Bounds on the Total Coefficient Size of Nullstellensatz Proofs of the Pigeonhole Principle

Aaron Potechin
University of Chicago, IL, USA

Aaron Zhang
The Voleon Group, Berkeley, CA, USA

Abstract

We show that the minimum total coefficient size of a Nullstellensatz proof of the pigeonhole principle on $n+1$ pigeons and n holes is $2^{\Theta(n)}$. We also investigate the ordering principle and construct an explicit Nullstellensatz proof for the ordering principle on n elements with total coefficient size $2^n - n$.

2012 ACM Subject Classification Theory of computation → Proof complexity

Keywords and phrases Proof complexity, Nullstellensatz, pigeonhole principle, coefficient size

Digital Object Identifier 10.4230/LIPIcs.ICALP.2024.117

Category Track A: Algorithms, Complexity and Games

Related Version *Full Version*: `https://arxiv.org/abs/2205.03577` [29]

Funding *Aaron Potechin*: NSF grant CCF-2008920
Aaron Zhang: NDSEG Fellowship 9422254702

1 Introduction

Given a system $\{p_i = 0 : i \in [m]\}$ of m polynomial equations, a Nullstellensatz proof of infeasibility is an equality of the form $1 = \sum_{i=1}^{m} p_i q_i$ for some polynomials $\{q_i = 0 : i \in [m]\}$. Hilbert's Nullstellensatz[1] says that the Nullstellensatz proof system is complete over algebraically closed fields, i.e., a system of polynomial equations has no solutions over an algebraically closed field if and only if there is a Nullstellensatz proof of infeasibility. However, Hilbert's Nullstellensatz does not give any bounds on the degree or size needed for Nullstellensatz proofs.

The degree of Nullstellensatz proofs has been extensively studied. Grete Hermann showed a doubly exponential degree upper bound for the ideal membership problem [24] which implies the same upper bound for Nullstellensatz proofs. Several decades later, W. Dale Brownawell gave an exponential upper bound on the degree required for Nullstellensatz proofs over algebraically closed fields of characterisic zero [11]. A year later, János Kollár showed that this result holds for all algebraically closed fields [27].

For specific problems, the degree of Nullstellensatz proofs can be analyzed using designs [14]. Using designs, Nullstellensatz degree lower bounds have been shown for many problems including the pigeonhole principle, the induction principle, the housesitting principle, and the mod m matching principles [6, 5, 15, 16, 12]. More recent work showed that there is a close connection between Nullstellensatz degree and reversible pebbling games [19] and that lower bounds on Nullstellensatz degree can be lifted to lower bounds on monotone span programs, monotone comparator circuits, and monotone switching networks [28].

[1] Technically, this is the weak form of Hilbert's Nullstellensatz. Hilbert's Nullstellensatz actually says that given polynomials $p_1, \ldots, p_m$ and another polynomial p, if $p(x) = 0$ for all x such that $p_i(x) = 0$ for each $i \in [m]$ then there exists a natural number r such that p^r is in the ideal generated by $p_1, \ldots, p_m$.

For analyzing the size of Nullstellensatz proofs (i.e., the number of monomials in the proof), a powerful technique is the size-degree relation shown by Russell Impagliazzo, Pavel Pudlák, and Jiří Sgall for polynomial calculus [25], which also holds for resolution proofs [7] and for sum of squares proofs with $\{0,1\}$ variables [3]. The size-degree relation says that if there is a size S polynomial calculus proof then there is a polynomial calculus proof of degree $O(\sqrt{n \log S})$. Thus, if we have an $\Omega(n)$ degree lower bound for polynomial calculus, this implies a $2^{\Omega(n)}$ size lower bound for polynomial calculus (which also holds for Nullstellensatz as Nullstellensatz is a weaker proof system). However, the size-degree relation does not give any size lower bound when the degree is $O(\sqrt{n})$, and we know of few other techniques for analyzing the size of Nullstellensatz proofs.

In this paper, instead of investigating the degree or size of Nullstellensatz proofs, we investigate the total coefficient size of Nullstellensatz proofs, i.e., the sum of the magnitudes of the coefficients of the monomials in the proof. Total coefficient size is a reasonably natural measure which is relatively unexplored (though as we discuss below, there has been considerable research on closely related measures such as unary Nullstellensatz size, unary Sherali-Adams size, and the total bit complexity of proofs [2, 1, 9, 20, 31]). There are several reasons why total coefficient size bounds in particular are interesting.

First, analyzing the total coefficient size of proofs may give insight into proof size in settings where we currently cannot prove size lower bounds. If we can prove a large total coefficient size lower bound, this shows that any proof must either have large size or involve large coefficients. Unless there is a reason to suspect that large coefficients are helpful for making the proof shorter, this gives considerable evidence for a lower bound on proof size.

Second, lower bounds on total coefficient size have some direct implications. As observed by [20], a total coefficient size lower bound for the stronger Sherali-Adams proof system implies a lower bound for the reversible resolution proof system which captures the Max-SAT resolution proof system (see [10]) for Max SAT. Similarly, [20] observes that a total coefficient size lower bound for Nullstellensatz implies a lower bound for the reversible resolution with terminals proof system, which is a weaker variant of reversible resolution.

Finally, investigating the total coefficient size of proofs gives insight into the following question. Are there natural examples where having fractional coefficients greatly reduces the total coefficient size needed for Nullstellensatz and/or Sherali-Adams proofs? We note that this question is not addressed by [20]. For example, [20] shows that there are n-variate CNF formulas F such that F can be refuted by constant width resolution proofs but any Sherali-Adams proof of F requires either exponentially many monomials or requires coefficients of exponential size (see Theorem 1 and the last paragraph of Section 1.1 in [20]). However, this does not rule out the existence of a proof where there are exponentially many monomials but the coefficient for each monomial is exponentially small so the total coefficient size is still small.

Proving total coefficient size lower bounds for a problem rules out this possibility. Conversely, if there is a natural example where the minimum proof size is large but the total coefficient size is small, that would be quite interesting.

1.1 Our results

In this paper, we show that the minimum total coefficient size of a Nullstellensatz proof of the pigeonhole principle is $2^{\Theta(n)}$. More precisely, we show the following bounds.

▶ **Theorem 1.** *For all $n \geq 1$, any Nullstellensatz proof of the pigeonhole principle with $n+1$ pigeons and n holes has total coefficient size $\Omega\left(n^{\frac{3}{4}}\left(\frac{2}{\sqrt{e}}\right)^n\right)$.*

We note that this lower bound also holds for the functional pigeonhole principle, where each pigeon must go to exactly one hole (instead of at least one hole).

▶ **Theorem 2.** *For all $n \geq 1$, there is a Nullstellensatz proof of the pigeonhole principle with $n+1$ pigeons and n holes with total coefficient size at most $2^{5(n+1)}$.*

▶ Remark 3. Note that Nullstellensatz size lower bounds do not imply total coefficient size lower bounds, because we could have a proof with many monomials but a small coefficient (absolute value less than 1) on each monomial. Indeed, in Appendix A, we show an example where the minimum total coefficient size of a Nullstellensatz proof is smaller than the minimum size of a Nullstellensatz proof. Thus, the exponential size polynomial calculus lower bounds for the pigeonhole principle from Razborov's $\Omega(n)$ degree lower bound for polynomial calculus [30] and the size-degree relation [25] do not imply total coefficient size lower bounds for the pigeonhole principle.

In addition, we investigate the total coefficient size of Nullstellensatz proofs of the ordering principle in Appendix C. We show the following upper bound by constructing an explicit Nullstellensatz proof.

▶ **Theorem 4.** *For all $n \geq 3$, there is a Nullstellensatz proof of the ordering principle on n elements with size and total coefficient size $2^n - n$. This upper bound is tight for $n \leq 5$.*

In the full version of this paper [29], we also discuss total coefficient size for the Sherali-Adams and sum of squares proof systems. We observe that even though resolution is a dynamic proof system, the $O(n^3)$ size resolution proof of the ordering principle found by Gunnar Stålmark [32] can be captured by a one line sum of squares proof with small size and coefficients.

1.2 Comparison with related work

Like previous resolution, polynomial calculus, and Sherali-Adams lower bounds for the pigeonhole principle (e.g. [22, 30, 18]), our analysis is inspired by the idea that if we only look at a small number of pigeons, we cannot detect a problem. That said, our analysis differs considerably from previous analyses of the pigeonhole principle as we need to bound the value of a linear program by constructing a dual certificate (see Proposition 12). To construct this dual certificate, we need to assign a value to every possible assignment of the variables, so we need to consider all n pigeons at once which requires a different analysis.

In terms of the overall framework, the work which is most similar to ours is that of De Rezende, Potechin, and Risse [31] which shows a total coefficient size Sherali-Adams lower bound for showing that a random graph does not contain a large clique. Like our paper, [31] constructs a dual certificate which assigns a value to every possible assignment of the variables. That said, [31] uses different techniques to construct and analyze their dual certificate. In particular, while the construction in [31] is inspired by the pseudo-calibration technique used to prove SoS lower bounds for planted clique [4] and the analysis heavily uses the fact that the graph is random, our construction and analysis is combinatorial and takes advantage of symmetry.

