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

Aaron Zhang
Advisor: Aaron Potechin

Abstract

In this paper, we investigate the total coefficient size of Nullstellensatz proofs. We show that Nullstellensatz proofs of the pigeonhole principle on n pigeons require total coefficient size $2^{\Omega(n)}$ and that there exist Nullstellensatz proofs of the ordering principle on n elements with total coefficient size $2^n - n$.

Acknowledgement: This research was supported by NSF grant CCF-2008920 and NDSEG fellowship F-9422254702.

1 Introduction

In this paper, we investigate the total coefficient size of Nullstellensatz proofs. We start by introducing the Nullstellensatz proof system and then define total coefficient size as a measure of the complexity of Nullstellensatz proofs.

A Nullstellensatz proof is a way to certify that a system of axioms is infeasible. In this paper, we take a system of axioms to be a system of equations of the following form.

Definition 1. Let $x_1, \dots, x_N$ be variables that take values in $\{0, 1\}$. Let $\bar{x}_i$ denote $1 - x_i$. We define a monomial to be a product of the form $\left(\prod_{i \in S} x_i\right)\left(\prod_{j \in T} \overline{x_j}\right)$ for some disjoint subsets S, T of $[N]$. We define a system of axioms to be a system of m equations $\{p_i = 0 : i \in [m]\}$ where each p_i is a monomial. We call each p_i an axiom.

For example, the following system of axioms expresses the pigeonhole principle, which says that if n pigeons are assigned to $n - 1$ holes, then some hole must have more than one pigeon. The pigeonhole principle is one of two systems of axioms we will consider in this paper, together with the ordering principle.

Definition 2. For $n \geq 1$, we define PHP_n to be the following system of axioms.

- For each $i \in [n]$ and $j \in [n-1]$, we have a variable $x_{i,j}$. $x_{i,j} = 1$ represents pigeon i being in hole j, and $x_{i,j} = 0$ represents pigeon i not being in hole j.
- For each $i \in [n]$, we have the axiom $\prod_{j=1}^{n-1} \bar{x}_{i,j} = 0$ representing the constraint that each pigeon must be in at least one hole.
- For each pair of distinct pigeons $i_1, i_2 \in [n]$ and each hole $j \in [n-1]$, 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 now define the notion of weakenings, which are used in Nullstellensatz proofs.

Definition 3. Let $\{p_i = 0 : i \in [m]\}$ be a system of axioms. If W is a monomial such that $W = rp_i$ for some monomial r and some p_i, we say that W is a weakening of p_i.

For example, if we have variables x_1, x_2, x_3 and an axiom $x_1x_2 = 0$, then the possible weakenings of this axiom are x_1x_2, $x_1x_2x_3$, and $x_1x_2\bar{x}_3$. We use the following notation:

Definition 4. If $x \in \{0,1\}^N$, we call x an assignment and view it as a choice of $\{0,1\}$-values for $x_1, \ldots, x_N$. For an axiom p_i or a weakening W, we let $p_i(x)$ or $W(x)$ denote p_i or W evaluated on x. Thus, if W is a weakening of p_i, then any assignment x satisfying $p_i(x) = 0$ also satisfies $W(x) = 0$.

Now we define the Nullstellensatz proof system.

Definition 5. Let $\{p_i = 0 : i \in [m]\}$ be a system of axioms. A Nullstellensatz proof that $\{p_i = 0 : i \in [m]\}$ is infeasible is an equality of the form

$$1 = \sum_W c_W W,$$

where W ranges over all possible weakenings of axioms and $c_W \in \mathbb{R}$.

If we can write $1 = \sum_W c_W W$, then $\{p_i = 0 : i \in [m]\}$ is infeasible because any assignment x satisfying $p_i(x) = 0$ for all axioms also satisfies $W(x) = 0$ for all weakenings. The example below is a Nullstellensatz proof for PHP_2. The axioms are:

$$\begin{aligned} 1 - x_{1,1} &= 0 \\ 1 - x_{2,1} &= 0 \\ x_{1,1}x_{2,1} &= 0 \end{aligned}$$

We have the Nullstellensatz proof:

$$\begin{aligned} 1 = &(1 - x_{1,1}) \\ &+ x_{1,1}\,(1 - x_{2,1}) \\ &+ (x_{1,1}x_{2,1})\,. \end{aligned}$$

Finally, we define total coefficient size, which is the complexity measure of Nullstellensatz proofs that we will be consider in this paper.

