Separations Above TFNP from Sherali-Adams Lower Bounds

Noah Fleming
Lund University, Sweden
Columbia University, New York, NY, USA

Anna Gál
University of Texas at Austin, TX, USA

Deniz Imrek
University of Texas at Austin, TX, USA

Christophe Marciot
Lund University, Sweden

Abstract

Unlike in TFNP, for which there is an abundance of problems capturing natural existence principles which are incomparable (in the black-box setting), Kleinberg et al. [15] observed that many of the natural problems considered so far in the second level of the total function polynomial hierarchy ($\mathsf{TF}\Sigma_2$) reduce to the Strong Avoid problem. In this work, we prove that the Linear Ordering Principle does not reduce to Strong Avoid in the black-box setting, exhibiting the first $\mathsf{TF}\Sigma_2$ problem that lies outside of the class of problems reducible to Strong Avoid.

The proof of our separation exploits a connection between total search problems in the polynomial hierarchy and proof complexity, recently developed by Fleming, Imrek, and Marciot [9]. In particular, this implies that to show our separation, it suffices to show that there is no small proof of the Linear Ordering Principle in a Σ_2-variant of the Sherali-Adams proof system. To do so, we extend the classical pseudo-expectation method to the Σ_2 setting, showing that the existence of a Σ_2 pseudo-expectation precludes a Σ_2 Sherali-Adams proof. The main technical challenge is in proving the existence of such a pseudo-expectation, we manage to do so by solving a combinatorial covering problem about permutations. We also show that the extended pseudo-expectation bound implies that the Linear Ordering Principle cannot be reduced to any problem admitting a low-degree Sherali-Adams refutation.

2012 ACM Subject Classification Theory of computation → Proof complexity; Theory of computation → Problems, reductions and completeness

Keywords and phrases TFNP, Range Avoidance, Linear Ordering Principle, Separation, Sherali-Adams, Pseudo-Expectation

Digital Object Identifier 10.4230/LIPIcs.CCC.2026.37

Funding *Noah Fleming*: Supported by the Swedish Research Council under grant number 2025-06762.

Acknowledgements We thank the anonymous referees for helpful comments.

1 Introduction

In recent years total search problems in the second level of the polynomial hierarchy ($\mathsf{TF}\Sigma_2$) have received considerable attention. A substantial reason for this is that this class contains a variety of important explicit construction problems, such as finding truth tables of functions with high circuit complexity, pseudo-random generators, rigid matrices, time-bounded Kolmogorov random strings, and extractors. The totality of these problems is witnessed by union-bound type arguments, which can be formalized as reductions to the $\mathsf{TF}\Sigma_2$ problem Avoid [15, 17].

41st Computational Complexity Conference (CCC 2026).
Editor: Dana Moshkovitz; Article No. 37; pp. 37:1–37:23


Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Avoid

Given $C : [2^n] \to [2^{n+1}]$ output $y \in [2^{n+1}]$ such that for every $x \in [2^n]$, $C(x) \neq y$.

By developing new algorithms for this problem, as well as the harder Linear Ordering Principle, researchers have obtained state-of-the-art circuit lower bounds [4, 18, 20] and data structure lower bounds [14, 19].

Linear Ordering Principle (LOP)

Given $\prec$: $[2^n] \times [2^n] \to \{0, 1\}$, output either

- x such that for every $y \neq x$, $x \prec y$, or *(Minimal element)*
- $x \neq y \neq z$ such that either (i) $x \prec x$, or (ii) $x \not\prec y$ and $y \not\prec x$, or (iii) $x \prec y, y \prec z$ but $x \not\prec z$.

(Linear ordering violation)

Kleinberg et al. [15] initiated the study of $\mathsf{TF}\Sigma_2$ as a class, introducing many of the aforementioned explicit construction problems along with a number of other natural search problems, including Ramsey-Erdős Completion which captures the proof that the n-th Ramsey number is at least $2^{n/2}$. In doing so they observed that, while TFNP has a variety of incomparable subclasses capturing various existence principles, every $\mathsf{TF}\Sigma_2$ problem that they considered admitted a reduction to the following strong variant of the Avoid problem.

Strong Avoid

Given $C : [2^n] \to [2^n + 1]$ output $y \in [2^n + 1]$ such that for every $x \in [2^n]$, $C(x) \neq y$.

[15] raised the question of whether there are any natural problems in $\mathsf{TF}\Sigma_2$ which do not reduce to the problem StrongAvoid. The main contribution of our work is to give the first $\mathsf{TF}\Sigma_2$ problem which is not contained within $\mathsf{StrongAvoid}^{dt}$, the class of problems which admit efficient black-box reductions to StrongAvoid. This answers the question in the black-box setting. Note that a separation in the Turing Machine setting would separate P from Σ_2^P.

▶ **Theorem 1.** LOP $\notin \mathsf{StrongAvoid}^{dt}$.

