Lower bounds on Hilbert's Nullstellensatz and propositional proofs

Paul Beame[*]
Dept. of Computer Science and Engineering
University of Washington
beame@cs.washington.edu

Russell Impagliazzo [†]
Dept. of Computer Science
University of California at San Diego
russell@cs.ucsd.edu

Jan Krajíček [†]
Mathematical Institute
Academy of Sciences, Praha
krajicek@earn.cvut.cz

Toniann Pitassi[§]
Dept. of Computer Science and Mathematics
University of Pittsburgh
toni@cs.pitt.edu

Pavel Pudlák[†]
Mathematical Institute
Academy of Sciences, Praha
pudlak@earn.cvut.cz

Abstract

The *weak form of the 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 iff 1 is in the ideal generated by the polynomials $Q_i(\bar{x})$. We shall prove a lower bound on the degrees of polynomials $P_i(\bar{x})$ such that $\sum_i P_i(\bar{x})Q_i(\bar{x}) = 1$.

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 $Count_q^N$. Ajtai [3] proved recently that, whenever p, q are two different primes, the propositional formulas $Count_q^{qn+1}$ do not have polynomial size, constant-depth Frege proofs from instances of $Count_p^m$, $m \not\equiv 0 \pmod p$. We give a new proof of this theorem based on the lower bound for the Hilbert's Nullstellensatz. Furthermore our technique enables us to extend the independence results for counting principles to *composite* numbers p and q. This results in an exact characterization of when $Count_q$ can be proven efficiently from $Count_p$, for all p and q.

Introduction

The problem of solvability of a system of algebraic equations

$$Q_i(\bar{x}) = 0, \quad i \in I, \tag{1}$$

in a fixed finite field $\mathbf{F}$ is one of the most natural $\mathcal{NP}$-complete problems. If we look for solutions in the algebraic closure of $\mathbf{F}$ 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 $\mathbf{F}$, iff there exist polynomials $P_i(\bar{x})$ from $\mathbf{F}[\bar{x}]$ such that

$$\sum_i P_i(\bar{x})Q_i(\bar{x}) = 1. \tag{2}$$

We can also use this to test solvability in the finite field $\mathbf{F}$ itself by adding the equations $\{x_j^{|\mathbf{F}|} - x_j = 0\}$ to (1) and we will usually assume that these equations have been included in (1).

The $\mathcal{NP}$-completeness of the solvability problem for (1) in $\mathbf{F}$ holds even if the degrees of the Q_i's are bounded by a constant. Suppose there existed P_i's 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_i's, we could construct these polynomials in polynomial time. Thus $\mathcal{P} \neq \mathcal{NP}$ implies that there are instances such that the degree of P_i's cannot be bounded by a constant. But the assump-

[*] supported by NSF grants CCR-8858799 and CCR-9303017

[†] partially supported by grant # 93025 of the joint *US-Czechoslovak Science and Technology Program*

[§] supported by an NSF postdoctoral fellowship and a UC presidential fellowship

tion $\mathcal{P} \neq \mathcal{NP}$ does not point to any concrete system of equations which requires P_i's of nonconstant degree.

The question of the degree of the polynomials P_i in (2) has also been studied in the context of the 'effective Nullstellensatz' of Brownawell and others ([8, 14, 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 $\mathbf{F}$ there is a matching lower bound example as well [8]. However, when using the Nullstellensatz to determine if solutions exist in $\mathbf{F}$ by extending the system of polynomials as described above, the degrees of the P_i for this example becomes constant.

In this paper we prove the first nonconstant lower bound on the degree of polynomials P_i in (2) for such an extended system. We prove this bound for equations which represent the counting principles mentioned in the abstract. Since we are interested in modular counting with an arbitrary modulus we also prove our results for some rings which are not fields. Our proof relies on a repeated use of Ramsey's theorem, so the bounds we obtain are extremely slowly growing functions of the number of variables. (There is a much simpler lower bound proof for the case of the real field which gives a much larger lower bound.)

Our results on the degrees of the witnessing polynomials P_i imply an improved lower bound for the lengths of the proofs of these counting principles in a class of propositional proof systems called constant-depth Frege proofs.

The interest in constant-depth Frege proofs is twofold. Cook and Reckhow [10] showed that the problem of whether there is a proof system in which all propositional tautologies have proofs of size polynomial in the size of the formula is equivalent to the question of whether $\mathcal{NP}$ equals co-$\mathcal{NP}$. Thus, a research program of proving lower bounds for more and more powerful propositional proof systems parallels the approach taken by boolean complexity of trying to prove lower bounds on the circuit size for restricted classes of circuits. Constant-depth Frege systems (which include Resolution as a special case) are the strongest proof systems for which a nontrivial lower bound is known. The second source of interest is the connection between the complexity of constant-depth Frege proofs and the inherent power of the systems $I\Delta_0(R)$ and $S_2(R)$ of bounded arithmetic (see [16] for details.) This connection shows that to demonstrate the unprovability of a principle in $I\Delta_0(R)$ or $S_2(R)$, it suffices to prove that it has no constant-depth Frege proofs of a certain size.

In a series of results, it has been shown that any constant-depth Frege proof of the *pigeonhole principle* (PHP) requires exponential size [12, 1, 15, 6, 17, 19]. These results were extended to show that even with the pigeonhole principle as an additional axiom constant-depth Frege systems require exponential size to prove the $Count_2$ tautologies [2, 7, 20]. (We use the notation $Count_q$ for the generic version of $Count_q^m$ where $m \not\equiv 0 \pmod q$.) Ajtai [3] studied the relationships among the various modular counting principles and showed that if p, q are different primes then $Count_q$ does not admit polynomial size, constant-depth Frege proofs from instances of $Count_p$. Riis [20] also considered the same problem.

In this paper we give a new proof of this lower bound which also applies when p and q are not primes. We also give upper bounds that result in an exact characterization of when polynomial-size constant-depth Frege proofs exist. The lower bound proof consists of two parts. The first part is similar to the proof strategy of [6, 17, 19] 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 involves the lower bound for Hilbert's Nullstellensatz.

1 Proof systems and counting principles

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

A *Frege system* [10] 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. 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.

Definition 1.1 *Let $N \geq r \geq 2$ and let V be a set of cardinality N. $[V]^r$ denotes the set of r-element subsets of V.*

Formula $Count_r^N$ is formed from atoms x_e, $e \in [V]^r$ and it is the formula:

$$\bigvee_{v \in V} \bigwedge_{v \in e} \neg x_e \quad \vee \quad \bigvee_{e \perp f} (x_e \wedge x_f)$$

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

Denote by $Count_{r,i}$ the set of formulas $Count_r^N$ for $N \equiv i \pmod r$. For R a set of pairs $\langle r, i \rangle$ such that $0 < i < r$, $Count_R$ denotes the union of the sets of formulas $Count_{r,i}$ such that $\langle r, i \rangle \in R$. Finally, $Count_r$ denotes the set of all formulas $Count_r^N$, for $N \not\equiv 0 \pmod r$; equivalently, $Count_r$ is $Count_R$ for $R = \{\langle r, i \rangle \mid 0 < i < r\}$.

