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

Aaron Zhang
Advisor: Aaron Potechin

Abstract

We study the total coefficient size of Nullstellensatz proofs. We show that any Nullstellensatz proof of the pigeonhole principle on n pigeons and $n-1$ holes has total coefficient size at least $2^{\Omega(n)}$, and that there is a Nullstellensatz proof of the ordering principle on n elements with total coefficient size $2^n - n$.

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

1 Introduction

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

Definition 1. Let $x_1, \dots, x_N \in \{0,1\}$ be Boolean variables. Let $\overline{x_i}$ 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.

Nullstellensatz can be defined in more general contexts, such as systems of polynomial equations instead of monomimal equations, but we will only need the definition above. The following system of axioms represents the pigeonhole principle and is one of the two systems of axioms we will consider, together with the ordering principle. Intuitively, the pigeonhole principle says that if n pigeons go to $n-1$ holes, then some hole must have more than one pigeon. Formally, the pigeonhole principle says that the following system of axioms is infeasible.

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

- For each pigeon $i \in [n]$ and each hole $j \in [n-1]$, 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 pigeon $i \in [n]$, we have the axiom $\prod_{j=1}^{n-1} \overline{x_{i,j}} = 0$, 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 $x_1x_2\overline{x_3}$. We call these weakenings because any assignment to the variables that makes an axiom evaluate to 0 also makes all weakenings of the axiom evaluate to 0. Now we define Nullstellensatz proofs.

Definition 4. 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 as a linear combination of weakenings, then $\{p_i = 0 : i \in [m]\}$ must be infeasible because any assignment to the variables that makes all p_i evaluate to 0 also makes all weakenings evaluate to 0. 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 the total coefficient size of a Nullstellensatz proof.

Definition 5. 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 one way to measure the complexity of a Nullstellensatz proof, but it is not the only way. The degree of Nullstellensatz proofs (i.e., the maximum degree of a weakening with nonzero coefficient) has been extensively studied [BIK+94, BP96, Bus96, CEI96, BIK+97, BCE+98]. On the other hand, very little is known about total coefficient size. We believe that total coefficient size is a natural complexity measure that is worth studying 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.

We show the following results, which are to the best of our knowledge the first results on Nullstellensatz total coefficient size. In section 2, 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 $\widetilde{\Omega}\left(\left(\frac{2}{\sqrt{e}}\right)^n\right)$.

We remark that our lower bound also holds for the functional pigeonhole principle, where each pigeon must go to exactly one hole (instead of at least one hole). In section 4, we introduce the ordering principle and show an exponential upper bound on total coefficient size:

Theorem 2. There is a Nullstellensatz proof of the ordering principle on n elements with total coefficient size $2^n - n$.

For both the pigeonhole principle and the ordering principle, we also obtain experimental results by implementing a linear program that computes minimum total coefficient size. We show these experimental results in sections 3 and 5 respectively. Finally, in section 6, we discuss open problems that our work raises.

Note: Another version of our work is available on arXiv [PZ22]. The main differences between this version and the arXiv version are the following.

- This version includes primal and dual solutions that we found experimentally for the linear program computing minimum total coefficient size in sections 3 and 5.
- The arXiv version defines Nullstellensatz total coefficient size in the more general context of systems of polynomial equations over $\mathbb{R}$.
- The arXiv version defines total coefficient size for a stronger proof system called resolution-like proofs and shows that our proof technique for Theorem 1 does not apply.
- The arXiv version defines total coefficient size for sum-of-squares proofs and shows that there is a sum-of-squares proof of the ordering principle with polynomial total coefficient size.

2 Total coefficient size lower bound for the pigeonhole principle

To prove a lower bound on total coefficient size, we use the following lemma, which applies to any system of axioms. First, we introduce the notation:

Definition 6. If $x \in \{0,1\}^N$, we call x an assignment and treat x as a choice of 0/1-values for $x_1, \ldots, x_N$. If W is a weakening and x is an assignment, $W(x)$ denotes W evaluated on x. (Note that $W(x) \in \{0,1\}$ because weakenings are monomials.)