Definition 6. Given a system of axioms $\{p_i = 0 : i \in [m]\}$ and a Nullstellensatz proof $1 = \sum_W c_W W$, we define the total coefficient size of this Nullstellensatz proof to be

$$\sum_W |c_W|.$$

The Nullstellsatz proof above for PHP_2 has total coefficient size 3. Total coefficient size is a complexity measure of Nullstellensatz proofs that is relatively explored, unlike degree which has been extensively studied [BIK+94, BP96, Bus96, CEI96, BIK+97, BCE+98]. We believe that total coefficient size is a natural complexity measure in its own right, and it is also an interesting open problem whether there are tradeoffs between total coefficient size and other complexity measures like degree.

In this paper, we show the following results. In section 2, we show how the problem of minimizing total coefficient size can be expressed as a linear program.

In section 3, we prove our main result, which is an exponential lower bound on total coefficient size for the pigeonhole principle:

Theorem 1. Any Nullstellensatz proof of the pigeonhole principle on n pigeons and $n-1$ holes has total coefficient size at least $\Omega\left(n^{\frac{3}{4}}\left(\frac{2}{\sqrt{e}}\right)^n\right)$.

In section 4, we introduce the ordering principle and show an exponential upper bound on total coefficient size:

Theorem 2. For all $n \geq 1$, there is a Nullstellensatz proof of the ordering principle on n elements with total coefficient size $2^n - n$.

Finally, in section 5 we discuss open problems that our work raises.

2 Linear program for total coefficient size

Given a set of axioms $\{p_i = 0 : i \in [m]\}$, we can express the problem of minimizing total coefficient size with the following linear program.

Variables: For each weakening W, we have a variable c_W and a variable c_W^{abs} representing the absolute value of c_W.

Constraints: For each weakening W, we have the two constraints

$$c_W^{\text{abs}} - c_W \geq 0,$$
$$c_W^{\text{abs}} + c_W \geq 0.$$

For each assignment x, we have the constraint

$$\sum_W c_W W(x) = 1$$

expressing the condition that $\sum_W c_W W$ evaluates to 1 on all assignments.

Objective:

$$\min \sum_W c_W^{\text{abs}}.$$

We now derive the dual linear program in the standard way.

Variables: For each weakening W, we have a variable $u_W^+ \geq 0$ (corresponding to the primal constraint $c_W^{\text{abs}} - c_W \geq 0$) and a variable $u_W^- \geq 0$ (corresponding to the primal constraint $c_W^{\text{abs}} + c_W \geq 0$).

For each assignment x, we have a variable v_x (corresponding to the primal constraint $\sum_W c_W W(x) = 1$).

Constraints: For each weakening W, we have a constraint corresponding to the primal variable c_W:

$$\left(\sum_x v_x W(x)\right) + (-u_W^+ + u_W^-) = 0.$$

For each weakening W, we have a constraint corresponding to the primal variable c_W^{abs}:

$$u_W^+ + u_W^- = 1.$$

Objective:

$$\max \sum_x v_x.$$

Because $u_W^+, u_W^- \geq 0$ and $u_W^+ + u_W^- = 1$, we have that $(-u_W^+ + u_W^-)$ can be any value in the range $[-1, 1]$. Therefore, the dual constraints can be rephrased as

$$\forall W : \sum_x v_x W(x) \in [-1, 1].$$

3 Total coefficient size lower bound for the pigeonhole principle

In this section, we prove Theorem 1, an exponential lower bound on total coefficient size for the pigeonhole principle PHP_n (Definition 2).

We will prove our lower bound by constructing and analyzing a dual solution D. For an assignment x, let $D(x)$ denote the value of v_x in our dual solution. The only assignments x for which $D(x) \neq 0$ will be those where each pigeon goes to exactly one hole (i.e., for each pigeon i, exactly one of the $x_{i,j}$ is 1). Note that there are $(n-1)^n$ such assignments. In the rest of this section, when we refer to assignments or write a summation over assignments x, we refer specifically to these $(n-1)^n$ assignments. Further, if W is a weakening of an axiom of the form $\prod_{j=1}^{n-1} \bar{x}_{i,j} = 0$, then D trivially satisfies the dual constraint $\sum_x D(x)W(x) \in [-1, 1]$ because $D(x)W(x) = 0$ for all x. Therefore, in the rest of this section when we refer to weakenings, we refer specifically to weakenings of the axioms $x_{i_1,j} x_{i_2,j} = 0$.