In the introduction we observed that if r divides s then are polynomial size constant-depth Frege proofs of $Count_r$ from $Count_s$.

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

This was solved in the negative by Ajtai [3] for the case when r, s are two different primes. The following theorem is a strengthened version of Ajtai's result that gives a complete characterization of this problem. Let symbol (a, b) denote the *greatest common divisor* of a and b.

Theorem 1.2 *Let $q \geq 2$ and $0 < i < q$. Let R be a set of pairs of integers $\langle p, j \rangle$ such that $0 < j < p$. Then there are constant-depth, polynomial size Frege proofs of formulas from $Count_{q,i}$ from instances of $Count_R$ if and only if there is a $\langle p, j \rangle \in R$ such that all prime divisors of $\frac{p}{(p,j)}$ also divide $\frac{q}{(q,i)}$.*

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 is an infinite set of $N \not\equiv 0 \pmod q$ such that there are no constant-depth, polynomial size Frege proofs of $Count_q^N$ from instances of $Count_p$. In particular, this holds for all sufficiently large N such that $(p, \frac{q}{(q,N)}) = 1$.*

On the other hand, if all prime factors of q also divide p then for every infinite set of $N \not\equiv 0 \pmod q$ there are constant-depth, polynomial size Frege proofs of $Count_q^N$ 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 $N \equiv q/r \pmod q$ we have $(p, \frac{q}{(q,N)}) = (p, r) = 1$. It follows that for all j, $0 < j < p$, $(\frac{p}{(p.j)}, \frac{q}{(q,N)}) = 1$. Thus for each j, $0 < j < p$, there is some prime factor of $\frac{p}{(p,j)}$ that is not a factor of $\frac{q}{(q,N)}$. Applying Theorem 1.2 we obtain our desired result.

For the second part, assume that all prime factors of q also divide p. For $N \not\equiv 0 \pmod q$, let s be any prime factor of $\frac{q}{(q,N)}$. By assumption s also divides p. Thus for $j = p/s$ all prime factors of $\frac{p}{(p,j)}$ also divide $\frac{q}{(q,N)}$ and applying Theorem 1.2 we are done. □

The bulk of our arguments are concerned with our extension of Ajtai's lower bounds but we deal with the upper bounds first. These are extensions of upper bounds due to Riis [20]. The informal arguments given below can be easily formalized by polynomial size, constant-depth Frege proofs.

Lemma 1.4 *Assume that $r \geq 2$, $0 < i < r$, and k is a positive integer. Then*

(a) *there are polynomial size constant-depth Frege proofs of $Count_{r,i}$ from instances of $Count_{rk,ik}$.*

(b) *if $ik \not\equiv 0 \pmod r$, there are polynomial size constant-depth Frege proofs of $Count_{r,i}$ from instances of $Count_{r,ik \bmod r}$.*

Proof: Suppose that we have an r-partition of N, $N \equiv i \pmod r$. 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. □

The proof of the following lemma is more interesting but too long to include here. The main idea in its proof, which we borrow from a similar construction due to Riis [20], is to use a property of a set N (existence of an rk-partition of N) to obtain a property of the set M of k-element subsets of N (existence of an r-partition) and to use the fact that the size of M modulo r depends on the size of N in a nice way.

Lemma 1.5 *Let $r \geq 2$ and k be a positive integer. There are polynomial size constant-depth Frege proofs of $Count_{rk,k}$ from instances of $Count_{r,1}$.*

Corollary 1.6 *Let $p \geq 2$ and $0 < j < p$.*

(a) *There are polynomial size constant-depth Frege proofs of $Count_{p,j}$ from instances of $Count_{\frac{p}{(p,j)},1}$, and*

(b) *There are polynomial size constant-depth Frege proofs of $Count_{\frac{p}{(p,j)},1}$ from instances of $Count_{p,j}$.*

Proof: For (a) part, we start with instances of $Count_{\frac{p}{(p,j)},1}$. By Lemma 1.5 we obtain $Count_{p,(p,j)}$. By definition there are integers k, l such that $(p,j) = kj + lp$. Applying Lemma 1.4 part (b) with this value of k, $r = p$ and $i = j$ we obtain $Count_{p,j}$. For (b) part, starting with instances of $Count_{p,j}$ and applying Lemma 1.4 part (a) with $r = \frac{p}{(p,j)}$, $i = j$, and $k = (p,j)$ we obtain $Count_{\frac{p}{(p,j)},\frac{j}{(p,j)}}$. Then, applying Lemma 1.4 part (b) with $r = \frac{p}{(p,j)}$, $i = 1$, and $k = \frac{j}{(p,j)}$ we obtain $Count_{\frac{p}{(p,j)},1}$. □

Lemma 1.7 *Let $p, q \geq 2$. If all prime factors of p also divide q then there are polynomial size constant-depth Frege proofs of $Count_{q,1}$ from instances of $Count_{p,1}$.*

Proof: If all prime factors of p also divide q then there is some integer k such that $p \mid q^k$. By splitting the classes of size p into pieces of size q^k, we can obtain polynomial size constant-depth Frege proofs of $Count_{q^k,1}$ from instances of $Count_{p,1}$. We will show how to obtain $Count_{q,1}$ from instances of $Count_{q^k,1}$.

Suppose that there is some q-partition E of $N \equiv 1 \pmod{q}$. The *Fermat-Euler Theorem* implies that $N^m \equiv 1 \pmod{q^k}$ where m is any multiple of $\phi(q^k)$. Fix some such multiple m with $m \geq k$. Define a q^m-partition E' of N^m by taking the the m-th Cartesian power of E, i.e. the classes of E' have the form: $e_1 \times \ldots \times e_m$ where e_i are classes of E. 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 $Count_{q^k,1}$. □

We are now ready to prove the forward direction of Theorem 1.2. Assuming that all prime divisors of $\frac{p}{(p,j)}$ divide $\frac{q}{(q,i)}$, we want to show that we can prove $Count_{q,i}$ from $Count_{p,j}$. By Corollary 1.6 part (b), we can prove $Count_{\frac{p}{(p,j)},1}$ from $Count_{p,j}$. Then by Lemma 1.7, we can prove $Count_{\frac{q}{(q,i)},1}$ from $Count_{\frac{p}{(p,j)},1}$. Lastly, by Corollary 1.6 part (a), we can derive $Count_{q,i}$ from $Count_{\frac{q}{(q,i)},1}$.