Another work which is closely related to ours is that of Göös et al. [20]. [20] analyzes the size of unary Nullstellensatz and Sherali-Adams proofs, which is equivalent to analyzing the total coefficient size of Nullstellensatz and Sherali-Adams proofs with the added restriction that all coefficients are integers. The authors show that there are deep connections between unary Nullstellensatz, unary Sherali-Adams, resolution, and total NP search problems (TFNP). In particular, they prove the following results (among others) which show that there are considerable advantages to having the restriction that all coefficients are integers.

1. Resolution is not polynomially simulated by unary Sherali-Adams, and reversible resolution is not polynomially simulated by Nullstellensatz. Since unary Sherali-Adams can simulate reversible resolution, this implies that resolution is not simulated by reversible resolution.
2. Roughly speaking, unary Nullstellensatz corresponds to the TFNP class PPAD which corresponds to the principle that every directed graph with an unbalanced node (i.e., a node whose indegree is not equal to its outdegree) must have another unbalanced node. Similarly, unary Sherali-Adams corresponds to the TFNP class PPADS which corresponds to the principle that every directed graph with a postitively unbalanced node (outdegree exceeds indegree) must have a negatively unbalanced node (indegree exceeds outdegree).
3. There is a reversible resolution refutation of a CNF F if and only if there is both a resolution refutation of F and a unary Sherali-Adams refutation of F. Similarly, there is a reversible resolution with terminals refutation of a CNF F if and only if there is both a resolution refutation of F and a unary Nullstellensatz refutation of F.

In this paper, we show that there are also advantages to allowing fractional coefficients. Proving a total coefficient size lower bound when fractional coefficients are allowed removes the possibility of having a proof with many monomials but a small total coefficient size. In addition, allowing fractional coefficients gives us a linear program for minimum total coefficient size which can be analyzed directly. As a result, while [20] needs several steps to show their separations, we show our bounds directly.

Finally, a natural alternative to analyzing the size or total coefficient size of proofs is to analyze the bit complexity of proofs. One way to prove a lower bound on the bit complexity of a proof is to show an exponentially larger lower bound on the total coefficient size of the proof. Hakoniemi [23] uses this approach to give an example where there is a polynomial size sum of squares proof of degree 2 but every sum of squares proof requires doubly exponential total coefficient size and thus exponential bit complexity.

While it is generally hard to lower bound the bit complexity of a proof without lower bounding the proof size or total coefficient size, this has been done for the binary value principle which says that a number written in binary with no minus sign must be non-negative. More precisely, if $x_1, \ldots, x_n \in \{0, 1\}$ then we cannot have that $1 + x_1 + 2x_2 + \ldots + 2^{n-1}x_n = 0$. By considering the primes $p \in [1, 2^n]$ and showing that the proof must involve a coefficient which is divisible by all such primes, [1] and [2] show bit complexity lower bounds for powerful proof systems, namely polynomial calculus with extensions and the ideal proof system (see [21]) where the latter bound is conditional on the Shub-Smale hypothesis. This technique is powerful but is specialized to this problem and is very different from our techniques.

2 Nullstellensatz total coefficient size

We start by defining total coefficient size for Nullstellensatz proofs and describing a linear program for finding the minimum total coefficient size of a Nullstellensatz proof. In this paper, we only consider problems on Boolean variables, so we give definitions which are specialized for this setting.

▶ **Definition 5.** *For each Boolean variable x_i, we define the twin variable $\bar{x}_i$ to be $\bar{x}_i = 1 - x_i$.*

▶ **Definition 6.** *Given Boolean variables $x_1, \ldots, x_N$, we define a monomial to be a product of the form $\left(\prod_{i \in S} x_i\right)\left(\prod_{j \in T} \bar{x}_j\right)$ for some disjoint subsets S, T of $[N]$.*

▶ **Definition 7.** *Given a polynomial f on Boolean variables $x_1, \dots, x_N$, we define the total coefficient size $T(f)$ of f to be the minimum sum of the magnitudes of coefficients when we decompose f into monomials. For example, if $f(x_1, x_2) = 1 - x_1 - x_2 + 2x_1x_2$, then $T(f) = 2$ as we can write $f = \bar{x}_1\bar{x}_2 + x_1x_2 = (1 - x_1)(1 - x_2) + x_1x_2$.*

We will use the following terminology:

▶ **Definition 8.** *Given a system $\{p_i = 0 : i \in [m]\}$ of polynomial equations, we call each of the p_i an axiom. We say that a polynomial W is a weakening of the axiom p_i if $W = rp_i$ for some monomial r.*

We now define Nullstellensatz proofs and their total coefficient size.

▶ **Definition 9.** *Given a system $\{p_i = 0 : i \in [m]\}$ of polynomial equations on Boolean variables $x_1, \dots, x_N$, a Nullstellensatz proof of infeasibility is an equality of the form*

$$1 = \sum_{i=1}^{m} p_i q_i + \sum_{j=1}^{N} (x_j^2 - x_j) g_j + \sum_{j=1}^{N} (x_j + \bar{x}_j - 1) h_j$$

for some polynomials $\{q_i : i \in [m]\}$, $\{g_j : j \in [N]\}$, and $\{h_j : j \in [N]\}$. We define the total coefficient size of such a Nullstellensatz proof to be $\sum_{i=1}^{m} T(q_i)$.

▶ Remark 10. We do not include the total coefficient size of p_i, g_j, or h_j in the total coefficient size of the proof as we want to focus on the complexity of the proof as opposed to the complexity of the axioms and manipulating the Boolean variables. That said, in this paper we only consider systems of polynomial equations where each p_i is a monomial, so this choice does not matter: in this setting $T(p_i) = 1$ for all i, and it is both possible and optimal to take $g_j = 0$ for all j.[2] In terms of weakenings, in this setting a Nullstellensatz proof is an equality

$$1 = \sum_{W} c_W W,$$

where W ranges over all possible weakenings of axioms and $c_W \in \mathbb{R}$. The total coefficient size of a Nullstellensatz proof is $\sum_W |c_W|$.

The minimum total coefficient size of a Nullstellensatz proof can be found using a linear program. To illustrate this, we now give an example.

▶ **Example 11.** Consider the following system of equations on two variables x_1, x_2:

$$1 - x_1 = 0$$
$$1 - x_2 = 0$$
$$x_1x_2 = 0$$

Given these axioms, the possible weakenings W (modulo the Boolean axioms) are $1 - x_1$, $(1 - x_1)x_2$, $(1 - x_1)(1 - x_2)$, $1 - x_2$, $x_1(1 - x_2)$, and x_1x_2.

[2] In other words, we can assume without loss of generality that all terms in a Nullstellensatz proof have degree at most 1 in each variable. If an axiom contains a variable x_i, there is no point in multiplying the axiom by x_i or $\bar{x}_i = (1 - x_i)$, because $x_i^2 = x_i$ and $x_i(1 - x_i) = 0$ modulo the Boolean axioms. The reasoning is similar in the case that an axiom contains a variable $\bar{x}_i = (1 - x_i)$.

To find a Nullstellensatz proof with minimum total coefficient size, we write a linear program with a variable c_W for each weakening W. We also have a variable b_W for each weakening W representing the absolute value of c_W with the constraints $b_W - c_W \geq 0$ and $b_W + c_W \geq 0$. The objective is to minimize $\sum_W b_W$.

To ensure that $\sum_W c_W W = 1$, we have a constraint for each of the 4 possible assignments of values to the variables. For example, one possible assignment is $x_1 = 0, x_2 = 0$. We ensure that $\sum_W c_W W$ evaluates to 1 on this assignment by having the constraint

$$c_{(1-x_1)} + c_{(1-x_2)} + c_{(1-x_1)(1-x_2)} = 1,$$

because the weakenings $1 - x_1$, $1 - x_2$, and $(1 - x_1)(1 - x_2)$ evaluate to 1 on this assignment while the other weakenings evaluate to 0. The analogous constraints for the other 3 possible assignments of values to the variables are as follows:

1. $c_{(1-x_1)} + c_{(1-x_1)x_2} = 1$ for the assignment $x_1 = 0$, $x_2 = 1$
2. $c_{(1-x_2)} + c_{x_1(1-x_2)} = 1$ for the assignment $x_1 = 1$, $x_2 = 0$
3. $c_{x_1x_2} = 1$ for the assignment $x_1 = 1$, $x_2 = 1$.

The set of optimal solutions to this linear program is

$$\{c_{x_1x_2} = 1, c_{(1-x_1)} = c_{x_1(1-x_2)} = a, c_{(1-x_1)x_2} = c_{(1-x_2)} = 1-a, c_{(1-x_1)(1-x_2)} = 0 : a \in [0,1]\}$$

which corresponds to the equality

$$1 = x_1x_2 + a\left((1 - x_1) + x_1(1 - x_2)\right) + (1 - a)\left((1 - x_1)x_2 + (1 - x_2)\right).$$

In the same way as the above example, we can find the minimum total coefficient size of any system of equations with a linear program. In order to show a lower bound on total coefficient size, we will analyze the dual of this linear program. Because the primal has a constraint for each assignment of values to the variables $x \in \{0,1\}^N$, the dual has a variable for each assignment $x \in \{0,1\}^N$. We will let $D : \{0,1\}^N \to \mathbb{R}$ denote the dual.