For convenience, we will construct a function D that does not necessarily satisfy $\forall W : \sum_x D(x)W(x) \in [-1, 1]$; we can then obtain a valid dual solution by dividing each $D(x)$ by $\max_W |\sum_x D(x)W(x)|$. Letting $\mathbb{E}$ denote expectation over a uniform assignment (where each pigeon goes to exactly one hole), we obtain a dual value of $\frac{\sum_x D(x)}{\max_W |\sum_x D(x)W(x)|} = \frac{\mathbb{E}(D)}{\max_W |\mathbb{E}(DW)|}$. So, we will construct D and analyze $\mathbb{E}(D)$ and $\max_W |\mathbb{E}(DW)|$.

First, we provide some intuition for our construction. Consider the following functions on assignments, which indicate whether a subset of pigeons is in different holes:

Definition 7. Let $S \subsetneq [n]$ be a subset of pigeons of size at most $n-1$. We define the function J_S that maps assignments to $\{0,1\}$. For an assignment x, $J_S(x) = 1$ if all pigeons in S are in different holes, 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 $\mathbb{E}(J_S) = \left(\prod_{k=1}^{|S|}(n-k)\right)/(n-1)^{|S|}$.

Naively, we might want D to be the indicator for whether all n pigeons are in different holes. Of course, this wouldn't work because it is impossible for all n pigeons to be in different holes. Perhaps the next best thing would be that if we only consider a subset S of $n-1$ pigeons, then D "mimics" J_S. More concretely, suppose p is a monomial that does not depend on some pigeon i (i.e., p does not contain any terms $x_{i,j}$ or $\bar{x}_{i,j}$). Then, we might hope that D mimics $J_{[n]\setminus\{i\}}$ in the sense that $\mathbb{E}(Dp) = \mathbb{E}\left(J_{[n]\setminus\{i\}}p\right)$. Given this intuition, we now construct D.

Definition 8. Our dual solution D is

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

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

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 the intuition for our construction:

Lemma 1. If p is a monomial that does not depend on pigeon i (i.e., p does not contain any terms $x_{i,j}$ or $\bar{x}_{i,j}$), then $\mathbb{E}(Dp) = \mathbb{E}(J_{[n]\setminus\{i\}}p)$.

Proof. Without loss of generality, suppose p does not contain any terms $x_{1,j}$ or $\bar{x}_{1,j}$. Let $T \subsetneq \{2,\dots,n\}$. Observe that

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

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

$$\begin{aligned} c_{T\cup\{1\}} &= \frac{(-1)^{n-2-|T|}(n-2-|T|)!}{(n-1)^{n-2-|T|}} \\ &= -\frac{n-1}{n-1-|T|}\cdot\frac{(-1)^{n-1-|T|}(n-1-|T|)!}{(n-1)^{n-1-|T|}} = -\frac{n-1}{n-1-|T|}c_T, \end{aligned}$$

we have that for all $T \subsetneq \{2,\dots,n\}$,

$$\mathbb{E}(c_{T\cup\{1\}}J_{T\cup\{1\}}p) + \mathbb{E}(c_T J_T p) = 0.$$

Thus, all terms in $\mathbb{E}(Dp)$ except $\mathbb{E}(c_{\{2,\dots,n\}}J_{\{2,\dots,n\}}p)$ cancel. Since $c_{\{2,\dots,n\}} = 1$, we have that $\mathbb{E}(Dp) = \mathbb{E}(J_{\{2,\dots,n\}}p)$, as needed. □

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

Corollary 1.

$$\mathbb{E}(D) = \frac{(n-2)!}{(n-1)^{n-2}}.$$

Proof. Let $p = 1$. By Lemma 1, $\mathbb{E}(D) = \mathbb{E}(J_{\{2,\dots,n\}}) = (n-2)!/(n-1)^{n-2}$. □

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

We introduce the following notation:

Definition 9 ($H_{W,i}$)**.** Given a weakening W, we define a set of holes $H_{W,i} \subseteq [n-1]$ for each pigeon $i \in [n]$ so that $W(x) = 1$ if and only if each pigeon $i \in [n]$ goes 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 h and h').
- 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 1:

Lemma 2. Let W be a weakening. If there exists some pigeon $i \in [n]$ such that $H_{W,i} = [n-1]$ (i.e., W does not contain any terms $x_{i,j}$ 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 1, $\mathbb{E}(DW) = \mathbb{E}(J_{\{2,\dots,n\}}W)$. However, $\mathbb{E}(J_{\{2,\dots,n\}}W) = 0$ because if $W = 1$ then pigeons 2 and 3 must both go to hole 1. □

We make the following definition and then state a corollary of Lemma 2.