2 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 [20] (the name *generic systems* was introduced in [20] 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 $N \not\equiv 0 \pmod{q}$. 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.

2.1 q-decision trees and k-evaluations

Definition 2.1 *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, (usually L will be the set $\{0,1\}$) and which satisfies two conditions:*

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.*

   We shall identify the edges by their labels.

2. *Let T^e be the proper subtree of T whose root is the end-point of edge e outgoing from the root. Then T^e is a q-decision tree over $V \setminus \{e\}$.*

The height *of tree T is the maximum number of edges on a path from the root to a leaf. Let* $\text{br}(T)$ *denote the set of branches of T and* $\text{br}_i(T)$ *denote the set of branches of T with leaf label $i \in L$. We shall also identify a branch with the conjunction of the variables indexed by its edges.*

We would like to prove the main theorem by finding an assignment making all formulas from $Count_R$ true but $Count_q^N$ false, demonstrating thus that $Count_q$ cannot be proved from $Count_R$. This is, of course, impossible as all formulas in $Count_q$ are tautologies. So we must find another way of 'evaluating' formulas which would permit such argument. A possible evaluation is to assign to each formula ϕ a usual decision tree deciding the value of the formula and consider the set of all branches of the tree for ϕ with leaf label 1. Thus ϕ is a tautology iff 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. To each formula ϕ we assign a q-decision tree T_ϕ. Intuitively, a formula ϕ is approximately true iff all branches in T_ϕ have leaf label 1. Furthermore, condition (1) of Definition 2.1 corresponds to the first conjunct in $\neg Count_q^N$ (i.e. every $v \in V$ is in some class e of the partition violating $Count_q^N$) 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_ϕ have leaf label 1 for $\phi = \neg Count_q^N$ and thus get a notion of evaluation in which $Count_q^N$ is false. The following is a reinterpretation of the definition in [17].

Definition 2.2 *Let Γ be a set of formulas formed from atoms of $Count_q^N$ (we shall not repeat this condition as we do not consider formulas formed from other atoms) and closed under subformulas.*

A k-evaluation *of Γ is a map $T : \phi \mapsto T_\phi$, defined on Γ such that:*

1. *the set of leaf labels L of T_ϕ is $\{0, 1\}$.*
2. *T_ϕ 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, $\mathrm{br}_1(T_0) = \emptyset$ and $\mathrm{br}_1(T_1) = \mathrm{br}(T_1)$*
4. *T_{x_e} is a q-decision tree having the property that every non-leaf vertex of T_{x_e} is labelled by some $v \in e$, and $\mathrm{br}_1(T_{x_e})$ is the unique branch of length 1 consisting of the edge labelled by e.*
5. *$\mathrm{br}(T_{\neg\phi}) = \mathrm{br}(T_\phi)$ and $\mathrm{br}_1(T_{\neg\phi}) = \mathrm{br}(T_\phi) \setminus \mathrm{br}_1(T_\phi)$*
6. *if $\phi = \bigvee_i \phi_i$ then T_ϕ refines and represents $\bigvee H$ where $H := \bigcup_i \mathrm{br}_1(T_{\phi_i})$, i.e.:*

   (a) $\mathrm{br}(T_\phi)$ refines H, *i.e. every $E \in \mathrm{br}(T_\phi)$ either is a superset of some $F \in H$ or $E \perp F$ for all $F \in H$ where $E \perp F$ iff $e \perp f$ for some $e \in E$ and $f \in F$*

   (b) $\mathrm{br}_1(T_\phi) = \{E \in \mathrm{br}(T_\phi) \mid \exists F \in H, F \subseteq E\}$.

The following lemma is completely analogous to a lemma from [17] treating the case of PHP_n in place of $Count_q^N$ and we shall not reprove it here.

Lemma 2.3 *Assume Γ 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* $\mathrm{br}_1(T_\phi) = \mathrm{br}(T_\phi)$
2. *equality* $\mathrm{br}_1(T_\phi) = \mathrm{br}(T_\phi)$ *is preserved by any sound inference rule of size at most s; for example if $3kq < N$,* $\mathrm{br}_1(T_\alpha) = \mathrm{br}(T_\alpha)$ *and* $\mathrm{br}_1(T_{\neg\alpha\vee\beta}) = \mathrm{br}(T_{\neg\alpha\vee\beta})$ *implies* $\mathrm{br}_1(T_\beta) = \mathrm{br}(T_\beta)$
3. *if $\phi = Count_q^N \in \Gamma$ then* $\mathrm{br}_1(T_\phi) = \emptyset \neq \mathrm{br}(T_\phi)$

Assume now that Π is a short constant-depth Frege proof of $Count_q^N$ from some instances of $Count_R$; that is from some formulas

$$Count_p^M(y_g/\psi_g) \ ,$$

where $M \equiv j \pmod p$, $0 < j < p$ and $\langle p, j\rangle \in R$, $g \in [M]^p$ and ψ_g are formulas in atoms x_e of $Count_q^N$. Other than the $Count_R$ axioms, there is a constant s such that all axiom schemes and inference rules used in Π are of size at most s. Suppose that there is a k-evaluation T of all subformulas in Π such that

(1) $kqs < N$ and

(2) for all instances ϕ of a $Count_R$ axiom in Π, $\mathrm{br}_1(T_\phi) = \mathrm{br}(T_\phi)$.

That would give a contradiction with Lemma 2.3 as $\mathrm{br}_1(T_\phi) = \mathrm{br}(T_\phi)$ would hold for all axioms and all inferences in π but not for the final formula.

This motivates the structure of our argument. More precisely, we first show that if Π is a short proof of $Count_q^N$ then there is a k-evaluation of all subformulas of Π satisfying (1) but not (2) and thus derive in Theorem 2.5 that a particular combinatorial object (a generic system) must exist. We then derive a contradiction in Lemma 2.10 by showing that this combinatorial object cannot exist.

Definition 2.4 *A* (p, q, ℓ, M)-generic system over V *is a collection of q-decision trees over V, T_i, $i \leq M$, with leaf labels from $[M]^p$ such that:*

(1) each T_i has height at most ℓ;

(2) each branch in T_i with leaf label g has $i \in g$;

(3) for all $g \in [M]^p$, $\mathrm{br}_g(T_i) = \mathrm{br}_g(T_j)$ for all $i, j \in g$.

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 $\{x_e\}$, $e \in [V]^q$. (Say that E and F are *compatible* if $E \not\perp F$. 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, $(p, q, 1, M)$-generic systems exist (take any system of height 1 q-decision trees $\{T_i\}$ and a p-partition π 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 2.5 *Let $0 < i < q$ and R be a set of pairs of integers $\langle p, j\rangle$ 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 sufficiently large N, $N \equiv i \pmod q$, there are depth d size $N^{k(N)}$ Frege proofs of $Count_q^N$ from instances of $Count_R$.*

Then for sufficiently large $N \equiv i \pmod q$ there is a $\langle p, j\rangle \in R$, an $N' \leq N$, $N' = N^{\Omega(1)}$ and $N' \equiv i \pmod q$, an $\ell = O(k(N))$ and a number $M = N^{O(\ell)} = (N')^{O(\ell)}$, $M \equiv j \pmod p$, such that there exists a (p, q, ℓ, M)-generic system over a set V of size N'.