Lemma 1. Let $\{p_i = 0 : i \in [m]\}$ be a system of axioms. Let $D : \{0,1\}^N \to \mathbb{R}$ be any real-valued function on assignments. The total coefficient size of a Nullstellensatz proof is at least

$$\frac{\sum_x D(x)}{\max_W |\sum_x D(x)W(x)|}.$$

To provide some motivation for this equation: if we consider the problem of finding a Nullstellensatz proof that minimizes the total coefficient size, we can write a linear program whose objective value is the minimum coefficient size. (We will show this linear program in section 3.1.) If we take the dual of this linear program, then a dual solution gives a lower bound on total coefficient size, and in fact this lemma is equivalent to weak duality. But we can also prove this lemma directly.

Proof. Let $1 = \sum_W c_W W$ be any Nullstellensatz proof. Then

$$\begin{aligned}
\sum_x D(x) &= \sum_x D(x)1(x) && \text{(where 1 denotes the constant function 1)} \\
&= \sum_x D(x) \sum_W c_W W(x) \\
&= \sum_W c_W \sum_x D(x)W(x) \\
&\leq \sum_W |c_W| \left| \sum_x D(x)W(x) \right|.
\end{aligned}$$

Dividing both sides by $\max_W |\sum_x D(x)W(x)|$ gives the desired lower bound on $\sum_W |c_W|$. □

2.1 Construction of D

Lemma 1 applies to any system of axioms, but in the rest of this section we will consider the pigeonhole principle in particular. We will construct a function D. In our construction, the only assignments x for which $D(x) \neq 0$ are those where each pigeon goes to exactly one hole: 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.)

Definition 7. In the rest of this section, when we refer to assignments, we refer specifically to assignments where each pigeon goes to exactly one hole. We let $\mathbb{E}$ denote expectation over a uniform assignment (i.e., the expectation when each of the n pigeons goes to one of the $n-1$ holes uniformly at random).

We rewrite Lemma 1 in terms of $\mathbb{E}$.

Corollary 1. Let D be any real-valued function on assignments. The total coefficient size of a Nullstellensatz proof of PHP_n is at least

$$\frac{\mathbb{E}(D)}{\max_W |\mathbb{E}(DW)|}.$$

Because we only consider assignments where each pigeon goes to exactly one hole, we can also restrict the class of weakenings we consider.

Definition 8. In the rest of this section, when we refer to weakenings, we refer specifically to weakenings of axioms of the form $x_{i_1,j}x_{i_2,j} = 0$ for pigeons i_1, i_2 and hole j. (i.e., we can ignore the axioms that constrain that each pigeon goes to at least one hole, because $\mathbb{E}(DW) = 0$ for any weakening W of these axioms.)

Now we provide intuition for our construction of D. Consider the following functions on assignments, which indicate whether a subset of pigeons goes to different holes.

Definition 9. 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 go to different holes, and $J_S(x) = 0$ otherwise.

If there were assignments x where all n pigeons go to different holes, we could get an arbitrarily good lower bound on total coefficient size–but of course there are no such assignments. Perhaps the next best thing would be that if we consider a subset S of $n-1$ pigeons, then D behaves as if those $n-1$ pigeons go to different holes. 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 $\overline{x_{i,j}}$). Our idea is to construct D so that

$$\mathbb{E}(Dp) = \mathbb{E}\left(J_{[n]\setminus\{i\}}p\right).$$

Thus, as far as p is concerned, D behaves as if the pigeons $[n]\setminus\{i\}$ go to different holes. We will construct D that satisfies this equation for all pigeons i and for all monomials p that do not depend on pigeon i, but we note that it is far from obvious that this property necessarily guarantees a good lower bound on total coefficient size.

Definition 10. Our function D is

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

where

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

The important point is not the actual values of the c_S, but rather the ratios between these values. Namely, observe that

- If $|S| = n-1$, then $c_S = 1$.
- If $|S| = |T|+1$, then
$$c_S = -\frac{n-1}{n-1-|T|}c_T.$$

We now prove that D satisfies the property we intuitively wanted, and the proof will make clear why we chose these ratios in this way.