We observe that D induces a linear map $\widehat{D}$ from polynomials to $\mathbb{R}$ in a natural way, by taking $\widehat{D}(f) = \sum_{x\in\{0,1\}^N} D(x)f(x)$. It turns out that the dual is equivalent to:

Maximize $\widehat{D}(1)$ subject to the constraint that for each weakening W, $|\widehat{D}(W)| \leq 1$.

Weak duality, which is what we need to prove lower bounds on total coefficient size, can be seen directly as follows.

▶ **Proposition 12.** *If $\widehat{D}$ is a linear map from polynomials to $\mathbb{R}$ such that $|\widehat{D}(W)| \leq 1$ for all weakenings W, then any Nullstellensatz proof has total coefficient size at least $\widehat{D}(1)$.*

Proof. Given a Nullstellensatz proof $1 = \sum_{i=1}^m p_iq_i$, applying $\widehat{D}$ to both sides gives $\widehat{D}(1) = \sum_{i=1}^m \widehat{D}(p_iq_i) \leq \sum_{i=1}^m T(q_i)$. The inequality holds because for any q_i, for any way of writing q_i in terms of monomials r as $q_i = \sum_r c_{ir}r$, we have $\widehat{D}(p_iq_i) = \sum_r c_{ir}\widehat{D}(rp_i) \leq \sum_r |c_{ir}|$ because rp_i is a weakening. ◀

3 Total coefficient size lower bound for the pigeonhole principle

In this section, we prove Theorem 1, our total coefficient size lower bound for the pigeonhole principle. We start by formally defining the pigeonhole principle.

▶ **Definition 13** (pigeonhole principle (PHP_n)). *Intuitively, the pigeonhole principle says that if $n+1$ pigeons are assigned to n holes, then some hole must have more than one pigeon. Formally, for $n \geq 1$, we define PHP_n to be the statement that the following system of axioms is infeasible:*

- *For each $i \in [n+1]$ and $j \in [n]$, we have a variable $x_{i,j}$ and the Boolean axiom $x_{i,j}^2 - x_{i,j} = 0$. $x_{i,j} = 1$ represents pigeon i being in hole j, and $x_{i,j} = 0$ represents pigeon i not being in hole j.*
- *For each $i \in [n+1]$, we have the axiom $\prod_{j=1}^{n} \bar{x}_{i,j} = 0$ representing the constraint that each pigeon must be in at least one hole.*
- *For each pair of distinct pigeons $i_1, i_2 \in [n+1]$ and each hole $j \in [n]$, we have the axiom $x_{i_1,j}x_{i_2,j} = 0$ representing the constraint that pigeons i_1 and i_2 cannot both be in hole j.*

We prove our lower bound on total coefficient size for PHP_n by constructing and analyzing a dual solution $D : \{0,1\}^{(n+1)n} \to \mathbb{R}$. In our dual solution, the only assignments of values to the variables $x \in \{0,1\}^{(n+1)n}$ for which $D(x) \neq 0$ are those where each pigeon goes to exactly one hole, i.e., for each pigeon i, exactly one of the $x_{i,j}$ is 1. As a result, Theorem 1 also applies to the functional pigeonhole principle. Note that there are n^{n+1} such assignments. In the rest of this section, when we refer to assignments or write a summation or expectation over assignments x, we refer specifically to these n^{n+1} assignments.

Recall that the dual constraints are

$$\widehat{D}(W) = \sum_{x \in \{0,1\}^N} D(x)W(x) \in [-1,1]$$

for all weakenings W. Note that since $D(x)$ is only nonzero for assignments x where each pigeon goes to exactly one hole, for any weakening W of an axiom of the form $\prod_{j=1}^{n} \bar{x}_{i,j} = 0$, we have $\widehat{D}(W) = 0$. Thus, it is sufficient to consider weakenings W of the axioms $x_{i_1,j}x_{i_2,j} = 0$.

For simplicity, in order to construct a dual solution, we first ignore the constraints $|\widehat{D}(W)| \leq 1$. Then, we obtain a dual solution by normalizing D, i.e., dividing D by $\max_W |\widehat{D}(W)|$. Thus, we can rewrite the objective value of the dual program as $\frac{\widehat{D}(1)}{\max_W |\widehat{D}(W)|}$. Letting $\mathbb{E}$ denote the expectation over a uniform assignment where each pigeon goes to exactly one hole, $\frac{\widehat{D}(1)}{\max_W |\widehat{D}(W)|} = \frac{\mathbb{E}(D)}{\max_W |\mathbb{E}(DW)|}$. Thus, it is sufficient to construct D and analyze $\mathbb{E}(D)$ and $\max_W |\mathbb{E}(DW)|$.

Before constructing and analyzing D, we provide some intuition for our construction. The idea is that if we consider a subset of n pigeons then D should behave like the indicator function for whether those n pigeons all go to different holes. More concretely, for any polynomial p which does not depend on some pigeon i (i.e., p does not contain $x_{i,j}$ or $\bar{x}_{i,j}$ for any $j \in [n]$), we want

$$\mathbb{E}(Dp) = \frac{n!}{n^n}\mathbb{E}(p \mid \text{all pigeons in } [n+1] \setminus \{i\} \text{ go to different holes})$$

Given this intuition, we now present our construction. Our dual solution D will be a linear combination of the following functions:

▶ **Definition 14** (functions J_S). *For each subset of pigeons $S \subsetneq [n+1]$ of size at most n, we define the function J_S that maps assignments to $\{0,1\}$ so that for each assignment x, $J_S(x) = 1$ if all pigeons in S are in different holes according to x and $J_S(x) = 0$ otherwise.*

Note that if $|S| = 0$ or $|S| = 1$, then J_S is the constant function 1. In general, the expectation of J_S over a uniform assignment is $\mathbb{E}(J_S) = \left(\prod_{k=1}^{|S|}(n+1-k)\right)/n^{|S|}$.

▶ **Definition 15** (dual solution D)**.** *Our dual solution D is:*

$$D = \sum_{S \subsetneq [n+1]} c_S J_S,$$

where the coefficients c_S are $c_S = \frac{(-1)^{n-|S|}(n-|S|)!}{n^{n-|S|}}$.

We will lower-bound the dual value $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ by computing $\mathbb{E}(D)$ and then upper-bounding $\max_W |\mathbb{E}(DW)|$. In both calculations, we will use the following key property of D which we introduced in our intuition for the construction:

▶ **Lemma 16.** *If p is a polynomial which does not depend on pigeon i (i.e., p does not contain any variables of the form $x_{i,j}$ or $\bar{x}_{i,j}$), then $\mathbb{E}(Dp) = \mathbb{E}(J_{[n+1]\setminus\{i\}}p)$.*

Proof. Without loss of generality, suppose p does not contain any variables of the form $x_{1,j}$ or $\bar{x}_{1,j}$. Let T be any subset of pigeons that does not contain pigeon 1 and that has size at most $n-1$. Observe that

$$\mathbb{E}(J_{T\cup\{1\}}p) = \frac{n-|T|}{n}\mathbb{E}(J_T p)$$

because when the pigeons in T go to different holes, the probability that pigeon 1 goes to a different hole is $\frac{n-|T|}{n}$, and p does not depend on the location of pigeon 1. Since

$$\begin{aligned} c_{T\cup\{1\}} &= \frac{(-1)^{n-1-|T|}(n-1-|T|)!}{n^{n-1-|T|}} \\ &= -\frac{n}{n-|T|}\cdot\frac{(-1)^{n-|T|}(n-|T|)!}{n^{n-|T|}} = -\frac{n}{n-|T|}c_T \end{aligned}$$

we have that for all $T \subsetneq \{2,\ldots,n+1\}$, $\mathbb{E}(c_{T\cup\{1\}}J_{T\cup\{1\}}p) + \mathbb{E}(c_T J_T p) = 0$. Thus, all terms in the sum $\mathbb{E}(Dp) = \sum_{S\subsetneq[n+1]}\mathbb{E}(c_S J_S p)$ cancel, except $J_{\{2,3,\ldots,n+1\}}$. Since $c_{\{2,3,\ldots,n+1\}} = 1$, we have that $\mathbb{E}(Dp) = \mathbb{E}(J_{\{2,3,\ldots,n+1\}}p)$, as needed. ◀

The value of $\mathbb{E}(D)$ follows immediately:

▶ **Corollary 17.**

$$\mathbb{E}(D) = \frac{n!}{n^n}.$$

Proof. Let $p = 1$. By Lemma 16, $\mathbb{E}(D) = \mathbb{E}(J_{\{2,\ldots,n+1\}}) = \frac{n!}{n^n}$. ◀

3.1 Upper bound on $\max_W |\mathbb{E}(DW)|$

We now upper bound $\max_W |\mathbb{E}(DW)|$. To do this, we introduce the following notation:

▶ **Definition 18** ($H_{W,i}$)**.** *Given a weakening W, we define a set of holes $H_{W,i} \subseteq [n]$ for each pigeon $i \in [n+1]$ so that $W(x) = 1$ if and only if each pigeon $i \in [n+1]$ is mapped to one of the holes in $H_{W,i}$. More precisely,*