For the proof we need to use *restrictions.*

Definition 2.6 *A* restriction *ρ is given by a set of disjoint q-element sets on a domain V. It determines an assignment to the variables over V as follows:*

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

We shall denote by V^ρ the set of vertices not covered by q-element sets of ρ. Denote by F^ρ the effect of applying the assignment to a Boolean formula F; we use similar notion for the effect of ρ on proofs, functions etc. Also, we can apply a restriction ρ to a q-decision tree T over V in the obvious way to obtain a q-decision tree T^ρ over V^ρ. Note that after applying ρ, we get essentially the same modulus q counting principle on V^ρ and that $|V^\rho| \equiv |V| \bmod q$.

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

Lemma 2.7 *Let Π 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 ρ such that:*

(1) $N' = |V^\rho| \geq N^{3/c_d}$,

(2) $N' \equiv N (\bmod q)$,

(3) Π restricted by ρ is a proof (over N') of $Count_q^{N'}$ from instances of $Count_R$, and

(4) there exists an k'-evaluation, T, of the subformulas in Π' for $k' \leq c_d k(N)$.

The above lemma is proven by inductively generating k_ℓ-evaluations (for some appropriate sequence of values k_ℓ) for the set of subformulas appearing at the ℓ bottom levels of every formula in Π. It is trivial to do this for the literals and constants on the leaves of the formulas. This k_ℓ-evaluation is extended to a $k_{\ell+1}$-evaluation of the set of subformulas in Π one level higher by applying a Håstad-style switching lemma [13] on an appropriate class of restrictions. The restrictions that are needed, together with the corresponding switching lemma, are stated below. A complete proof of this switching lemma and a sketch of how to apply it to obtain the above lemma can be found in [5].

Definition 2.8 *Define the set of restrictions M_m^V to be the set of all partial q-partitions ρ of V which cover all but $qm + j$ nodes of V where $j = |V| \bmod q$.*

Lemma 2.9 *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 $n = \lfloor |V|/q \rfloor$ and let $u = m/n$. If ρ is a restriction chosen uniformly at random from M_m^V, then with probability at least $1 - (4e^q r^{1/q} u^q n^{q-1/q})^s$, there is a q-decision tree over V^ρ of depth at most s refining and representing $\bigvee_i \mathrm{br}_1(T_i^\rho)$.*

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 $\phi^\rho \to T_\phi^\rho$ is a k-evaluation of S^ρ over V^ρ. With this observation and the appropriate choice of parameters we obtain Lemma 2.7 from repeated applications of Lemma 2.9.

After applying Lemma 2.7, we are left with a new proof of $Count_q^{N'}$ from instances of $Count_R$: $Count_{p_i}^{M_i}(\psi)$ such that $M_i \equiv j_i \pmod{p_i}$ where $\langle p_i, j_i\rangle \in R$, together with a k'-evaluation, T, for all subformulas in the proof, where $N' = N^\epsilon$, $\epsilon > 0$, and $M_i \leq (N')^{k'/\epsilon}$ for $k' = O(k(N))$.

Lemma 2.3 implies (as $\mathrm{br}_1(T_\eta) \neq \mathrm{br}(T_\eta)$ holds for the final formula η of Π') that $\mathrm{br}_0(T_\phi) \neq \emptyset$ for at least one $Count_R$ axiom ϕ in Π'. Take one such ϕ, and any $G \in \mathrm{br}_0(T_\phi)$ and restrict Π' further by G. In particular, ϕ is reduced to some instance $\phi^G = Count_p^M(\psi)$ for which $\mathrm{br}(T_{\phi^G}) = \mathrm{br}_0(T_{\phi^G})$. 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 $\mathrm{br}(T_\phi) = \mathrm{br}_0(T_\phi)$ for the axiom $\phi = Count_p^M(\psi)$.

For each formula ψ_g in $Count_p^M(\psi)$, let T_{ψ_g} denote the image of ψ_g under T.

Claim. *For all incompatible $g, h \in [M]^p$, if E is a branch of T_{ψ_g} labelled by 1, and F is a branch of T_{ψ_h} labelled by 1, then $E \perp F$.*

Proof: Suppose the claim fails, and let $g, h \in [M]^p$ be incompatible, E, F be branches labelled 1 in T_{ψ_g} and $T\psi_h$, respectively, where E is compatible with F. Apply the restriction $\sigma = E \cup F$ to the entire proof to obtain a new proof of $Count_q$ on the smaller universe of size $N' - q|\sigma|$. It is not too hard to show that after applying σ to $\phi = Count_p^M(\psi)$, the tree, T_{ϕ^σ} representing ϕ^σ will have all 1 labels.

This is, however, impossible as already $\mathrm{br}_1(T_\phi) = \emptyset$. □

We still need to apply another restriction to Π'; this restriction will be used to obtain q-decision trees, T_i for each $i \in [M]$, whose leaf labels are in $[M]^p$. We obtain the q-decision trees T_i, $i \in [M]$, by applying Lemma 2.9 with $r = k'$, $s = O(k'/\epsilon)$ and $m = (N')^\delta$ for some constant $\delta > 0$, to the sets of trees $F_i = \{T_{\psi_g} \mid i \in g\}$. Let ρ' be the restriction constructed and $N'' = (N')^\delta$ be the size of the resulting domain. A direct application of the switching lemma to the F_i, $i \in [M]$, yields q-decision trees over N'', T_i, with 0-1 labels, that refine and represent $\bigvee_{g,\ i \in g} \mathrm{br}_1(T_{\psi_g}^{\rho'})$. We will modify the 1 labels as follows. If σ is a branch of T_i labelled by 1, then by the switching lemma, it extends a branch labelled 1 of some $T_{\psi_g}^{\rho'}$, $i \in g$. Furthermore, it extends a branch of exactly one $T_{\psi_g}^{\rho'}$ by the above claim. Therefore, we will label σ by the unique $g \in [M]^p$ such that σ extends a branch of $T_{\psi_g}^{\rho'}$. By construction, all labels of T_i will now be either 0 or will be labelled by some $g \in [M]^p$, $i \in g$.