Lemma 2. If p is a monomial that does not depend on pigeon i (i.e., p does not contain any terms $x_{i,j}$ or $\overline{x_{i,j}}$), then

$$\mathbb{E}(Dp) = \mathbb{E}(J_{[n]\setminus\{i\}}p).$$

Proof. Without loss of generality, suppose p does not depend on pigeon 1. Let $T \subsetneq \{2, \ldots, n\}$. Compare $\mathbb{E}(J_T p)$ and $\mathbb{E}(J_{T\cup\{1\}}p)$:

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

because p does not depend on pigeon 1, and if all pigeons in T go to different holes, then the probability that pigeon 1 goes to a different hole than all of them is $\frac{n-1-|T|}{n-1}$. But this ratio is the negative reciprocal of the ratio between $c_{T\cup\{1\}}$ and c_T, so

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

Thus,

$$\begin{aligned}
\mathbb{E}(Dp) &= \sum_{S\subsetneq[n]} \mathbb{E}(c_S J_S p) \\
&= \mathbb{E}(c_{\{2,\ldots,n\}} J_{\{2,\ldots,n\}} p) + \sum_{T\subsetneq\{2,\ldots,n\}} \big(\mathbb{E}(c_T J_T p) + \mathbb{E}(c_{T\cup\{1\}} J_{T\cup\{1\}} p)\big) \\
&= \mathbb{E}(J_{\{2,\ldots,n\}} p)
\end{aligned}$$

as desired. □

Recall that our lower bound on total coefficient size is $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ (Corollary 1). From Lemma 2, the value of $\mathbb{E}(D)$ follows immediately.

Corollary 2.

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

Proof. Let $p = 1$. By Lemma 2, $\mathbb{E}(D) = \mathbb{E}(J_{\{2,\ldots,n\}}) = (n-1)!/(n-1)^{n-1}$, which is $\widetilde{\Theta}(e^{-n})$ by Stirling's formula. □

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

Now we will prove an upper bound on $\max_W |\mathbb{E}(DW)|$. The key property of D we will use is again Lemma 2. We have the following corollary.

Corollary 3. If W is a weakening that does not depend on some pigeon i, 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 W does not depend on pigeon 1. By Lemma 2, $\mathbb{E}(DW) = \mathbb{E}(J_{\{2,\ldots,n\}}W) = 0$, because if W evaluates to 1, then pigeons 2 and 3 must both go to hole 1. □

We make the following definitions and then state a corollary of Lemma 3.

Definition 11. Given a weakening W, for each pigeon i we define a subset of holes $H_{W,i}$ such that $W(x) = 1$ if and only if each pigeon i goes to one of the holes in $H_{W,i}$. For example, if $n = 4$ and $W = x_{1,1}x_{2,1}\overline{x_{3,1}}$, the figure below shows the subsets $H_{W,i}$.

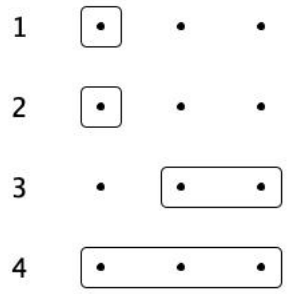

Figure 1: For $W = x_{1,1}x_{2,1}\overline{x_{3,1}}$, $W(x) = 1$ iff each pigeon goes to one of the specified holes

Formally,

- If W contains terms x_{i,j_1} and x_{i,j_2} for distinct holes j_1, j_2, then $H_{W,i} = \emptyset$.
- If W contains exactly one term of the form $x_{i,j}$, then $H_{W,i} = \{j\}$.
- 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 $\overline{x_{i,j}}$.

Definition 12. 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, which is also a weakening of the axiom $x_{i_1,j}x_{i_2,j} = 0$. We define W_S in terms of the subsets $H_{W_S,i}$.

- For each pigeon $i_3 \in S$, $H_{W_S,i_3} = [n-1] \setminus H_{W,i_3}$.
- For each pigeon $i_3 \notin S$, $H_{W_S,i_3} = H_{W,i_3}$.

(If multiple monomials W_S satisfy these properties, we choose one of them arbitrarily.)

Corollary 4. 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

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

Proof. Consider the case $|S| = 1$. The figure below shows the argument when W is the weakening on the left and $S = \{4\}$. The weakening in the middle is $W_{\{4\}}$.