- *If W contains terms x_{i,j_1} and x_{i,j_2} for distinct holes j_1, j_2, then $H_{W,i} = \emptyset$ (i.e., it is impossible that $W(x) = 1$ because pigeon i cannot go to both holes j_1 and j_2).*
- *If W contains exactly one term of the form $x_{i,j}$, then $H_{W,i} = \{j\}$. (i.e., for all x such that $W(x) = 1$, pigeon i goes to hole j).*

- *If W contains no terms of the form $x_{i,j}$, then $H_{W,i}$ is the subset of holes j such that W does not contain the term $\bar{x}_{i,j}$. (i.e., if W contains the term $\bar{x}_{i,j}$, then for all x such that $W(x) = 1$, pigeon i does not go to hole j.)*

The key property we will use to bound $\max_W |\mathbb{E}(DW)|$ follows immediately from Lemma 16:

▶ **Lemma 19.** *Let W be a weakening. If there exists some pigeon $i \in [n+1]$ such that $H_{W,i} = [n]$ (i.e., W does not contain any terms of the form $x_{i,j}$ or $\bar{x}_{i,j}$), then $\mathbb{E}(DW) = 0$.*

Proof. Without loss of generality, suppose W is a weakening of the axiom $x_{2,1}x_{3,1} = 0$ and $H_{W,1} = [n]$. By Lemma 16, $\mathbb{E}(DW) = \mathbb{E}(J_{\{2,\dots,n+1\}}W)$. However, $\mathbb{E}(J_{\{2,\dots,n+1\}}W) = 0$ because if $W(x) = 1$, then pigeons 2 and 3 must both go to hole 1. ◀

We now make the following definition and then state a corollary of Lemma 19.

▶ **Definition 20** (W_S^{flip})**.** *Let W be a weakening of the axiom $x_{i_1,j}x_{i_2,j} = 0$ for pigeons i_1, i_2 and hole j. Let $S \subseteq [n+1] \setminus \{i_1, i_2\}$. We define W_S^{flip}, which is also a weakening of the axiom $x_{i_1,j}x_{i_2,j} = 0$, as follows.*

- *For each pigeon $i_3 \in S$, we define W_S^{flip} so that $H_{W_S^{\text{flip}},i_3} = [n] \setminus H_{W,i_3}$.*
- *For each pigeon $i_3 \notin S$, we define W_S^{flip} so that $H_{W_S^{\text{flip}},i_3} = H_{W,i_3}$.*

Note: Technically, there are multiple possible weakenings W_S^{flip} which satisfy these properties (e.g. if $n = 2$, $W = x_{1,1}x_{2,1}x_{3,1}$, and $S = \{3\}$, then W_S^{flip} can be $x_{1,1}x_{2,1}\bar{x}_{3,1}$ or $x_{1,1}x_{2,1}x_{3,2}$ or even $x_{1,1}x_{2,1}\bar{x}_{3,1}x_{3,2}$, among others). We arbitrarily choose any such weakening W_S^{flip}.

In other words, W_S^{flip} is obtained from W by flipping the sets of holes that the pigeons in S can go to in order to make the weakening evaluate to 1.

▶ **Corollary 21.** *Let W be a weakening of the axiom $x_{i_1,j}x_{i_2,j} = 0$ for pigeons i_1, i_2 and hole j. Let $S \subseteq [n+1] \setminus \{i_1, i_2\}$. Then*

$$\mathbb{E}\left(DW_S^{\text{flip}}\right) = (-1)^{|S|} \cdot \mathbb{E}(DW).$$

Proof. It suffices to show that for $i_3 \in [n+1] \setminus \{i_1, i_2\}$, we have $\mathbb{E}\left(DW_{\{i_3\}}^{\text{flip}}\right) = -\mathbb{E}(DW)$. Indeed, let W' be a weakening such that $W'(x) = W(x) + W_{\{i_3\}}^{\text{flip}}(x)$ for all assignments x where each pigeon goes to exactly one hole. (For example, if $n = 2$, $W = x_{1,1}x_{2,1}x_{3,1}$, and $i_3 = 3$, then we can take $W_{\{3\}}^{\text{flip}}$ to be $x_{1,1}x_{2,1}x_{3,2}$, in which case $W' = x_{1,1}x_{2,1}$.) Then $\mathbb{E}(DW') = 0$ by Lemma 19 because $H_{W',i_3} = [n]$, so $\mathbb{E}\left(DW_{\{i_3\}}^{\text{flip}}\right) = -\mathbb{E}(DW)$. ◀

Using Corollary 21, we can bound $\max_W |\mathbb{E}(DW)|$ using Cauchy-Schwarz. We first show an approach that does not give a strong enough bound. We then show how to modify the approach to achieve a better bound.

▶ **Definition 22.** *Given functions F, G on the assignments mapping each pigeon to exactly one hole, we define $\langle F, G\rangle = \mathbb{E}(FG)$. We define $\|F\| = \sqrt{\langle F, F\rangle} = \sqrt{\mathbb{E}(F^2)}$.*

3.1.1 Unsuccessful approach to upper bound $\max_W |\mathbb{E}(DW)|$

Consider $\max_W |\mathbb{E}(DW)|$. By Lemma 21, it suffices to take the max over weakenings W such that, if W is a weakening of the axiom $x_{i_1,j}x_{i_2,j} = 0$, then for all pigeons $i_3 \in [n+1] \setminus \{i_1, i_2\}$, we have $|H_{W,i_3}| \leq \lfloor n/2 \rfloor$ (because if $|H_{W,i_3}| > \lfloor n/2 \rfloor$, we can flip H_{W,i_3} without changing $|\mathbb{E}(DW)|$). For any such W, we have

$$\|W\| = \sqrt{\mathbb{E}(W^2)} \leq \sqrt{\left(\frac{1}{n}\right)^2 \left(\frac{1}{2}\right)^{n-1}} = n^{-1}2^{-(n-1)/2}.$$

By Cauchy-Schwarz,

$$\begin{aligned} |\mathbb{E}(DW)| &\leq \|D\|\|W\| \\ &\leq \|D\|n^{-1}2^{-(n-1)/2}. \end{aligned}$$

Using the value of $\mathbb{E}(D)$ from Corollary 17, the dual value $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ is at least

$$\frac{n!}{n^n} \cdot \frac{n2^{(n-1)/2}}{\|D\|} = \widetilde{\Theta}\left(\left(\frac{e}{\sqrt{2}}\right)^{-n} \cdot \frac{1}{\|D\|}\right)$$

by Stirling's formula. Thus, in order to achieve an exponential lower bound on the dual value, we would need $1/\|D\| \geq \Omega(c^n)$ for some $c > e/\sqrt{2}$. However, this requirement is too strong, as we will show in Lemma 26 that $1/\|D\| = \widetilde{\Theta}\left((\sqrt{e})^n\right)$. Directly applying Cauchy-Schwarz results in too loose of a bound on $\max_W |\mathbb{E}(DW)|$, so we now modify our approach.

3.1.2 Successful approach to upper bound $\max_W |\mathbb{E}(DW)|$

▶ **Definition 23** ($W_{i_1,i_2}^{\{-1,0,1\}}$). *Let W be a weakening of the axiom $x_{i_1,j}x_{i_2,j} = 0$ for pigeons i_1, i_2 and hole j. We define a function $W_{i_1,i_2}^{\{-1,0,1\}}$ that maps assignments to $\{-1, 0, 1\}$. For an assignment x,*

- *If pigeons i_1 and i_2 do not both go to hole j, then $W_{i_1,i_2}^{\{-1,0,1\}}(x) = 0$.*
- *Otherwise, let $V(x) = |\{i_3 \in [n+1] \setminus \{i_1, i_2\} : \text{pigeon } i_3 \text{ does not go to } H_{W,i_3}\}|$. Then $W_{i_1,i_2}^{\{-1,0,1\}}(x) = (-1)^{V(x)}$.*

Note that $W_{i_1,i_2}^{\{-1,0,1\}}$ is a linear combination of the W_S^{flip}:

▶ **Lemma 24.** *Let W be a weakening of the axiom $x_{i_1,j}x_{i_2,j} = 0$ for pigeons i_1, i_2 and hole j. We have:*

$$W_{i_1,i_2}^{\{-1,0,1\}} = \sum_{S \subseteq [n+1] \setminus \{i_1,i_2\}} (-1)^{|S|} \cdot W_S^{\text{flip}}.$$

It follows that:

$$\mathbb{E}\left(DW_{i_1,i_2}^{\{-1,0,1\}}\right) = 2^{n-1} \cdot \mathbb{E}(DW).$$

Proof. To prove the first equation, consider any assignment x. If pigeons i_1 and i_2 do not both go to hole j, then both $W_{i_1,i_2}^{\{-1,0,1\}}$ and all the W_S^{flip} evaluate to 0 on x. Otherwise, exactly one of the $W_S^{\text{flip}}(x)$ equals 1, and for this choice of S we have $W_{i_1,i_2}^{\{-1,0,1\}}(x) = (-1)^{|S|}$.