We now show that for all i, no branch of T_i is labelled by 0. (This argument is very similar to the proof of the above claim.) Assume there exists a branch σ of T_i with label 0. By the switching lemma, this implies that σ is not compatible with any branch labelled 1 in $\bigcup_{g,\ i \in g} \mathrm{br}_1(T_{\psi_g}^{\rho'})$. Applying the restriction σ to the entire proof, we obtain a new proof of $Count_q$ on the smaller universe of size $N'' - q|\sigma|$. But it is not hard to show (by similar reasoning as in the above claim) that $\mathrm{br}_1(T_\phi^{\rho'\sigma}) = \mathrm{br}(T_\phi^{\rho'\sigma})$ since i is not contained in any partition of $[M]$. But again, this contradicts the fact that $\mathrm{br}_1(T_\phi) = \emptyset$.

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 $\ell = pk$, where k was the height of the original trees T_i.)

We are now ready to complete the proof of Theorem 2.5. We will show that the set of q-partition decision trees T_i', $i \in [M]$ form a (p, q, ℓ, M)-generic system for $\ell = O(k(N))$. 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 $N'' = N^{\epsilon\delta} = N^{\Omega(1)}$. 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 $\mathrm{br}(T_i)$ with leaf label g, is equal to the set of branches in $\mathrm{br}(T_j)$ with leaf label g. Thus, the decision trees T_i', $i \in [M]$ is a (p, q, ℓ, M)-generic system as required by Theorem 2.5. □

We shall apply Theorem 2.5 only for $k(N)$ a constant because we are able to prove the next lemma only for ℓ a constant independent of N. It is our main combinatorial result and its proof occupies sections 3 and 4.

Lemma 2.10 *Let $\ell > 0$. For all sufficiently large N such that $(p^{(\ell-1)^2+1}, q) \mid N$, there does not exist a (p, q, ℓ, M)-generic system over N, with $M \not\equiv 0 \pmod p$.*

Proof of Theorem 1.2:
Let $q' = \frac{q}{(q,i)}$ and $R' = \{\frac{p}{(p,j)} \mid \langle p, j \rangle \in R\}$. By Corollary 1.6 it suffices to show that there are polynomial size constant-depth Frege proofs of $Count_{q',1}$ from instances of $Count_{p',1}$, $p' \in R'$, 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 1.7.

Now suppose instead that for every $p' \in R'$ there is some prime factor of p' that does not divide q'. Let k be a constant. Assume that, for sufficiently large $N \equiv 1 \pmod{q'}$, there are depth d size N^k Frege proofs of $Count_{q'}^N$ from instances of $Count_{p',1}$. Then by Theorem 2.5 for some $p' \in R'$ there is a (p', q', ℓ, M)-generic system over $N' \equiv 1 \pmod{q'}$ for constant ℓ, $M \equiv 1 \pmod{p'}$, and N' arbitrarily large.

Let r be the prime factor of p' that does not divide q'. From the (p', q', ℓ, M)-generic system over N' we can derive an (r, q', ℓ, M)-generic system over N' by splitting sets of size p' in a canonical way into p'/r sets each of size r. More precisely, for each $i \in M$ and

$g \in [M]^{p'}$ for each leaf of T_i labelled by g we replace the label g by the canonical r-subset of g that contains i. Since r and q' are relatively prime, the existence of this (r, q', ℓ, M)-generic system over N' contradicts Lemma 2.10 and we have completed the proof of the theorem. □

3 Counting principles and systems of polynomial equations

It is possible to express the propositional formula $Count_q^N$ by a system of polynomial equations, $Q_i(\overline{x}) = 0$, over the ring $\mathbf{Z}_p$. We will first describe these polynomial equations and then show that Lemma 2.10 follows from a nonconstant lower bound on the degree of any linear combination of the Q_i's that equals 1 modulo p.

Definition 3.1 *Assume that* $N \not\equiv 0 \pmod{q}$. *An* (N, q)-polynomial system *expressing the modulo q counting principle is the following system of polynomial equations in variables x_e, $e \in [V]^q$, $|V| = N$:*

$$(v) \quad (\sum_{v \in e} x_e) - 1 = 0$$

one for each $v \in V$, and

$$(e, f) \qquad x_e \cdot x_f = 0$$

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

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

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 nonzero for $v \in e$ and the equation (v) then implies that the unique nonzero 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 3.2 *Let $\mathbf{F}$ be any field. There are polynomials P_v, $v \in V$, and $P_{e,f}$, $e, f \in [V]^q$, $e \perp f$ from the ring $\mathbf{F}[\overline{x}_e]$ 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 $\mathbf{F}[\overline{x}_e]$.

We note that, although $x_e^{|\mathbf{F}|} - x_e$ 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 $x_e \cdot Q_v - \sum_{v \in e', e' \neq e} Q_{e,e'}$ for any $v \in e$. Thus a non-constant degree lower bound on the P_v and $P_{e,f}$ 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 $\mathbf{Z}_p$ of counting modulo p, where we do not assume that p is prime. Henceforth a *linear combination* L means a polynomial of the form:

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

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

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

Lemma 3.3 *Let T be a q-decision tree of height ℓ and assume $\ell q < N$. Then the polynomial*

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

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

(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 $\ell = 1$

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

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

Assume $\ell > 1$ and let v be the label of the root of T. Then:

$$u_T = \sum_{E \in \mathrm{br}(T)} x_E \quad - 1 = \sum_{v \in e} x_e(\sum_{F \in \mathrm{br}(T^e)} x_F) \quad - 1 \,.$$

By the induction hypothesis there are linear combinations L_e of degree at most $\ell - 2$ such that:

$$L_e = \sum_{F \in \mathrm{br}(T^e)} x_F \quad - 1$$

and so

$$u_T = \sum_{v \in e} x_e(L_e + 1) \quad - 1 = \sum_{v \in e} x_e \cdot L_e \; + \; \sum_{v \in e} x_e \quad - 1 \,.$$

The quantity $\sum_{v\in e} x_e \cdot L_e$ is a linear combination of degree at most $\ell - 1$ and the remaining quantity is just the polynomial Q_v; hence u_t is also a linear combination of degree at most $\ell - 1$. □

The next lemma is an important property of generic systems.

Lemma 3.4 *Let T_i, $i \in [M]$ be a (p, q, ℓ, M)-generic system. Then in the ring $\mathbf{Z}_p[\bar{x}_e]$ we have*

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

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

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

□

The lemma below follows from the previous two lemmas.

Lemma 3.5 *If there is a (p, q, ℓ, M)-generic system T_i, $i \in [M]$, such that $\ell \cdot q < N$, then there is a linear combination L of degree at most $\ell - 1$ such that $L + M = 0$ in the ring $\mathbf{Z}_p[\bar{x}_e]$.*

Proof: By Lemma 3.3, we can write the sum

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

as:

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

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

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

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

The following is the main technical result of this paper.