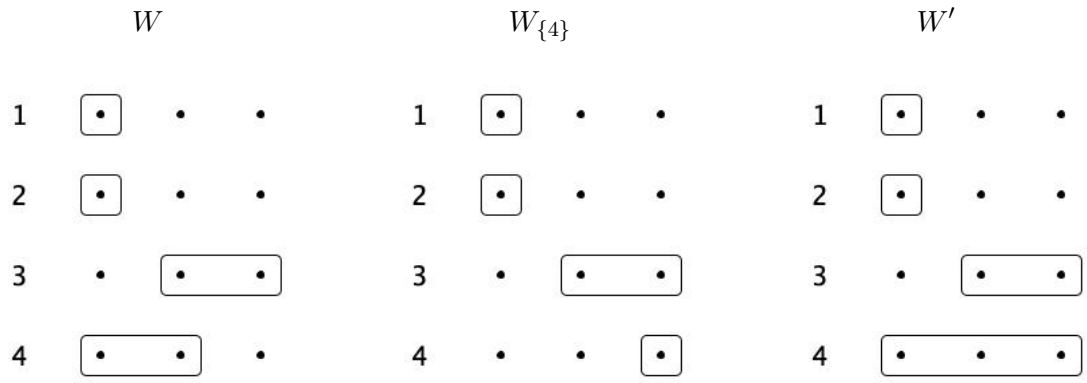

Figure 2: $\mathbb{E}(DW) + \mathbb{E}(DW_{\{4\}}) = \mathbb{E}(DW') = 0$

We have $\mathbb{E}(DW)+\mathbb{E}(DW_{\{4\}}) = \mathbb{E}(DW')$, where W' is the weakening on the right. But $\mathbb{E}(DW') = 0$ by Corollary 3, so $\mathbb{E}(DW_{\{4\}}) = -\mathbb{E}(DW)$. The general case $\mathbb{E}\left(DW_S\right) = (-1)^{|S|}\mathbb{E}(DW)$ follows. □

Using Corollary 4, we will upper 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 4 and Cauchy-Schwarz can be useful. We then modify the approach.

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

By Cauchy-Schwarz,

$$|\mathbb{E}(DW)| \leq \|D\|\|W\|$$

where

$$\|D\| = \sqrt{\mathbb{E}(D^2)},$$
$$\|W\| = \sqrt{\mathbb{E}(W^2)} = \sqrt{\mathbb{E}(W)}.$$

Corollary 4 allows us to bound $\|W\|$. Consider $\max_W |\mathbb{E}(DW)|$. By Corollary 4, instead of taking the maximum over all weakenings W, it suffices to take the maximum over weakenings W such that for all pigeons i, $|H_{W,i}| \leq \lfloor (n-1)/2 \rfloor$. For such weakenings W, we have $\|W\| = \sqrt{\mathbb{E}(W)} \leq \sqrt{2^{-n}} = 2^{-n/2}$.

Recall that our lower bound on total coefficient size is $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ (Corollary 1), and that $\mathbb{E}(D) = \widetilde{\Theta}(e^{-n})$ (Corollary 2). We have

$$\frac{\mathbb{E}(D)}{\max_W |\mathbb{E}(DW)|} \geq \frac{\widetilde{\Theta}(e^{-n})}{\|D\|2^{-n/2}}.$$

Thus, in order to obtain an exponential lower bound on total coefficient size, we would need $\|D\| \leq O(c^{-n})$ for some $c > e/\sqrt{2}$. However, this requirement is too strong:

Lemma 3.

$$\|D\| = \widetilde{\Theta}((\sqrt{e})^{-n}).$$

We defer the proof to Appendix A. The approach we showed here does not yield the desired lower bound on total coefficient size, but it does show the main tools we will use: Corollary 4, Cauchy-Schwarz, and Lemma 3. Now we show how to apply these tools in a better way.

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

Definition 13. 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, choose one of those axioms arbitrarily.) We define the function $W^{\{-1,0,1\}}$ on assignments:

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

Lemma 4. $W^{\{-1,0,1\}}$ takes values in $\{-1, 0, 1\}$. Thus, $\|W^{\{-1,0,1\}}\| \leq 1$.

Proof. If x is an assignment where pigeons i_1 and i_2 do not both go to hole j, then $W^{\{-1,0,1\}}(x) = 0$. Otherwise, exactly one of the W_S evalutes to 1 on x. Namely, if S is the set of pigeons $i_3 \in [n] \setminus \{i_1, i_2\}$ such that pigeon i_3 does not go to H_{W,i_3}, then $W_S(x) = 1$, and $W_{S'}(x) = 0$ for all other $S' \subseteq [n] \setminus \{i_1, i_2\}$. □

Finally, we prove our lower bound on total coefficient size (Theorem 1). For any weakening W, we have

$$\begin{aligned}
\mathbb{E}\left(DW^{\{-1,0,1\}}\right) &= \sum_{S \subseteq [n] \setminus \{i_1,i_2\}} (-1)^{|S|} \mathbb{E}\left(DW_S\right) \\
&= \sum_{S \subseteq [n] \setminus \{i_1,i_2\}} (-1)^{|S|}(-1)^{|S|} \mathbb{E}(DW) && \text{(Corollary 4)} \\
&= 2^{n-2}\mathbb{E}(DW).
\end{aligned}$$