The second equation follows because:

$$\begin{aligned}\mathbb{E}\left(DW_{i_1,i_2}^{\{-1,0,1\}}\right) &= \sum_{S\subseteq[n+1]\setminus\{i_1,i_2\}} (-1)^{|S|}\cdot\mathbb{E}\left(DW_S^{\text{flip}}\right)\\ &= \sum_{S\subseteq[n+1]\setminus\{i_1,i_2\}} (-1)^{|S|}(-1)^{|S|}\cdot\mathbb{E}(DW) && \text{(Corollary 21)}\\ &= 2^{n-1}\cdot\mathbb{E}(DW).\end{aligned}$$

◀

Using Lemma 24, we now improve on the approach to upper-bound $\max_W |\mathbb{E}(DW)|$ from section 3.1.1:

▶ **Lemma 25.** *The dual value* $\mathbb{E}(D)/\max_W|\mathbb{E}(DW)|$ *is at least* $\frac{n!}{n^n}\cdot\frac{n2^{n-1}}{\|D\|}$.

Proof. If W is a weakening of the axiom $x_{i_1,j}x_{i_2,j}=0$ for pigeons i_1,i_2 and hole j,

$$\begin{aligned}|\mathbb{E}(DW)| &= 2^{-(n-1)}\cdot\left|\mathbb{E}\left(DW_{i_1,i_2}^{\{-1,0,1\}}\right)\right| && \text{(Lemma 24)}\\ &\le 2^{-(n-1)}\cdot\|D\|\|W_{i_1,i_2}^{\{-1,0,1\}}\| && \text{(Cauchy-Schwarz)}\\ &= 2^{-(n-1)}\cdot\|D\|\sqrt{\mathbb{E}\left(\left(W_{i_1,i_2}^{\{-1,0,1\}}\right)^2\right)}\\ &= n^{-1}2^{-(n-1)}\cdot\|D\|.\end{aligned}$$

Using the value of $\mathbb{E}(D)$ from Corollary 17, the dual value $\mathbb{E}(D)/\max_W|\mathbb{E}(DW)|$ is at least $\frac{n!}{n^n}\cdot\frac{n2^{n-1}}{\|D\|}$. ◀

It only remains to compute $\|D\|$:

▶ **Lemma 26.**

$$\|D\|^2 = \frac{n!}{n^n}\cdot(n+1)!\cdot\sum_{c=0}^{n}\frac{(-1)^{n-c}}{n+1-c}\cdot\frac{1}{n^{n-c}c!}$$

Proof. Recall the definition of D (Definition 15):

$$\begin{aligned}D &= \sum_{S\subsetneq[n+1]} c_S J_S,\\ c_S &= \frac{(-1)^{n-|S|}(n-|S|)!}{n^{n-|S|}}.\end{aligned}$$

We compute $\|D\|^2=\mathbb{E}(D^2)$ as follows.

$$\mathbb{E}(D^2) = \sum_{S\subsetneq[n+1]}\sum_{T\subsetneq[n+1]} c_S c_T\mathbb{E}(J_S J_T).$$

Given $S,T\subsetneq[n+1]$, we have:

$$\begin{aligned}\mathbb{E}(J_S J_T) &= \mathbb{E}(J_S)\mathbb{E}(J_T\mid J_S=1)\\ &= \left(\left(\prod_{i=1}^{|S|}(n+1-i)\right)/n^{|S|}\right)\left(\left(\prod_{j=|S\cap T|+1}^{|T|}(n+1-j)\right)/n^{|T\setminus S|}\right).\end{aligned}$$

Therefore,

$$c_S c_T \mathbb{E}(J_S J_T) = \left(c_S \left(\prod_{i=1}^{|S|} (n+1-i) \right) / n^{|S|} \right) \left(c_T \left(\prod_{j=|S\cap T|+1}^{|T|} (n+1-j) \right) / n^{|T\setminus S|} \right).$$

Note that the product of $(-1)^{n-|S|}$ (from the c_S) and $(-1)^{n-|T|}$ (from the c_T) is $(-1)^{-|S|-|T|} = (-1)^{|S|-|T|}$, so the above equation becomes:

$$c_S c_T \mathbb{E}(J_S J_T) = (-1)^{|S|-|T|} \left(\frac{n!}{n^n} \right) \left(\frac{(n-|S\cap T|)!}{n^{n-|S\cap T|}} \right).$$

Now, we rearrange the sum for $\mathbb{E}(D^2)$ in the following way:

$$\begin{aligned} \mathbb{E}(D^2) &= \sum_{S\subsetneq[n+1]} \sum_{T\subsetneq[n+1]} c_S c_T \mathbb{E}(J_S J_T) \\ &= \frac{n!}{n^n} \sum_{c=0}^{n} \frac{(n-c)!}{n^{n-c}} \sum_{\substack{S,T\subsetneq[n+1], \\ |S\cap T|=c}} (-1)^{|S|-|T|}. \end{aligned}$$

To evaluate this expression, fix $c \le n$ and consider the inner sum. Consider the collection of tuples $\{(S,T) \mid S,T \subsetneq [n+1], |S\cap T| = c\}$. We can pair up most of these tuples in the following way. For each S, let m_S denote the minimum element in $[n+1]$ that is not in S (note that m_S is well defined because S cannot be $[n+1]$). We pair up the tuple (S,T) with the tuple $(S, T\triangle\{m_S\})$, where $\triangle$ denotes symmetric difference. The only tuples (S,T) that cannot be paired up in this way are those where $|S| = c$ and $T = [n+1]\setminus\{m_S\}$, because T cannot be $[n+1]$. There are $\binom{n+1}{c}$ unpaired tuples (S,T), and for each of these tuples, we have $(-1)^{|S|-|T|} = (-1)^{n-c}$. On the other hand, each pair $(S,T), (S, T\triangle\{m_S\})$ contributes 0 to the inner sum. Therefore, the inner sum equals $(-1)^{n-c}\binom{n+1}{c}$, and we have:

$$\begin{aligned} \mathbb{E}(D^2) &= \frac{n!}{n^n} \sum_{c=0}^{n} \frac{(-1)^{n-c}(n-c)!}{n^{n-c}} \binom{n+1}{c} \\ &= \frac{n!}{n^n} \sum_{c=0}^{n} \frac{(-1)^{n-c}(n-c)!}{n^{n-c}} \cdot \frac{(n+1)!}{c!(n+1-c)!} \\ &= \frac{n!}{n^n} \cdot (n+1)! \cdot \sum_{c=0}^{n} \frac{(-1)^{n-c}}{n+1-c} \cdot \frac{1}{n^{n-c}c!}. \end{aligned}$$

◀

▶ **Corollary 27.** $\mathbb{E}(D^2) \le \frac{(n+1)!}{n^n}$

Proof. Observe that the sum $\sum_{c=0}^{n} \frac{(-1)^{n-c}}{n+1-c} \cdot \frac{1}{n^{n-c}c!}$ is an alternating series where the magnitudes of the terms decrease as c decreases. The two largest magnitude terms are $\frac{1}{n!}$ and $-\frac{1}{2}\cdot\frac{1}{n!}$. Therefore, the sum is at most $\frac{1}{n!}$, and we conclude that $\mathbb{E}(D^2) \le \frac{n!}{n^n} \cdot \frac{(n+1)!}{n!} = \frac{(n+1)!}{n^n}$, as needed. ◀

We can now complete the proof of Theorem 1.

Proof of Theorem 1. By Lemma 25, any Nullstellensatz proof for PHP_n has total coefficient size at least $\frac{n!}{n^n} \cdot \frac{n2^{n-1}}{\|D\|}$. By Corollary 27, $\|D\| \le \sqrt{\frac{(n+1)!}{n^n}}$. Combining these results, any Nullstellensatz proof for PHP_n has total coefficient size at least

$$\frac{n!}{n^n} \cdot \frac{n2^{n-1}}{\sqrt{\frac{(n+1)!}{n^n}}} = \frac{n2^{n-1}}{\sqrt{(n+1)}} \cdot \frac{\sqrt{n!}}{n^{\frac{n}{2}}}$$
$$= \frac{n2^{n-1}}{\sqrt{n+1}} \sqrt{\frac{n!}{n^n}}$$

Using Stirling's approximation that $n!$ is approximately $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, $\sqrt{\frac{n!}{n^n}}$ is approximately $\sqrt[4]{2\pi n}\left(\frac{1}{\sqrt{e}}\right)^n$, and this expression is $\Omega\left(n^{\frac{3}{4}}\left(\frac{2}{\sqrt{e}}\right)^n\right)$, as needed. ◀

4 Open problems

Our work raises a number of open problems. First, while we showed that the minimum total coefficient size of a Nullstellensatz proof of the pigeonhole principle on $n+1$ pigeons and n holes is $2^{\Theta(n)}$, it is natural to ask what happens when we increase the number of pigeons.

1. If we increase the number of pigeons from $n+1$ to $n+2$ while still having n holes, our lower bound proof no longer applies. Can we prove a total coefficient size lower bound on Nullstellensatz when there are m pigeons where $m \geq n+2$? More ambitiously, how does the minimum total coefficient size of a proof depend on m and whether or not we add the axioms that pigeons can only go to one hole (i.e., considering the functional pigeonhole principle rather than the pigeonhole principle)?

