of constant-depth polynomial size proofs of $\text{Count}_q$ from $\text{Count}_p$ to a question about uniform sequences of symmetric linear equations over $\mathbb{Z}_p$. Then the results of [3] are used to derive the conclusion that this system has no solution. The proofs in [3] use a detailed analysis of the structure theory of representations of the symmetric group to argue that as the objects in this theory are built up, the constructions made are appropriately uniform.

Assuming that constant-degree polynomials suffice, we conclude that the equations for the coefficients of the Nullstellensatz polynomials for the systems we generate are uniform sequences of symmetric linear equations in the sense considered in [3]. However, the proof that we obtained by looking at the polynomials themselves is considerably simpler than would be obtained by applying an argument of the form of [3]. Also, because the reduction in [4] critically uses the property of 'covering sets', there is no natural way to extend the methods of [4] to obtain stronger lower bounds for the size of constant-depth proofs of $\text{Count}_q$ from $\text{Count}_p$ as seems possible with our methods.

A $(p, q, l, M)$-generic system over $V$ formed by $q$-decision trees $T_1, \ldots, T_M$ has the property that every possible branch appears $0 \mod p$-times in the trees. Riis [19] conjectured (for $p, q$ different primes) that if $T_1, \ldots, T_M$ is any collection of $q$-decision trees over $V$, where $|V| \not\equiv 0 \mod q$, having the property

(a) every branch $F \in \bigcup_{i \leq M} \text{br}(T_i)$ appears $0 \mod p$-times,

(b) $M = 0 \mod p$.

The conjecture implies the non-existence of $(p, q, \lfloor |V|/q \rfloor, M)$-generic systems and thus would yield (by the remarks above) an exponential lower bound in the main theorem. On the other hand, the proof of Lemma 3.10 from Lemma 4.6 can be modified to prove the conjecture for collections of trees of a constant height.

Acknowledgement

We thank Jui-Lin Lee and Carlos Parra for comments on the manuscript of this paper.

References


Paul Beame  
Department of Computer Science and Engineering  
University of Washington  
Seattle  
Washington 98195  
U.S.A.  
E-mail: beame@cs.washington.edu

Russell Impagliazzo  
Department of Computer Science  
University of California at San Diego  
La Jolla  
California 92093  
U.S.A.  
E-mail: russell@cs.ucsd.edu

Jan Krajíček and Pavel Pudlák  
Mathematical Institute  
Academy of Sciences  
Zitná 25  
115 67 Prague  
Czech Republic  
E-mail: krajicek@earn.cvut.cz  
E-mail: pudlak@earn.cvut.cz

Toniann Pitassi  
Department of Computer Science  
University of Pittsburgh  
Pittsburgh  
Pennsylvania 15260  
U.S.A.  
E-mail: toni@cs.pitt.edu