Therefore, we obtain a lower bound on total coefficient size of

$$\begin{aligned}
\frac{\mathbb{E}(D)}{\max_W |\mathbb{E}(DW)|} &= \frac{\mathbb{E}(D)}{2^{-(n-2)} \max_W \left|\mathbb{E}\left(DW^{\{-1,0,1\}}\right)\right|} \\
&\geq \frac{\widetilde{\Theta}(e^{-n})}{2^{-(n-2)} \max_W (\|D\| \|W^{\{-1,0,1\}}\|)} && \text{(Corollary 2, Cauchy-Schwarz)} \\
&\geq \widetilde{\Omega}\left(\left(\frac{2}{\sqrt{e}}\right)^n\right), && \text{(Lemma 3, Lemma 4)}
\end{aligned}$$

completing the proof of Theorem 1.

3 Experimental results for the pigeonhole principle

3.1 Linear program to compute minimum total coefficient size

As mentioned in the discussion following Lemma 1, given a system of axioms $\{p_i = 0 : i \in [m]\}$, we can express the problem of minimizing total coefficient size as a linear program. We show the primal and dual linear programs below, followed by some results for the pigeonhole principle for small values of n that may be of interest. The primal is 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

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

Objective:

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

We now derive the dual linear program by having a dual variable for each primal constraint and vice versa.

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

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

We note that weak duality directly implies Lemma 1, because given any function $D : \{0,1\}^N \to \mathbb{R}$, we can obtain a dual solution by setting $v_x = D(x)/\max_W |\sum_x D(x)W(x)|$.

3.2 Minimum total coefficient size for the pigeonhole principle

For the pigeonhole principle, for small n we computed the dual values shown below. The first column is the optimal dual value for $n = 3, 4$. The second column is the optimal dual value for $n = 3, 4, 5, 6$ under the restriction that the only nonzero v_x are assignments x where each pigeon goes to exactly one hole. The third column is $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ for our function D, which is the dual value we obtain by setting $v_x = D(x)/\max_W |\sum_x D(x)W(x)|$.

n	dual value	dual value, exactly one hole	$\mathbb{E}(D)/\max_W \lvert\mathbb{E}(DW)\rvert$
3	11	6	4
4	$41.4\overline{69}$	27	18
5	-	100	64
6	-	293.75	210.674...

In general, the optimal primal and dual solutions we obtain from the linear program are hard to understand, with many variables having messy decimal values. (We note that even if a clean optimal solution exists, the linear program might not find it.) We were able to find clean dual solutions in some cases, which we now describe. The solutions are invariant under permutations of pigeons and holes (so each point below may correspond to multiple assignments). For $n = 3$, the following dual solution has value 11.

- If all $x_{i,j}$ are 0, then $v_x = -1$.
- If $x_{1,1} = x_{2,2} = 1$ and all other $x_{i,j}$ are 0, then $v_x = 1$.
- If $x_{1,1} = x_{2,1} = x_{3,2} = 1$ and all other $x_{i,j}$ are 0, then $v_x = 1$.

For $n = 4$, we were not able to find a clean dual solution with value $41.4\overline{69}$, but the following dual solution has value 40.

- If all $x_{i,j}$ are 0, then $v_x = 1$.
- If $x_{1,1} = x_{2,2} = 1$ and all other $x_{i,j}$ are 0, then $v_x = -\frac{1}{3}$.
- If $x_{1,1} = x_{2,2} = x_{3,3} = 1$ and all other $x_{i,j}$ are 0, then $v_x = 1$.
- If $x_{1,1} = x_{2,1} = x_{3,2} = x_{4,3} = 1$ and all other $x_{i,j}$ are 0, then $v_x = 1$.
- If $x_{1,1} = x_{2,1} = x_{3,2} = x_{4,2} = 1$ and all other $x_{i,j}$ are 0, then $v_x = -\frac{1}{2}$.