Second, we are still far from understanding the total coefficient size of Nullstellensatz proofs of the ordering principle. In Appendix C we construct an explicit Nullstellensatz proof for the ordering principle on n elements with total coefficient size $2^n - n$, but we have no non-trivial lower bounds.

2. Can we prove superpolynomial lower bounds on the total coefficient size of Nullstellensatz proofs of the ordering principle and/or improve the $O(2^n)$ upper bound?

In the full version of this paper [29], we also discuss total coefficient size for the Sherali-Adams and sum of squares proof systems. Some questions regarding these related proof systems are:

3. Are there Sherali-Adams proofs for the ordering principle with polynomial total coefficient size? If so, this shows that the seemingly dynamic $O(n^3)$ size resolution proof of the ordering principle [32] can be captured by a one line Sherali-Adams proof. If not, this gives a natural example separating resolution proof size and the total coefficient size of Sherali-Adams proofs. We note that this separation has been shown by [20] for unary Sherali-Adams using pebbling principles.
4. Are there natural examples where the minimum total coefficient size is very different (either larger or smaller) than the minimum size for Nullstellensatz, Sherali-Adams, or sum of squares proofs?
5. Can the minimum total coefficient size of a strong proof system be used to lower bound the size of another proof system? For example, can resolution proof size be lower bounded by the minimum total coefficient size of a sum of squares proof, or can we find an example where there is a polynomial size resolution proof but any sum of squares proof has superpolynomial total coefficient size?

References

1. Yaroslav Alekseev. A Lower Bound for Polynomial Calculus with Extension Rule. In *Proceedings of the 36th Computational Complexity Conference*, CCC '21, Dagstuhl, DEU, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2021.21.
2. Yaroslav Alekseev, Dima Grigoriev, Edward A. Hirsch, and Iddo Tzameret. Semi-algebraic proofs, IPS lower bounds, and the tau-conjecture: can a natural number be negative? In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, pages 54–67, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3357713.3384245.
3. Albert Atserias and Tuomas Hakoniemi. Size-Degree Trade-Offs for Sums-of-Squares and Positivstellensatz Proofs. In *Proceedings of the 34th Computational Complexity Conference*, CCC '19, Dagstuhl, DEU, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2019.24.
4. Boaz Barak, Samuel Hopkins, Jonathan Kelner, Pravesh K Kothari, Ankur Moitra, and Aaron Potechin. A Nearly Tight Sum-of-Squares Lower Bound for the Planted Clique Problem. *SIAM Journal on Computing*, 48(2):687–735, 2019.
5. P. Beame, R. Impagliazzo, J. Krajicek, T. Pitassi, and P. Pudlak. Lower Bounds on Hilbert's Nullstellensatz and Propositional Proofs. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 794–806, 1994. doi:10.1109/SFCS.1994.365714.
6. Paul Beame, Stephen Cook, Jeff Edmonds, Russell Impagliazzo, and Toniann Pitassi. The Relative Complexity of NP Search Problems. *J. Comput. Syst. Sci.*, 57(1):3–19, August 1998. doi:10.1006/jcss.1998.1575.
7. Eli Ben-Sasson and Avi Wigderson. Short Proofs are Narrow—Resolution made Simple. *J. ACM*, 48(2):149–169, March 2001. doi:10.1145/375827.375835.
8. Olaf Beyersdorff, Nicola Galesi, and Massimo Lauria. A lower bound for the pigeonhole principle in tree-like resolution by asymmetric Prover–Delayer games. *Information Processing Letters*, 110(23):1074–1077, 2010. doi:10.1016/j.ipl.2010.09.007.
9. Ilario Bonacina and Maria Luisa Bonet. On the Strength of Sherali-Adams and Nullstellensatz as Propositional Proof Systems. In *Proceedings of the 37th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '22, New York, NY, USA, 2022. Association for Computing Machinery. doi:10.1145/3531130.3533344.
10. María Luisa Bonet, Jordi Levy, and Felip Manyà. Resolution for Max-SAT. *Artificial Intelligence*, 171(8):606–618, 2007. doi:10.1016/j.artint.2007.03.001.
11. W. Dale Brownawell. Bounds for the Degrees in the Nullstellensatz. *Annals of Mathematics*, 126(3):577–591, 1987. URL: http://www.jstor.org/stable/1971361.
12. S. Buss, R. Impagliazzo, J. Krajíček, P. Pudlák, A. A. Razborov, and J. Sgall. Proof Complexity in Algebraic Systems and Bounded Depth Frege Systems with Modular Counting. *Comput. Complex.*, 6(3):256–298, December 1997. doi:10.1007/BF01294258.
13. Sam Buss and Toniann Pitassi. Resolution and the weak pigeonhole principle. In Mogens Nielsen and Wolfgang Thomas, editors, *Computer Science Logic*, pages 149–156, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg.
14. Samuel R. Buss. Lower Bounds on Nullstellensatz Proofs via Designs. In Paul Beame and Samuel R. Buss, editors, *Proof Complexity and Feasible Arithmetics, Proceedings of a DIMACS Workshop, New Brunswick, New Jersey, USA, April 21-24, 1996*, volume 39 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 59–71. DIMACS/AMS, 1996. doi:10.1090/dimacs/039/04.
15. S.R. Buss and T. Pitassi. Good Degree Bounds on Nullstellensatz Refutations of the Induction Principle. In *Proceedings of Computational Complexity (Formerly Structure in Complexity Theory)*, pages 233–242, 1996. doi:10.1109/CCC.1996.507685.
16. Matthew Clegg, Jeffery Edmonds, and Russell Impagliazzo. Using the Groebner basis algorithm to find proofs of unsatisfiability. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pages 174–183, New York, NY, USA, 1996. Association for Computing Machinery. doi:10.1145/237814.237860.

17 S. Dantchev and S. Riis. Tree resolution proofs of the weak pigeon-hole principle. In *Proceedings 16th Annual IEEE Conference on Computational Complexity*, pages 69–75, 2001. doi:10.1109/CCC.2001.933873.

18 Stefan Dantchev, Barnaby Martin, and Mark Rhodes. Tight rank lower bounds for the Sherali–Adams proof system. *Theoretical Computer Science*, 410(21):2054–2063, 2009. doi:10.1016/j.tcs.2009.01.002.

19 Susanna F. de Rezende, Jakob Nordström, Or Meir, and Robert Robere. Nullstellensatz Size-Degree Trade-offs from Reversible Pebbling. In Amir Shpilka, editor, *34th Computational Complexity Conference (CCC 2019)*, volume 137 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 18:1–18:16, Dagstuhl, Germany, 2019. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2019.18.

20 M. Goos, A. Hollender, S. Jain, G. Maystre, W. Pires, R. Robere, and R. Tao. Separations in Proof Complexity and TFNP. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1150–1161, Los Alamitos, CA, USA, November 2022. IEEE Computer Society. doi:10.1109/FOCS54457.2022.00111.

21 Joshua A. Grochow and Toniann Pitassi. Circuit Complexity, Proof Complexity, and Polynomial Identity Testing: The Ideal Proof System. *J. ACM*, 65(6), November 2018. doi:10.1145/3230742.

22 Armin Haken. The intractability of resolution. *Theoretical Computer Science*, 39:297–308, 1985. Third Conference on Foundations of Software Technology and Theoretical Computer Science. doi:10.1016/0304-3975(85)90144-6.

23 Tuomas Hakoniemi. Monomial size vs. Bit-Complexity in Sums-of-Squares and Polynomial Calculus. In *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '21, New York, NY, USA, 2021. Association for Computing Machinery. doi:10.1109/LICS52264.2021.9470545.

24 Grete Hermann. Die Frage der endlich vielen Schritte in der Theorie der Polynomideale. *Mathematische Annalen*, 95:736–788, 1926.

25 Russell Impagliazzo, Pavel Pudlák, and Jiří Sgall. Lower Bounds for the Polynomial Calculus and the Gröbner Basis Algorithm. *Comput. Complex.*, 8(2):127–144, November 1999. doi:10.1007/s000370050024.

26 Kazuo Iwama and S. Miyazaki. Tree-Like Resolution is Superpolynomially Slower than DAG-Like Resolution for the Pigeonhole Principle. In *International Symposium on Algorithms and Computation*, 1999.

27 János Kollár. Sharp Effective Nullstellensatz. *Journal of the American Mathematical Society*, pages 963–975, 1988.

28 Toniann Pitassi and Robert Robere. Lifting Nullstellensatz to Monotone Span Programs over Any Field. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, pages 1207–1219, New York, NY, USA, 2018. Association for Computing Machinery. doi:10.1145/3188745.3188914.

29 Aaron Potechin and Aaron Zhang. Bounds on the Total Coefficient Size of Nullstellensatz Proofs of the Pigeonhole Principle and the Ordering Principle, 2022. arXiv:2205.03577.

30 Alexander A Razborov. Lower bounds for the polynomial calculus. *computational complexity*, 7(4):291–324, 1998.

31 S. F. De Rezende, A. Potechin, and K. Risse. Clique is Hard on Average for Unary Sherali-Adams. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 12–25, Los Alamitos, CA, USA, November 2023. IEEE Computer Society. doi:10.1109/FOCS57990.2023.00008.