Definition 10 (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 $S \subseteq [n] \setminus \{i_1, i_2\}$. 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 $H_{W_S^{\text{flip}},i_3} = [n-1] \setminus H_{W,i_3}$.
- For each pigeon $i_3 \notin S$, we define W_S^{flip} so that $H_{W_S^{\text{flip}},i_3} = H_{W,i_3}$.

(Technically, there may be multiple ways to define W_S^{flip} to satisfy these properties; we can arbitrarily choose any such definition.)

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. Now we state a corollary of Lemma 2:

Corollary 2. 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 $S \subseteq [n] \setminus \{i_1, i_2\}$. Then

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

Proof. It suffices to show that for $i_3 \in [n] \setminus \{i_1, i_2\}$, we have $\mathbb{E}\left(DW_{\{i_3\}}^{\text{flip}}\right) = -\mathbb{E}(DW)$. Indeed, $W + W_{\{i_3\}}^{\text{flip}}$ is a weakening satisfying $H_{W+W_{\{i_3\}}^{\text{flip}},i_3} = [n-1]$. Therefore, by Lemma 2, $\mathbb{E}\left(D\left(W + W_{\{i_3\}}^{\text{flip}}\right)\right) = 0$. □

Using Corollary 2, we will bound $\max_W |\mathbb{E}(DW)|$ using Cauchy-Schwarz. We first show an approach that does not give a strong enough bound but shows how Corollary 2 and Cauchy-Schwarz can be useful. We then show how to improve the bound.

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

Consider $\max_W |\mathbb{E}(DW)|$. By Corollary 2, it suffices to take the max only 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 $i_3 \in [n] \setminus \{i_1, i_2\}$, we have $|H_{W,i_3}| \leq \lfloor (n-1)/2 \rfloor$. For any such W, we have

$$\begin{aligned}\|W\| &= \sqrt{E(W^2)} \\ &\leq \sqrt{\left(\frac{1}{n-1}\right)^2 \left(\frac{1}{2}\right)^{n-2}} \\ &= (n-1)^{-1} \cdot 2^{-(n-2)/2}.\end{aligned}$$

By Cauchy-Schwarz,

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

Using the value of $\mathbb{E}(D)$ from Corollary 1, the dual value $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ is at least

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

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 that $1/\|D\| = \widetilde{\Theta}\left((\sqrt{e})^n\right)$. We now improve our approach.

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

Definition 11 ($W^{\{-1,0,1\}}$)**.** 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. (If W is a weakening of multiple such axioms, we choose one of those axioms arbitrarily.) We define the function $W^{\{-1,0,1\}}$ 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 $W^{\{-1,0,1\}}(x) = 0$.

- Otherwise, let $V(x) = |\{i_3 \in [n]\backslash\{i_1, i_2\} : \text{pigeon } i_3 \text{ does not go to } H_{W,i_3}\}|$. Then $W^{\{-1,0,1\}}(x) = (-1)^{V(x)}$.

$W^{\{-1,0,1\}}$ is a linear combination of the W_S^{flip}:

Lemma 3. 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

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

It follows that

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

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 $W^{\{-1,0,1\}}$ 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 $W^{\{-1,0,1\}}(x) = (-1)^{|S|}$.

The second equation follows because

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

□

Using Lemma 3, we now upper bound $\max_W |\mathbb{E}(DW)|$:

Lemma 4. The dual value $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ is at least $\frac{(n-2)!}{(n-1)^{n-2}} \cdot \frac{(n-1)2^{n-2}}{\|D\|}$.

Proof. For any W, we have

$$\begin{aligned}
\mathbb{E}(DW) &= 2^{-(n-2)} \cdot \mathbb{E}\left(DW^{\{-1,0,1\}}\right) && \text{(Lemma 3)} \\
&\leq 2^{-(n-2)} \cdot \|D\|\|W^{\{-1,0,1\}}\| && \text{(Cauchy-Schwarz)} \\
&= 2^{-(n-2)} \cdot \|D\|\sqrt{\mathbb{E}\left(\left(W^{\{-1,0,1\}}\right)^2\right)} \\
&= (n-1)^{-1}2^{-(n-2)} \cdot \|D\|.
\end{aligned}$$

Using the value of $\mathbb{E}(D)$ from Corollary 1, the dual value $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ is at least $\frac{(n-2)!}{(n-1)^{n-2}} \cdot \frac{(n-1)2^{n-2}}{\|D\|}$. □

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

Lemma 5.