Lemma 3.6 (main) *Let d be a constant, let N be sufficiently large and suppose that N satisfies $(p^{d^2+1}, q) \mid N$ and $M \not\equiv 0 \pmod p$. Then every linear combination L such that $L = M$ in $\mathbf{Z}_p[\bar{x}_e]$ 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 2.10 from it.

Proof of Lemma 2.10 from Lemma 3.6

Assume that for some constant ℓ there exists a (p, q, ℓ, M)-generic system T_i, $i \in [M]$, $M \not\equiv 0 \pmod p$, over some N such that $(p^{(\ell-1)^2+1}, q) \mid N$.

If $N \equiv 0 \pmod q$ then there is some perfect q-partition π 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 $M \not\equiv 0 \pmod p$.

Suppose now that $N \not\equiv 0 \pmod q$. By Lemma 3.5 the existence of this (p, q, ℓ, M)-generic system over N implies the existence of a linear combination, L, of degree at most $\ell - 1$ such that $L = -M$ in the ring of polynomials $\mathbf{Z}_p[\bar{x}_e]$. But this contradicts Lemma 3.6 because $-M$ is not congruent to 0 mod p. □

4 Proof of the lower bound on the degree of the polynomials

In this section we prove main Lemma 3.6. It is an immediate corollary of the following lemma.

Lemma 4.1 *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, $v \in V$, $|V| = N$, are of degree $\leq d$, then there exists a 0-1 assignment a such that for every $e \perp f$*

$$Q_{e,f}(a) = 0, \tag{3}$$

and

$$\sum_{v\in V} P_v(a) Q_v(a) = 0, \tag{4}$$

where we count in $\mathbf{Z}_p$.

The rest of the section is devoted to outlining the proof of this lemma. We first need some preliminary concepts.

A 0-1 assignment corresponds to a set of q-element sets (those for which $x_e(a) = 1$). For this to satisfy (3) these q-element sets must be disjoint, so without loss of generality we restrict ourselves to such assignments. Also, 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 $x_{e_j}(a) = 0$ 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 nonzero coefficient.

Fix d. All systems $\mathcal{P}$ we consider have degree d (i.e. all P_v's have degree $\leq d$.) For a monomial $\Xi = x_{e_1} \ldots x_{e_c}$, we denote by $supp(\Xi) = e_1 \cup \ldots \cup e_c$ the *support* of Ξ.

The system of polynomials $\mathcal{P} = \{P_v\}_{v \in V}$ is determined by a sequence of coefficients of the form $a_\gamma \in Z_p$ for $\gamma = (v, \Xi)$, where $v \in V$, $\Xi = x_{e_1} \ldots x_{e_c}$ for $c \leq d$, $e_1, \ldots, e_c \in [V]^q$, and $\{v\}, e_1, \ldots, e_c$ are disjoint. We shall assume that the q-element sets are ordered by their least elements, i.e. $\min e_1 < \ldots < \min e_c$. Thus we get a 1-1 correspondence between (v, Ξ) with Ξ of degree c and $qc+1$-tuples $\langle v, e_{1,1}, .., e_{1,q}, \ldots, e_{c,1}, .., e_{c,q} \rangle$ of distinct elements from V such that $e_{i,1} < e_{j,1}$ for $i < j$ and $e_{i,j} < e_{i,k}$ for $j < k$.

Definition 4.2
(1) $type(v, \Xi)$ denotes the isomorphism type of the structure

$$(\{v\} \cup supp(\Xi);\ v, e_1, \ldots, e_c, \leq);$$

(2) k-$type(v, \Xi)$ over V denotes the isomorphism type of the structure

$$(\{v\} \cup supp(\Xi);\ v, e_1, \ldots, e_c, \leq, R_0, \ldots, R_{k-1}),$$

where $R_0, \ldots, R_{k-1}$ are unary predicates defined by

$$R_i(x) \Leftrightarrow_{df} x = v_j \text{ and } i = j \bmod k$$

where $V = \{v_1, \ldots, v_N\}$, $v_1 < \ldots < v_N$.

Definition 4.3
(1) $\mathcal{P}$ is symmetric, *if for every type T the a_γ are the same for all $\gamma = (v, \Xi)$ of type T.*
(2) $\mathcal{P}$ is k-symmetric over V, *if for every k-type T over V the a_γ are the same for all $\gamma = (v, \Xi)$ of k-type T over V.*

Lemma 4.4 *If $\mathcal{P}$ is p^k-symmetric over V, $|V| = N$, $(p^{k+1}, q) \mid N$ and $N \geq p^{k+1}q$, then there exists an assignment a such that*

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

Proof: Take $t \geq 0$ such that

$$N - qt \geq 0 \text{ and } N - qt \equiv 0 \pmod{p^{k+1}}.$$

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 Ξ 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 Ξ 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

$$\sum_{v \leq N - qt,\ v \equiv i \pmod{p^k}} P_v(a) Q_v(a) \equiv 0 \pmod{p},$$

whence the lemma follows. □

We note that in the argument above we actually used relatively little 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 employ restrictions as described in section 2. When we try to apply Ramsey theorem to get a p^k-symmetric system of polynomials after applying some restriction ρ we encounter two problems: First, a single application of Ramsey's theorem will only allow us 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 large degrees and then relaxing it as the smaller degrees are handled and being careful about the exact details of constructing the restriction.

Suppose that $|V| = N$, $V' = V^\rho = \{v_1, \ldots, v_m\}$, $v_1 < \ldots < v_m$ for some restriction ρ 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 argue that for any p^s-type T' over V' of degree d', the contribution (modulo p) to the coefficient of any (v, Ξ_1) 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^r-symmetry of the larger degree terms over V we need only count the number of $(v, \Xi_1\Xi_2)$ of p^r-type T over V that contribute to the coefficient of a given (v, Ξ_1) of p^s-type T' over V'.

We note that for a given (v, Ξ_1), the number of Ξ_2 such that $\Xi_2^\rho = 1$ and $(v, \Xi_1\Xi_2)$ is of a given p^r-type over V is affected by the relative order of $\{v\} \cup$ *supp*(Ξ_1) among the elements of the q-sets in ρ. Thus, after we choose the set V' using Ramsey's theorem, we have to be careful, when we choose ρ such that $V' = V^\rho$, to consider how the q-element sets in ρ cover the nodes in $V \setminus V'$ that are between elements of V'; we say any such q-element set *interleaves* the set V'.

The next lemma whose proof is in the full paper formalizes the base case of the proof of Lemma 4.1.