If we add the restriction that each pigeon goes to exactly one hole, our function D seems close to optimal, but we also found other clean dual solutions that may be of interest. We will describe these solutions in the following way. When each pigeon goes to exactly one hole, up to permutations of pigeons and holes, an assignment x can be represented as a tuple of integers $(a_1, a_2, \ldots, a_n)$, where a_i is the number of holes with exactly i pigeons. For $n = 4$, the following dual solution has value 27.

- If $a_2 = 2$, then $v_x = -\frac{1}{2}$.
- If $a_1 = 2, a_2 = 1$, then $v_x = 1$.

For $n = 5$, the following dual solution has value 100.

- If $a_1 = 1, a_2 = 2$, then $v_x = -\frac{1}{6}$.
- If $a_1 = 3, a_2 = 1$, then $v_x = \frac{2}{3}$.

For $n = 6$, we were not able to find a clean dual solution with value 293.75, but the following dual solution has value $283.\overline{3}$.

- If $a_2 = 3$, then $v_x = \frac{7}{108}$.
- If $a_1 = 2, a_2 = 2$, then $v_x = -\frac{13}{216}$.
- If $a_1 = 4, a_2 = 1$, then $v_x = \frac{11}{36}$.

3.3 Conjecture on $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$

We conjecture that $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ is $\widetilde{\Theta}(2^n)$ (whereas the lower bound we proved in Theorem 1 is $\widetilde{\Omega}\left(\left(\frac{2}{\sqrt{e}}\right)^n\right)$). 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 and holes without changing $|\mathbb{E}(DW)|$, and we can replace W by W_S 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 holds for general n, then it appears that $\mathbb{E}(D)/\max_W |\mathbb{E}(DW)|$ is $\widetilde{\Theta}(2^n)$, but we do not have a proof of this pattern or this implication.

4 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 $2^n - n$. Intuitively, the ordering principle says that any well-ordering on n elements must have a minimum element. Formally:

Definition 14. For $n \geq 1$, we define ORD_n to be the following system of axioms.

- For each pair of elements $i, j \in [n]$ with $i < j$, we have a variable $x_{i,j}$. $x_{i,j} = 1$ represents element i being less than element j, and $x_{i,j} = 0$ represents element i being more than element j. We write $x_{j,i}$ as shorthand for $\overline{x_{i,j}}$.
- For each element $i \in [n]$, 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 of elements $i, j, k \in [n]$ 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 is 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 transitivity weakenings (i.e., weakenings of transitivity axioms) W that can have $c_W = 1$ are of the following form.

Definition 15. Let W be a transitivity weakening for elements $\{i, j, k\}$. We define $G(W)$ to 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 transitivity weakening for elements $\{i, j, k\}$, 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 $W(x) = 0$ because every vertex in $G(W)$ has in-degree at least 1.

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 weakening of a transitivity axiom 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 is easy to verify that $\sum_W c_W W = 1$ (i.e., for each assignment x, there is exactly one weakening W such that $c_W = 1$ and $W(x) = 1$) and that this Nullstellensatz proof satisfies the properties of Theorem 3.

Now assume 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 a Nullstellensatz proof for ORD_{n+1} by setting the following c_W to 1 and all other c_W to 0. We will refer to Figure 3 below, where the first row is S_3 and the other three rows are the transitivity weakenings we obtain for ORD_4.

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 $c_{W'} = 1$. (See the second row in Figure 3.)
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 $x_{i+1,j+1}$. Then, we define $W''' = W'' \cdot x_{n+1,1}$ and set $c_{W'''} = 1$. (See the third row in Figure 3.)
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:

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

In other words, $W(x) = 1$ if and only if $A(x) = 1$ and i is the minimum element among the elements $\{2, \ldots, n\}$. (See the fourth row in Figure 3.)