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

Proof. Recall the definition of D (Definition 8):

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

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

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

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

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

Therefore,

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

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

$$c_S c_T \cdot \mathbb{E}(J_S J_T) = (-1)^{|S|-|T|}\left(\frac{(n-2)!}{(n-1)^{n-2}}\right)\left(\frac{(n-1-|S\cap T|)!}{(n-1)^{n-1-|S\cap T|}}\right).$$

We rearrange the sum for $\mathbb{E}(D^2)$ based on $|S \cap T|$:

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

To evaluate this expression, fix $c \leq n-1$ and consider the inner sum. Consider the collection of tuples $\{(S,T) \mid S, T \subsetneq [n], |S \cap T| = c\}$. We can pair up (most of) these tuples in the following way. For each S, let m_S denote the minimum element in $[n]$ that is not in S (note that m_S is well defined because S cannot be $[n]$). 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 $T = [n] \setminus \{m_S\}$, because T cannot be $[n]$. There are $\binom{n}{c}$ unpaired tuples (S,T), and for each of these tuples, we have $(-1)^{|S|-|T|} = (-1)^{n-1-c}$. On the other hand, each pair

$(S,T),(S,T\triangle\{m_S\})$ contributes 0 to the inner sum. Therefore, the inner sum equals $(-1)^{n-1-c}\binom{n}{c}$, and we have

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

□

Corollary 3. $\mathbb{E}(D^2)\le\frac{n!}{(n-1)^{n-1}}$

Proof. Observe that the sum

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

is an alternating series where the magnitudes of the terms decrease as c decreases. The two largest magnitude terms are $1/(n-1)!$ and $-(1/2)\cdot 1/(n-1)!$. Therefore, the sum is at most $\frac{1}{(n-1)!}$, and we conclude that

$$\mathbb{E}(D^2)\le\frac{(n-2)!}{(n-1)^{n-2}}\cdot\frac{n!}{(n-1)!}=\frac{n!}{(n-1)^{n-1}}$$

as needed. □

We can now complete the proof of Theorem 1.

Proof of Theorem 1. By Lemma 4, any Nullstellensatz proof for PHP_n has total coefficient size at least $\frac{(n-2)!}{(n-1)^{n-2}}\cdot\frac{(n-1)2^{n-2}}{\|D\|}$. By Corollary 3, $\|D\|\le\sqrt{\frac{n!}{(n-1)^{n-1}}}$. Combining these results, any Nullstellensatz proof for PHP_n has total coefficient size at least

$$\begin{aligned}\frac{(n-2)!}{(n-1)^{n-2}}\cdot\frac{(n-1)2^{n-2}}{\sqrt{\frac{n!}{(n-1)^{n-1}}}} &= \frac{2^{n-2}}{\sqrt{n}}\cdot\frac{\sqrt{(n-1)!}}{(n-1)^{\frac{n}{2}-\frac{3}{2}}}\\ &= \frac{2^{n-2}(n-1)}{\sqrt{n}}\sqrt{\frac{(n-1)!}{(n-1)^{n-1}}}\end{aligned}$$

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

3.2 Experimental results

For small n, we computed the optimal dual values shown below. The first column of values is the optimal dual value for $n = 3, 4$. The second column of values is the optimal dual value for $n = 3, 4, 5, 6$ under the restriction that the only nonzero assignments are those where each pigeon goes to exactly one hole.

n	dual value	dual value, each pigeon goes to exactly one hole
3	11	6
4	$41.4\overline{69}$	27
5	-	100
6	-	293.75

For comparison, the table below shows the value we computed for our dual solution and the lower bound of $\frac{2^{n-2}(n-1)}{\sqrt{n}}\sqrt{\frac{(n-1)!}{(n-1)^{n-1}}}$ that we showed in the proof of Theorem 1. (Values are rounded to 3 decimals.)

n	value of D	proven lower bound on value of D
3	4	1.633
4	18	2.828
5	64	4.382
6	210.674	6.4

It is possible that our lower bound on the value of D can be improved. The following experimental evidence suggests that the dual value $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ of D may actually be $\widetilde{\Theta}(2^n)$. For $n = 3, 4, 5, 6$, we found that the weakenings W that maximize $|\mathbb{E}(DW)|$ are of the following form, up to symmetry. (By symmetry, we mean that we can permute pigeons/holes without changing $|\mathbb{E}(DW)|$, and we can flip sets of holes as in Corollary 2 without changing $|\mathbb{E}(DW)|$.)