Lemma 4.5 *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, $|V| = N$, there exists a restriction ρ such that*

1. *the monomials of degree d' in $\mathcal{P}$ in variables over the set V^ρ form a symmetric system;*
2. *$|\ V^\rho\ | = m$;*
3. *for every $v, v' \in V^\rho$, $v \equiv v' \pmod{k'}$;*
4. *if we let $U = \{u_1, \ldots, u_z\} \subseteq V \setminus V^\rho$ be the set of elements that are between elements of V^ρ then ρ 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^ρ and working upwards, or starting from the smallest element of V^ρ working downwards.*

Lemma 4.6 *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 $\mathcal{P}$ of degree d over V, $|V| = N$, if the monomials of $\mathcal{P}$ of degrees $d'+1, \ldots, d$ form a k-symmetric system of polynomials, then there is a restriction ρ such that $|V^\rho| = m$ and the monomials of degree $d', d'+1, \ldots, d$ of $\mathcal{P}^\rho$ form a kp^d-symmetric system over V^ρ.*

To prove this we need:

Lemma 4.7 *Let $0 \leq i_1, \ldots, i_l < k$ 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*

$$\begin{aligned} b_1 &\equiv i_1 \pmod k \\ &\vdots \\ b_l &\equiv i_l \pmod k. \end{aligned} \qquad (5)$$

Then the residue class modulo p of C is determined by $n \bmod kp^l$.

Proof of Lemma 4.6. By Lemma 4.5 there is an N_0 such that for any V, $|V| \geq N_0$ and $|V| = N \equiv m \pmod q$, and any $\mathcal{P}$ over V, there is a restriction ρ satisfying conclusions 1-4 of Lemma 4.5 with $k' = kp^d$. Choose this N_0 and ρ. By conclusion 1 of Lemma 4.5, the monomials of degree d' in $\mathcal{P}$ over V^ρ appear symmetrically. Thus, since the monomials of degree $> d'$ in $\mathcal{P}$ form a k-symmetric system it is sufficient to show that if $\mathcal{P}'$ is any k-symmetric system of monomials of degree at most d over V then $\mathcal{P}'^\rho$ is kp^d-symmetric over V^ρ.

Let $V' = V^\rho = \{v_1, \ldots, v_m\}$ and write $\rho = \rho_1\rho_2$ where ρ_1 is the portion of ρ that interleaves V'.

Consider first $\mathcal{P}'^{\rho_2}$ over V^{ρ_2}. Note that V^{ρ_2} is a consecutive sequence of elements of V. Thus, for any two (v, Ξ_1) and (v', Ξ_1') of k-type T over V^{ρ_2}, if $(v, \Xi_1\Xi_2)$ is of k-type T' over V of degree $> d'$ then so is $(v', \Xi_1'\Xi_2)$. Thus $\mathcal{P}'^{\rho_2}$ is k-symmetric over V^{ρ_2}.

It remains to see what happens with monomials after applying ρ_1. Since p'^{ρ_2} is k-symmetric over V^{ρ_2} and the elements of V^{ρ_2} are consecutive we can ignore the difference between $\mathcal{P}'^{\rho_2}$ and $\mathcal{P}'$ and between V^{ρ_2} and V. Thus we want to show that for an arbitrary k-symmetric system $\mathcal{P}'$ over V of degree at most d such that $V^{\rho_1} = V'$, $\mathcal{P}'^{\rho_1}$ is kp^d-symmetric over V'.

Claim. *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, Ξ_1) of kp^d-type T' over V', the number (modulo p) of Ξ_2 such that $(v, \Xi_1\Xi_2)$ has k-type T over V and $\Xi_2^\rho = 1$ is the same.*

Let $\{v\} \cup$ *supp*$(\Xi_1) = \{v_{i_0}, \ldots, v_{i_{qc}}\}$ in increasing order (c is the degree of Ξ_1). Consider possible monomials Ξ_2 of degree b such that $b + c \leq d$ and $\Xi_2{}^{\rho_1} = 1$. Each q-set fixed by ρ_1 is determined by a single representative $u \in [v_1, v_m] \setminus V'$. Thus consider the representatives $u_{j_1} < \ldots < u_{j_b}$ in $[v_1, v_m] \setminus V'$ which determine the monomial Ξ_2. Since Ξ_1 is fixed, by construction of ρ_1 the k-type of $(v, \Xi_1\Xi_2)$ over V depends only on two properties:

(A) the position of $u_{j_1}, \ldots, u_{j_b}$ w.r.t. $v_{i_0}, v_{i_1} \ldots, v_{i_{qc}}$ (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_t} is either always the order of the u_{j_t} or always the reverse);

(B) the residue classes of $\{u_{j_1}, u_{j_1}^{(1)}, u_{j_1}^{(2)}, \ldots, u_{j_1}^{(q-1)}\}$, $, \ldots, \{u_{j_b}, u_{j_b}^{(1)}, u_{j_b}^{(2)}, \ldots, u_{j_b}^{(q-1)}\}$ modulo k.

The *key observation for* (B) is that the residue classes of $u_{j_t}^{(1)}, u_{j_t}^{(2)}, \ldots, u_{j_t}^{(q-1)}$ are precisely determined by the residue class modulo k of u_{j_t} and the residue class modulo k of the number of vertices in V' less than u_{j_t}. This is because the residue classes of $u_{j_t}^{(1)}, u_{j_t}^{(2)}, \ldots, u_{j_t}^{(q-1)}$ are determined just by j_t modulo k and the difference between the residue classes of u_{j_t} and j_t depends on the number of vertices in V' less than u_{j_t}. (The difference is either positive or negative depending upon whether ρ_1 matches the elements $[v_1, v_m] \setminus V'$ above or below $[v_1, v_m]$.)

Since we only consider Ξ_1 such that $(v, \Xi_1\Xi_2)$ has some fixed k-type T over V, for each α, $0 \leq \alpha \leq qc$, we have fixed the indices $\beta, \ldots, \gamma$ such that $v_{i_\alpha} < u_{j_\beta} < \ldots < u_{j_\gamma} < v_{i_{\alpha+1}}$ where β and γ depend only on α. We shall handle each such interval $v_{i_\alpha}, v_{i_{\alpha+1}}$ separately and show that (modulo p) the number of choices of such $u_{j_\beta}, \ldots, u_{j_\gamma}$ between v_{i_α} 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_α and $i_{\alpha+1}$ modulo kp^d. This will be sufficient since the total number of choices of Ξ_2 is the product of the number of choices in each of these intervals and the kp^d-*type*(v, Ξ_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

$$\begin{aligned} u_{j_t} &\equiv k_t \pmod{k}; \qquad \text{and} \\ \ell_t &\equiv k'_t \pmod{k}, \end{aligned}$$

where ℓ_t is the index such that $v_{\ell_t} < u_{j_t} < v_{\ell_t+1}$.