32 Gunnar Stålmarck. Short resolution proofs for a sequence of tricky formulas. *Acta Informatica*, 33(3):277–280, 1996.

A Nullstellensatz total coefficient size can be smaller than size

The following example shows that Nullstellensatz total coefficient size can be smaller than Nullstellensatz proof size. (See Section 2 for the definition of total coefficient size.)

The idea behind our example is as follows. If we have three points p_1, p_2, p_3 and three polynomials f_1, f_2, f_3 such that

1. $f_1(p_1) = 1$, $f_1(p_2) = 1$, $f_1(p_3) = 0$
2. $f_2(p_1) = 1$, $f_2(p_2) = 0$, $f_2(p_3) = 1$
3. $f_3(p_1) = 0$, $f_3(p_2) = 1$, $f_3(p_3) = 1$

then given the axioms $f_1 = 0$, $f_2 = 0$, and $f_3 = 0$, the equality $\frac{1}{2}f_1 + \frac{1}{2}f_2 + \frac{1}{2}f_3 = 1$ is a Nullstellensatz proof of infeasibility which has total coefficient size $\frac{3}{2}$. However, if we want to use integer coefficients then we need coefficient size 2 as we need two of f_1, f_2, and f_3 in order to cover the three points p_1, p_2, p_3.

Our actual example is as follows. We have variables $x_1, x_2, x_3, x_4, x_5, x_6$ and we have the following axioms:

1. For all $I \subseteq \{4,5,6\}$, $x_1 \left(\prod_{i\in I} x_i\right)\left(\prod_{j\in\{4,5,6\}\setminus I} \bar{x}_j\right) = 0$
2. For all $I \subseteq \{4,5,6\}$, $x_2 \left(\prod_{i\in I} x_i\right)\left(\prod_{j\in\{4,5,6\}\setminus I} \bar{x}_j\right) = 0$
3. For all $I \subseteq \{4,5,6\}$, $x_3 \left(\prod_{i\in I} x_i\right)\left(\prod_{j\in\{4,5,6\}\setminus I} \bar{x}_j\right) = 0$
4. $x_1x_2x_3 = 0$
5. $\bar{x}_1\bar{x}_2 = 0$, $\bar{x}_1\bar{x}_3 = 0$, $\bar{x}_2\bar{x}_3 = 0$

We now observe that $1 = \frac{1}{2}\bar{x}_1\bar{x}_2 + \frac{1}{2}\bar{x}_1\bar{x}_3 + \frac{1}{2}x_1\bar{x}_2\bar{x}_3 + \frac{1}{2}(x_1 + x_2 + x_3) - \frac{1}{2}x_1x_2x_3$. We can show this by checking that the right hand side is 1 for all $(x_1, x_2, x_3) \in \{0,1\}^3$.

1. If $x_1 = x_2 = x_3 = 0$ then the first two terms are $\frac{1}{2}$ and the remaining terms are 0.
2. If $x_1 + x_2 + x_3 = 1$ then the fourth term and exactly one of the first three terms are $\frac{1}{2}$ and the remaining terms are 0.
3. If $x_1 + x_2 + x_3 = 2$ then the fourth term is 1 and the remaining terms are 0.
4. If $x_1 = x_2 = x_3 = 1$ then the fourth term is $\frac{3}{2}$, the fifth term is $-\frac{1}{2}$, and the remaining terms are 0.

Using this equation, we have that

$$1 = \frac{1}{2}\bar{x}_1\bar{x}_2 + \frac{1}{2}\bar{x}_1\bar{x}_3 + \frac{1}{2}x_1\bar{x}_2\bar{x}_3 - \frac{1}{2}x_1x_2x_3 + \frac{1}{2}(x_1 + x_2 + x_3)\sum_{I\subseteq\{4,5,6\}}\left(\prod_{i\in I} x_i\right)\left(\prod_{j\in\{4,5,6\}\setminus I} \bar{x}_j\right)$$

This Nullstellensatz proof has total coefficient size $4 * \frac{1}{2} + \frac{3*8}{2} = 14$. However, for each $I \subseteq \{4,5,6\}$, two of the axioms of the form $x_k \left(\prod_{i\in I} x_i\right)\left(\prod_{j\in\{4,5,6\}\setminus I} \bar{x}_j\right) = 0$ for $k \in \{1,2,3\}$ are needed to prove infeasibility. We also need one of the three axioms $\bar{x}_1\bar{x}_2 = 0$, $\bar{x}_1\bar{x}_3 = 0$, and $\bar{x}_2\bar{x}_3 = 0$. Thus, any Nullstellensatz proof of infeasibility must have size at least $2 * 8 + 1 = 17$.

B Total coefficient size upper bound for the pigeonhole principle

In this section, we use a divide and conquer approach to give a unary Nullstellensatz proof of the pigeonhole principle with size $2^{O(n)}$. Before giving our proof, we discuss other potential approaches for constructing a Nullstellensatz proof for the pigeonhole principle and why they were insufficient for our purposes.

One approach is to use the observation that if we have a tree-like resolution proof with S leaves, this gives us a Nullstellensatz proof of size S where every coefficient is 1.

There is a simple tree-like resolution proof of size $O((n+1)!)$ which works as follows. For each pigeon i, we query the variables $\{x_{ij} : j \in [n]\}$ one by one and stop when we find a j such that $x_{i,j} = 1$ or we have queried all of these variables. If we already had that $x_{i',j} = 1$ for some $i' < i$ then this contradicts the axiom $\neg x_{i',j} \vee \neg x_{i,j}$. If none of the $x_{i,j}$ are 1 then this contradicts the axiom $\bigvee_{j\in[n]} x_{i,j}$. If pigeon i was placed in a new hole j then we continue on to pigeon $i+1$.

This gives an upper bound of $O((n+1)!)$. However, it has been shown [8, 17, 26] that every tree-like resolution proof for the pigeonhole principle has size $n^{\Omega(\log(n))}$ so this is essentially the best that we can do with this approach.

Buss and Pitassi [13] showed that there is a resolution proof of size $O(n^3 2^n)$ for the pigeonhole principle. The idea behind this proof is as follows.

▶ **Definition 28.** *Given $S = \{s_1, \ldots, s_j\} \subseteq [n+1]$, define $C_{S,[j,n]}$ to be the clause*

$$C_{S,[j,n]} = \bigvee_{i\in[j],k\in[j,n]} x_{s_i,k}.$$

In other words, the clause $C_{S,[j,n]}$ says that at least one of the j pigeons in S must go into one of the holes in $[j,n]$.

At stage j, we start with the clauses $\{C_{S,[j,n]} : S \subseteq [n+1], |S| = j\}$ and derive the clauses $\{C_{S',[j+1,n]} : S' \subseteq [n+1], |S'| = j+1\}$. After stage n, this gives us the empty clause, which proves that the pigeonhole axioms are infeasible. However, it is not clear how to translate this proof into a Nullstellensatz proof without blowing up the size and total coefficient size.

We now give a unary Nullstellensatz proof of the pigeonhole principle which has size $2^{O(n)}$.

▶ **Theorem 29.** *For all $n \in \mathbb{N}$, there is a unary Nullstellensatz proof of size at most $2^{5(n+1)}$ for the pigeonhole principle with $n+1$ pigeons and n holes.*

Proof. We construct this proof recursively as follows. Given $k \in \mathbb{N}$, a set $S = \{p_1, \ldots, p_{k+1}\}$ of $k+1$ pigeons, and a set $H = \{h_1, \ldots, h_k\}$ of k holes, we want to show the equality

$$\prod_{a=1}^{k+1}\left(1 - \prod_{b=1}^{k}(1 - x_{p_a,h_b})\right) = \prod_{a=1}^{k+1}\left(1 - \prod_{b=1}^{k}\bar{x}_{p_a,h_b}\right) = 0$$

using the hole axioms. Note that this equality coresponds to the statement that there is at least one pigeon in S which does not go to any of the holes in H.