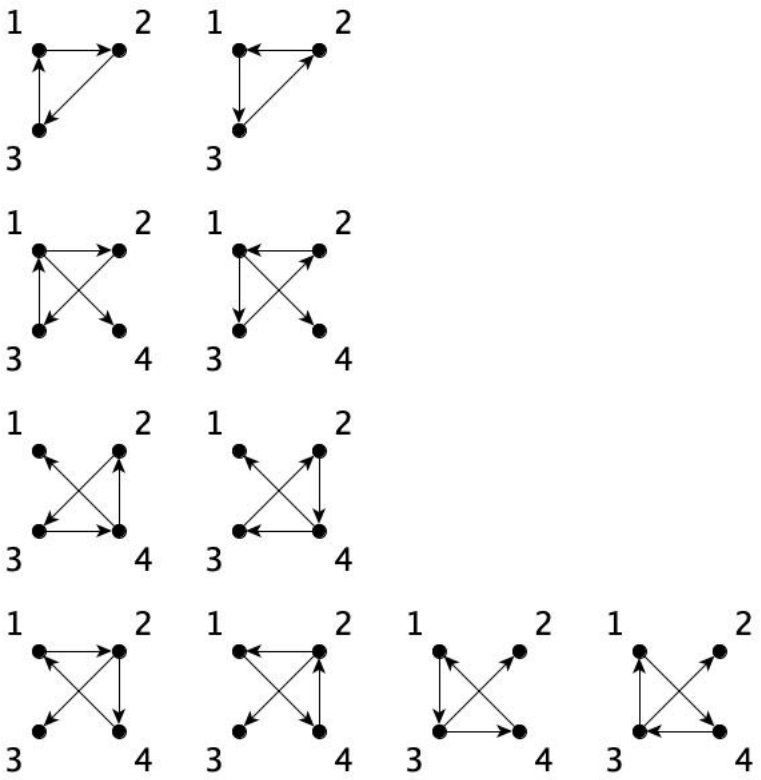

Figure 3: Transitivity weakenings for ORD_3 and ORD_4

The desired properties 1 through 4 in Theorem 3 can easily 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 $i \in \{2, \ldots, n\}$, i is a minimum element among the elements $\{1, \ldots, n\}$ and $x_{n+1,i} = 1$ (because i is not a minimum element among the elements $[n+1]$). 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. □

5 Experimental results for the ordering principle

5.1 Minimum total coefficient size

Using the linear program we described in section 3.1, we computed the minimum total coefficient size for the ordering principle for $n = 3, 4, 5, 6$. 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 $n = 6$, the minimum total coefficient size is 52 instead of $2^6 - 6 = 58$. For $n = 6$, a somewhat clean optimal dual solution is the following. Up to permutations of elements, there are 7 types of assignments x for which $v_x \neq 0$.

1. For all pairs $i < j \in [6]$, $x_{i,j} = 1$. $v_x = 1/24$.
2. $x_{1,2} = x_{2,3} = x_{3,1} = 1$. For $j \in \{4, 5, 6\}$, for all $i < j$, $x_{i,j} = 1$. $v_x = 1/6$.
3. Element 1 is more than all other elements. $x_{2,6} = x_{3,6} = 1$, and for all other pairs $i < j \in \{2, \ldots, 6\}$, $x_{i,j} = 0$. $v_x = -1/24$.
4. The same as case 3, except element 1 is less than all other elements. $v_x = -1/24$.
5. Element 1 is more than all other elements. $x_{2,5} = x_{2,6} = x_{3,6} = 1$, and for all other pairs $i < j \in \{2, \ldots, 6\}$, $x_{i,j} = 0$. $v_x = 1/24$.
6. The same as case 5, except element 1 is less than all other elements. $v_x = 1/24$.
7. $x_{1,6} = x_{2,5} = x_{3,5} = x_{3,6} = x_{5,6} = 1$, and for all other pairs $i < j \in [6]$, $x_{i,j} = 0$. $v_x = 1/24$.

5.2 Minimum total coefficient size under the restriction of no minimum element

When analyzing the pigeonhole principle, we considered restricting assignments so that each pigeon goes to exactly one hole (in which case we only need to consider the axioms $x_{i_1,j}x_{i_2,j} = 0$ for pigeons i_1, i_2 and hole j). For the ordering principle, it is natural to try restricting assignments so that there is no minimum element (in which case we only need to consider transitivity axioms). In fact, the following lemma shows that the minimum total coefficient size decreases by at most a factor of $O(n)$ under this restriction. (For the pigeonhole principle, we do not know of such a relationship.)

Lemma 5. Suppose we have coefficients c_W satisfying $\sum_W c_W W(x) = 1$ for all assignments x that have no minimum element (but it is possible that $\sum_W c_W W(x) \neq 1$ if x has a minimum element). Then there exists a Nullstellensatz proof $\sum_W c'_W W = 1$ with

$$\sum_W |c'_W| \leq (n+1)\left(\sum_W |c_W|\right) + n.$$

Proof. Let C denote the function on weakenings that maps W to c_W. For $i \in [n]$, we will define a function C_i on weakenings satisfying:

- If x is an assignment where i is a minimum element, then $\sum_W C_i(W)W(x) = \sum_W C(W)W(x)$.
- Otherwise, $\sum_W C_i(W)W(x) = 0$.