We would like to apply Lemma 4.7 to the equations for u_{j_t} and ℓ_t above and argue that the number of solutions only depends on i_α and $i_{\alpha+1}$ modulo kp^d. We cannot do so immediately since it is possible, if $k'_t = k'_{t+1}$, that $\ell_t = \ell_{t+1}$ and Lemma 4.7 does not apply in this case. Instead we break up the cases into the possible partitions π of the interval $[\beta, \gamma]$ into intervals $[\mu_1, \nu_1], \ldots, [\mu_\xi, \nu_\xi]$ so that for t, t' in in the same interval $\ell_t = \ell_{t'}$ and for t, t' in different intervals $\ell_t \neq \ell_{t'}$. It is now sufficient to argue two things for each fixed partition π:

- The number of choices (modulo p) of the sequence ℓ_t consistent with the k-type T and the partition π depends only on i_α and $i_{\alpha+1}$ modulo kp^d.

- The number of choices (modulo p) of $u_{j_\beta}, \ldots, u_{j_\gamma}$, consistent with k-type T and a fixed sequence of ℓ_t that is consistent with T, is independent of the choice of (v, Ξ_1).

Given the fixed partition π, the sequence of ℓ_t for $t \in [\beta, \gamma]$ is precisely determined by $\ell_{\mu_1}, \ldots, \ell_{\mu_\xi}$ where $i_\alpha \leq \ell_{\mu_1} < \ldots < \ell_{\mu_\xi} < i_{\alpha+1}$ and $\ell_{\mu_i} \equiv k'_{\mu_i} \pmod{k}$ for $i = 1, \ldots, \xi$. By Lemma 4.7 the number of such solutions depends only on $i_{\alpha+1} - i_\alpha$ modulo kp^ξ which is determined by $i_{\alpha+1} - i_\alpha$ modulo kp^d since $\xi \leq d$.

Now consider the fixed sequence ℓ_t and its associated partition π. For each interval $[\mu_i, \nu_i] = [r, s]$ in π 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 $v_{\ell_r} + 1 \leq u_{j_r} < \ldots < u_{j_s} < v_{\ell_r+1}$ and $u_{j_t} \equiv k_t \pmod{k}$ for $t \in [r, s]$. By Lemma 4.7 the number of such choices modulo p depends only $v_{\ell_r+1} - (v_{\ell_r} + 1)$ modulo kp^{s-r+1}. Since all elements of V' are equivalent modulo kp^d and $d \geq s - r + 1$, $v_{\ell_r+1} - (v_{\ell_r} + 1)$ is always congruent to -1 modulo kp^{s-r+1} and thus the number of choices modulo p in each interval of π is independent of the choice of (v, Ξ_1). Therefore the number of choices modulo p of $u_{j_\beta}, \ldots, u_{j_\gamma}$ consistent with the sequence of ℓ_t and the k-type T is independent of the choice of (v, Ξ_1) as required.

Thus we have proved the claim and hence the system $\mathcal{P}'^\rho$ is kp^d-symmetric, which finishes the proof of Lemma 4.6. □

Proof of Lemma 4.1. Let $m \geq p^{d^2+1}q$ and $(p^{d^2+1}, q)|m$. Choose V, $|V| = N \equiv m \pmod{q}$, large enough to apply Lemma 4.6 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 constructed have $|V^\rho| = m$. Choose any such system $\mathcal{P}$ and note that Lemma 4.6 implies that there is a restriction ρ with $|V^\rho| = m$ and $\mathcal{P}^\rho$ is p^{d^2}-symmetric over V^ρ. By Lemma 4.4 there is an assignment a on which $\mathcal{P}^\rho$ vanishes. Combining ρ with a, we get an assignment on which the $\mathcal{P}$ vanishes. Finally, recall that all Q_{ef}'s vanish too, due to the fact that ρ and a are partial partitions. □

5 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 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 nonconstant lower bound. Note that an exponential lower bound on the size of constant-depth Frege proofs of $Count_q^N$ from $Count_p$ instances would follow from an improvement of the degree lower bound to n^ϵ, for some $\epsilon > 0$.

We note that, at the same time as Ajtai's proof of the separation of the counting principles appeared [3], Riis [20] also announced a more detailed separation but its proof was incomplete. More recently [21], he has developed another proof of the separation using a Ramsey theory argument but not the Nullstellensatz polynomials that we use here.

Finally, Edmonds, Impagliazzo, and Pitassi [11] have recently shown a lower bound of $\Omega(n^{1/4})$ on the degree of the witnessing polynomials for a different unsolvable system of polynomial equations.

References

[1] Ajtai, M. (1988) The complexity of the pigeonhole principle, in: *Proc. IEEE 29^{th} Annual Symp. on Foundation of Computer Science*, pp. 346-355.

[2] ——(1990) Parity and the pigeonhole principle, in: *Feasible Mathematics*, Eds. S.R.Buss and P.J.Scott, pp.1-24. Birkhaüser.

[3] ——(1993) The independence of the modulo p counting principles, preprint.

[4] ——(1993) Symmetric systems of linear equations modulo p, preprint.

[5] Beame, P. (1993) A switching lemma primer, preprint.

[6] Beame, P., Impagliazzo, R., Krajíček, J., Pitassi, T., Pudlák, P., and Woods, A. (1992) Exponential lower bounds for the pigeonhole principle, in: *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, pp.200-221. ACM Press.

[7] Beame, P., and Pitassi, T. (1993) An exponential separation between the matching principles and the pigeonhole principle, to appear in: *Annals of Pure and Applied Logic*. Preliminary version: University of Washington Technical Report, April 1993.

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

[9] Caniglia, L., Galligo, A., and Heintz, J. (1988) Some new effectivity bounds in computational geometry, in: *Proceedings 6th International Conference on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, Ed. T.Mora, pp. 131-151. Lecture Notes in Computer Science 357 (Springer Verlag, 1989).

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

[11] Edmonds, J., Impagliazzo, R., and Pitassi, T. (1994) personal communication.

[12] Haken, A. (1985) The intractability of resolution, *Theoretical Computer Science*, **39**:297-308.

[13] Håstad, J. (1987) *Computation limits of small depth circuits*. ACM dissertation award, 1986. MIT Press.

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

[15] Krajíček, J. (1991) Lower bounds to the size of constant-depth propositional proofs, *Journal of Symbolic Logic*, in print.

[16] ——(1993) Bounded arithmetic, propositional calculus and complexity theory, preliminary version of a monograph prepared for the *Cambridge University Press*.

[17] Krajíček, J.,Pudlák, P. and Woods, A. (1991) Exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle, submitted.

[18] Paris, J. B., and Wilkie, A. J. (1985) Counting problems in bounded arithmetic, in: *Methods in Mathematical Logic*, LNM 1130, pp.317-340.Springer.

[19] Pitassi, T., Beame, P., and Impagliazzo, R. (1993) Exponential lower bounds for the pigeonhole principle, in: *Computational Complexity*, **3**:97-308.

[20] Riis, S. (1993) Independence in bounded arithmetic, PhD. Thesis, Oxford University.

[21] Riis, S. (1994) *Count*(q) does not imply *Count*(p), preprint.