We do this as follows. If $k = 1$ then this equality is a hole axiom. If $k > 2$ then

1. For each $a \in [k]$, we decompose the term $\left(1 - \prod_{b=1}^{k}\bar{x}_{p_a,h_b}\right)$ as

$$\left(1 - \prod_{b=1}^{k}\bar{x}_{p_a,h_b}\right) = \left(1 - \prod_{b=1}^{\lfloor\frac{k+1}{2}\rfloor}\bar{x}_{p_a,h_b}\right) + \left(\prod_{b=1}^{\lfloor\frac{k+1}{2}\rfloor}\bar{x}_{p_a,h_b}\right)\left(1 - \prod_{b=\lfloor\frac{k+1}{2}\rfloor+1}^{k}\bar{x}_{p_a,h_b}\right)$$

This gives us the equality

$$\prod_{a=1}^{k+1}\left(1-\prod_{b=1}^{k}\bar{x}_{p_a,h_b}\right)=$$
$$\sum_{A\subseteq[k+1]}\left(\prod_{a\in A}\left(1-\prod_{b=1}^{\lfloor\frac{k}{2}\rfloor}\bar{x}_{p_a,h_b}\right)\right)\left(\prod_{a\in[k+1]\setminus A}\left(\prod_{b=1}^{\lfloor\frac{k+1}{2}\rfloor}\bar{x}_{p_a,h_b}\right)\right)\left(1-\prod_{b=\lfloor\frac{k+1}{2}\rfloor+1}^{k}\bar{x}_{p_a,h_b}\right)\right)$$

2. For each of the 2^{k+2} resulting terms, we check whether $|A|\geq\lfloor\frac{k+1}{2}\rfloor+1$ or $|A|\leq\lfloor\frac{k+1}{2}\rfloor$. If $|A|\geq\lfloor\frac{k+1}{2}\rfloor+1$ then letting A' be the first $\lfloor\frac{k+1}{2}\rfloor+1=\lceil\frac{k+2}{2}\rceil$ elements of A, we recursively construct a proof that $\prod_{a\in A'}\left(1-\prod_{b=1}^{\lfloor\frac{k+1}{2}\rfloor}\bar{x}_{p_a,h_b}\right)=0$. If $|A|\leq\lfloor\frac{k+1}{2}\rfloor$ then $|[k+1]\setminus A|\geq\lceil\frac{k+1}{2}\rceil$ so letting A'' be the first $\lceil\frac{k+1}{2}\rceil$ elements of $[k+1]\setminus A$, we recursively construct a proof that $\prod_{a\in A''}\left(1-\prod_{b=\lfloor\frac{k+1}{2}\rfloor+1}^{k}\bar{x}_{p_a,h_b}\right)=0$.

To obtain our Nullstellensatz proof, we construct this proof for $S=[n+1]$ and $H=[n]$. We then use the following equality (recall that the pigeon axioms are $\{\prod_{b=1}^{n}\bar{x}_{a,b}=0:a\in[n+1]\}$):

$$1=\prod_{a=1}^{n+1}\left(1-\prod_{b=1}^{n}\bar{x}_{a,b}\right)+\sum_{j=1}^{n+1}\left(\prod_{b=1}^{n}\bar{x}_{j,b}\right)\left(\prod_{a=1}^{j-1}\left(1-\prod_{b=1}^{n}\bar{x}_{a,b}\right)\right)$$

The size of the resulting unary Nullstellensatz proof can be upper bounded by $S(n)+2^{n+1}$ where $S(n)$ is the solution to the recurrence relation $S(n)=2^{2(n+1)}S(\lceil\frac{n+2}{2}\rceil)$ where $S(1)=1$. It is not hard to show by induction that $S(n)\leq 2^{5(n+1)}-2^{(n+1)}$ so this gives an upper bound of $2^{5(n+1)}$. ◀

C Total coefficient size upper bound for the ordering principle

In this section, we construct an explicit Nullstellensatz proof for the ordering principle on n elements with total coefficient size 2^n-n. In the full version of this paper [29], we also present experimental results obtained by implementing the linear program for minimum total coefficient size. One of our experimental results is that the 2^n-n upper bound is tight for $n\leq 5$.

We start by formally defining the ordering principle.

▶ **Definition 30** (ordering principle (ORD_n)). *Intuitively, the ordering principle says that any well-ordering on n elements must have a minimum element. Formally, for $n\geq 1$, we define ORD_n to be the statement that the following system of axioms is infeasible:*

- *We have a variable $x_{i,j}$ for each pair $i,j\in[n]$ with $i<j$, with the Boolean axiom $x_{i,j}^2-x_{i,j}=0$. $x_{i,j}=1$ represents element i being less than element j in the well-ordering, and $x_{i,j}=0$ represents element i being more than element j in the well-ordering. We write $x_{j,i}$ as shorthand for $\bar{x}_{i,j}=1-x_{i,j}$.*
- *For each $i\in[n]$, we have the axiom $\prod_{j\in[n]\setminus\{i\}}x_{i,j}=0$ which represents the constraint that element i is not a minimum element. We call these axioms non-minimality axioms.*
- *For each triple $i,j,k\in[n]$ where $i<j<k$, we have the two axioms $x_{i,j}x_{j,k}x_{k,i}=0$ and $x_{k,j}x_{j,i}x_{i,k}=0$ which represent the constraints that elements i,j,k satisfy transitivity. We call these axioms transitivity axioms.*

In our Nullstellensatz proof $1=\sum_W c_W W$, each c_W is either 0 or 1. Non-minimality axioms have coefficient 1, and all weakenings of transitivity axioms that have coefficient 1 must have a special form:

▶ **Definition 31** (nice transitivity weakening). *Let W be a weakening of the axiom $x_{i,j}x_{j,k}x_{k,i}$ or the axiom $x_{k,j}x_{j,i}x_{i,k}$ for some $i < j < k$. Let $G(W)$ be the following directed graph. The vertices of $G(W)$ are $[n]$. For distinct $i', j' \in [n]$, $G(W)$ has an edge from i' to j' if W contains the term $x_{i',j'}$. We say that W is a nice transitivity weakening if $G(W)$ has exactly n edges and all vertices are reachable from vertex i.*

In other words, if W is a weakening of the axiom $x_{i,j}x_{j,k}x_{k,i}$ or the axiom $x_{k,j}x_{j,i}x_{i,k}$, then $G(W)$ contains a 3-cycle on vertices $\{i, j, k\}$. W is a nice transitivity weakening if and only if contracting this 3-cycle results in a directed spanning tree rooted at the contracted vertex. Note that if W is a nice transitivity weakening and x is an assignment with a minimum element, then $W(x) = 0$.

▶ **Theorem 32.** *There is a Nullstellensatz proof for* ORD_n *satisfying:*

1. *The total coefficient size is $2^n - n$.*
2. *Each c_W is either 0 or 1.*
3. *If A is a non-minimality axiom, then $c_A = 1$, and $c_W = 0$ for all other weakenings W of A.*
4. *If W is a transitivity weakening but not a nice transitivity weakening, then $c_W = 0$.*

Proof. We prove Theorem 32 by induction on n. When $n = 3$, the desired Nullstellensatz proof sets $c_A = 1$ for each axiom A. It can be verified that $\sum_W c_W W$ evaluates to 1 on each assignment, and that this Nullstellensatz proof satisfies the properties of Theorem 32.

Now suppose we have a Nullstellensatz proof for ORD_n satisfying Theorem 32, and let S_n denote the set of transitivity weakenings W for which $c_W = 1$. The idea to obtain a Nullstellensatz proof for ORD_{n+1} is to use two copies of S_n, the first copy on elements $\{1, \ldots, n\}$ and the second copy on elements $\{2, \ldots, n+1\}$. Specifically, we construct the Nullstellensatz proof for ORD_{n+1} by setting the following c_W to 1 and all other c_W to 0.

1. For each non-minimality axiom A in ORD_{n+1}, we set $c_A = 1$.
2. For each $W \in S_n$, we define the transitivity weakening W' on $n+1$ elements by $W' = W \cdot x_{1,n+1}$ and set $c_{W'} = 1$.
3. For each $W \in S_n$, first we define the transitivity weakening W'' on $n+1$ elements by replacing each variable $x_{i,j}$ that appears in W by $x_{i+1,j+1}$ (e.g., if $W = x_{1,2}x_{2,3}x_{3,1}$, then $W'' = x_{2,3}x_{3,4}x_{4,2}$). Then, we define $W''' = W''x_{n+1,1}$ and set $c_{W'''} = 1$.
4. For each $i \in \{2, \ldots, n\}$, for each of the 2 transitivity axioms A for elements $\{1, i, n+1\}$, we set $c_W = 1$ for the following weakening W of A:

$$W = A\left(\prod_{j \in \{2,\ldots,n\} \setminus \{i\}} x_{i,j}\right).$$

In other words, $W(x) = 1$ if and only if $A(x) = 1$ and i is the minimum element among $\{2, \ldots, n\}$.

The desired properties 1 through 4 in Theorem 32 can be verified by induction. It remains to show that for each assignment x, there is exactly one nonzero c_W for which $W(x) = 1$. If x has a minimum element $i \in [n+1]$, then the only nonzero c_W for which $W(x) = 1$ is the non-minimality axiom for i. Now suppose that x does not have a minimum element. Consider two cases: either $x_{1,n+1} = 1$, or $x_{n+1,1} = 1$. Suppose $x_{1,n+1} = 1$. Consider the two subcases:

1. Suppose that, if we ignore element $n+1$, then there is still no minimum element among the elements $\{1, \ldots, n\}$. Then there is exactly one weakening W in point 2 of the construction for which $W(x) = 1$, by induction.
2. Otherwise, for some $i \in \{2, \ldots, n\}$, we have that i is a minimum element among $\{1, \ldots, n\}$ and $x_{n+1,i} = 1$. Then there is exactly one weakening W in point 4 of the construction for which $W(x) = 1$ (namely, the weakening W of the axiom $A = x_{i,1}x_{1,n+1}x_{n+1,i}$).

The case $x_{n+1,1} = 1$ is handled similarly by considering whether there is a minimum element among $\{2, \ldots, n+1\}$. Assignments that do have a minimum element among $\{2, \ldots, n+1\}$ are handled by point 3 of the construction, and assignments that do not are handled by point 4 of the construction. ◀