Let A_i be the non-minimality axiom for element i. Intuitively, we want to define C_i as follows. If W is a weakening that is not of the form $W = A_i W'$ for some weakening W', then $C_i(W) = 0$. Otherwise, $C_i(A_i W') = C(W')$. The only technicality is that multiple weakenings W' may become the same when multiplied by A_i (because all variables are Boolean), so we actually define $C_i(A_i W') = \sum_{W'' : A_i W' = A_i W''} C(W'')$.

Finally, we use the functions C_i to define the function C':

$$C' = C - \left(\sum_{i=1}^{n} C_i \right) + \left(\sum_{i=1}^{n} A_i \right).$$

By taking $c'_W = C'(W)$, the c'_W are a Nullstellensatz proof with the desired upper bound on total coefficient size. □

Under the restriction that assignments have no minimum element, for $n = 3, 4, 5, 6$, the optimal value of the linear program equals the number of transitivity axioms, $2\binom{n}{3}$. For $n = 7$, we did not compute the exact value of the linear program, but we found that the value must be more than the number of transitivity axioms. For $n = 3, 4, 5, 6$, there are somewhat clean primal solutions with value equal to the number of transitivity axioms, which we describe below. (There is a trivial dual solution with this value: for each transitivity axiom, set $v_x = 1$ for an assignment x that satisfies all other axioms.) The solutions are invariant under permutations of elements. For $n = 3$, the following primal solution has value 2.

- If $W = x_{1,2}x_{2,3}x_{3,1}$, then $c_W = 1$.

For $n = 4$, the following primal solution has value 8.

- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}$, then $c_W = \frac{1}{3}$.

For $n = 5$, the following primal solution has value 20.

- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}$, then $c_W = \frac{1}{15}$.
- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}x_{1,5}$, then $c_W = \frac{1}{15}$.
- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}x_{2,4}x_{2,5}$, then $c_W = \frac{1}{30}$.
- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}x_{1,5}x_{2,4}x_{2,5}$, then $c_W = \frac{1}{15}$.

For $n = 6$, the following primal solution has value 40.

- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}$, then $c_W = \frac{1}{45}$.
- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}x_{1,5}x_{1,6}$, then $c_W = \frac{1}{15}$.
- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}x_{2,5}x_{2,6}$, then $c_W = \frac{1}{45}$.
- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}x_{1,5}x_{2,4}x_{2,6}$, then $c_W = \frac{1}{90}$.
- If $W = x_{1,2}x_{2,3}x_{3,1} \cdot x_{1,4}x_{1,5}x_{1,6}x_{2,4}x_{2,5}$, then $c_W = \frac{1}{45}$.

6 Open problems

We proved an exponential lower bound on total coefficient size for the pigeonhole principle (Theorem 1) and an exponential upper bound on total coefficient size for the ordering principle (Theorem 2). To the best of our knowledge, these are the first results on Nullstellensatz total coefficient size. Our work leaves many open problems, 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 variants of the pigeonhole principle, such as when the number of pigeons exceeds the number of holes by more than 1, or we require that each hole must have at least one pigeon?
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.

[PZ22] Aaron Potechin and Aaron Zhang. Bounds on the total coefficient size of nullstellensatz proofs of the pigeonhole principle and the ordering principle. *arXiv preprint arXiv:2205.03577*, 2022.

A Proof of Lemma 3

Proof. Recall the definition of D (Definition 10):

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

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

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

For $S, T \subsetneq [n]$,

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

Therefore,

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

The product of $(-1)^{n-1-|S|}$ (from the c_S) and $(-1)^{n-1-|T|}$ (from the c_T) equals $(-1)^{|S|-|T|}$. Substituting the values of c_S and c_T,

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

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

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

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. 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

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

The sum is an alternating series where the magnitudes of the terms decrease as c decreases. The two largest magnitude terms are $\frac{1}{(n-1)!}$ and $-\frac{1}{2(n-1)!}$. Therefore, the sum is $\Theta(\frac{1}{(n-1)!})$. By Stirling's formula, we conclude that $\mathbb{E}(D^2) = \widetilde{\Theta}(e^{-n})$, and $\|D\| = \widetilde{\Theta}((\sqrt{e})^{-n})$, as desired. □