- For odd n ($n = 3, 5$): W is the weakening of the axiom $x_{1,1}x_{2,1} = 0$ where, for $i = 3, \ldots, n$, we have $H_{W,i} = \{2, \ldots, (n+1)/2\}$.
- For even n ($n = 4, 6$): W is the following weakening of the axiom $x_{1,1}x_{2,1} = 0$. For $i = 3, \ldots, n/2+1$, we have $H_{W,i} = \{2, \ldots, n/2\}$. For $i = n/2+2, \ldots, n$, we have $H_{W,i} = \{n/2+1, \ldots, n-1\}$.

If this pattern continues to hold for larger n, then experimentally it seems that $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ is $\widetilde{\Theta}(2^n)$, although we do not have a proof of this.

4 Total coefficient size upper bound for the ordering principle

In this section, we construct an explicit Nullstellensatz proof of infeasibility for the ordering principle on n elements with total coefficient size $2^n - n$.

Definition 12. 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 following system of axioms.

- We have a variable $x_{i,j}$ for each pair $i, j \in [n]$ with $i < j$. $x_{i,j} = 1$ 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 $\bar{x}_{i,j}$.
- For each $i \in [n]$, we have the axiom $\prod_{j \in [n] \setminus \{i\}} x_{i,j} = 0$ representing the constraint that element i is not a minimum element. We call these axioms non-minimality axioms.
- For each triple $i, j, k \in [n]$ with $i < j < k$, we have the two axioms $x_{i,j}x_{j,k}x_{k,i} = 0$ and $x_{k,j}x_{j,i}x_{i,k} = 0$ representing the constraint that elements i, j, k satisfy transitivity. We call these axioms transitivity axioms.

In our Nullstellensatz proof, each c_W will be either 0 or 1. If A is a non-minimality axiom, then $c_A = 1$ and $c_W = 0$ for all other weakenings W of A. The only weakenings of transitivity axioms that can have coefficient 1 are of the following form.

Definition 13. 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 3. There is a Nullstellensatz proof of infeasibility 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 3 by induction on n. When $n \leq 3$, the desired Nullstellensatz proof sets $c_A = 1$ for each axiom A. It can be verified that $\sum_W c_W W = 1$ and that this Nullstellensatz proof satisfies the properties of Theorem 3.

Now suppose we have a Nullstellensatz proof for ORD_n satisfying Theorem 3, 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 on elements $1, i, n+1$, we set $c_W = 1$ for the following weakening W of A:

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

In other words, $W(x) = 1$ if and only if $A(x) = 1$ and i is the minimum element among the elements $[n+1] \setminus \{1, n+1\}$.

The desired properties 1 through 4 in Theorem 3 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 of the axiom $x_{n+1,i}x_{i,1}x_{1,n+1}$).

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

4.1 Experimental results

For small values of n, we computed the minimum total coefficient size for the ordering principle. For $n = 3, 4, 5$, the minimum total coefficient size is $2^n - n$, so the primal solution given by Theorem 3 is optimal. However, for $n = 6$ this solution is not optimal as the minimum total coefficient size is 52 rather than $2^6 - 6 = 58$.

5 Open problems

In this paper, we proved an exponential lower bound on total coefficient size for the pigeonhole principle and an exponential upper bound on total coefficient size for the ordering principle. The total coefficient size of Nullstellensatz proofs is still relatively unexplored, and our work leaves many open questions including the following.

1. For the pigeonhole principle, can we improve our lower bound (Theorem 1) or prove any nontrivial upper bound?

2. What is the minimum total coefficient size for the pigeonhole principle on n pigeons when the number of holes is less than $n-1$?
3. For the ordering principle, can we improve our upper bound (Theorem 2) or prove any nontrivial lower bound?
4. Are there tradeoffs between Nullstellensatz total coefficient size and other complexity measures for Nullstellensatz or other proof systems?

References

[BCE+98] 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, aug 1998.

[BIK+94] 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.

[BIK+97] 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, dec 1997.

[BP96] 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.

[Bus96] 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.

[CEI96] 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, page 174–183, New York, NY, USA, 1996. Association for Computing Machinery.