This complements the work of Korten and Pitassi [18], who show a separation in the other direction. Using our technique, we are also able to show that LOP does not reduce to the $\mathsf{TF}\Sigma_2$ problem LeastNumber, introduced by Thapen [23].

Least Number

Given $f : \{0, 1\}^n \to \{0, 1\}$, output either $\perp$ if $f(x) = 0$ for all x, or x such that $f(x) = 1$ and for all $y < x$, $f(y) = 0$.

▶ **Theorem 2.** LOP $\notin \mathsf{LeastNumber}^{dt}$.

Our separations, along with those known already in the literature, are represented pictorially in Figure 1.

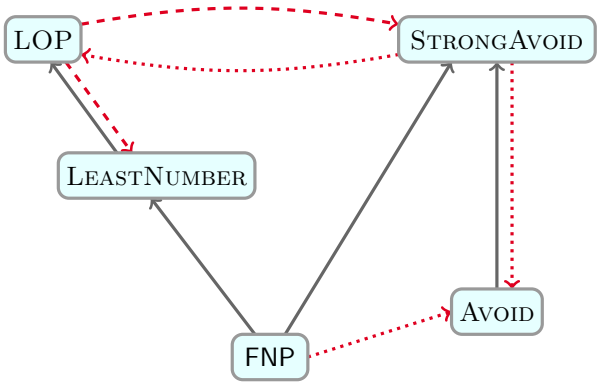

Figure 1 Relationships of some $\mathsf{TF}\Sigma_2^{dt}$ classes. A black arrow from a class A to a class B means that $A \subseteq B$ ([15]). A dashed or dotted arrow from a class A to a class B means that $A \not\subseteq B$. The dashed separations are proved in this paper. [18] proves that StrongAvoid does not reduce to Avoid and LOP. [7] proves that Avoid^{dt} does not contain all of TFNP^{dt}. This implies that none of the other problems in the diagram reduces to Avoid.

1.1 Technical Highlights

To prove Theorem 1 we make use of a connection between $\mathsf{TF}\Sigma_2^{dt}$ (where the superscript dt denotes black-box reductions) and proof complexity, developed in [9]. In doing so we provide the first proof of a separation between black-box $\mathsf{TF}\Sigma_2$ classes that uses proof complexity, confirming that this connection can indeed be useful. In particular, [9] shows that black-box reductions to StrongAvoid are equivalent to efficient proofs in a Σ_2-variant of the Sherali-Adams proof system (see Definition 8) which equips it with a Σ_2-weakening step, extending the weakening rule for resolution to depth-2 formulas, as follows.

▶ **Definition 3.** *Let D be a DNF formula, a* Σ_2-weakening *of D is a collection of DNFs $\{D_i\}_{i\in[m]}$ such that $D \implies D_i$ for every $i \in [m]$. A Σ_2-weakening of a collection of DNFs is a union of Σ_2-weakening of those DNFs.*

This connection with proof complexity reduces proving Theorem 1 to showing that the unsatisfiable formula which encodes the totality of LOP does not have efficient Σ_2-Sherali-Adams proofs (Definition 8).

Pseudo-Expectations

To prove our Sherali-Adams lower bound, we extend the well-developed pseudo-expectation technique for Sherali-Adams proofs to the Σ_2 setting. A pseudo-expectation is an object which appears to be a probability distribution over satisfying assignments to our (unsatisfiable) formula when examining low-degree marginals. The existence of a pseudo-expectation precludes a low-degree Sherali-Adams proof. A Σ_2-pseudo-expectation for LOP is a collection of pseudo-expectations, one for each set of weakenings of the constraints of LOP. See Definition 10 for details. The main technical challenge is to construct such a Σ_2-pseudo-expectation. While Dantchev et al. [5] showed that LOP itself has a high-degree pseudo-expectation, it is not clear how to extend their argument to all Σ_2-weakenings.

A Covering Problem of Permutations

We observe that constructing a Σ_2-pseudo-expectation requires us to solve the following covering problem. Let Ord be the set of all total orders on the set of numbers $\{1, \ldots, n\}$, and let $\text{Ord}^{*1} \subseteq \text{Ord}$ be the set of total orders for which at least one number comes before the element 1. For a fixed $d \leq n$, a set $S \subseteq [n]$ of size d and an order σ on S, let $\mathcal{C}_{S,\sigma} = \{\pi \in \text{Ord} \mid \pi \text{ induces the order } \sigma \text{ on } S\}$ be the set of all total orders on $[n]$ that are consistent with σ on S. Is there a set of strictly less than $d!$ pairs (S_i, σ_i), with each S_i of size d such that Ord^{*1} is contained in the union $\bigcup_i \mathcal{C}_{S_i,\sigma_i}$?

Note that one can cover the set of all total orders with exactly $d!$ such collections $\mathcal{C}_{S,\sigma}$. The question is thus "can we do better if we do not need to cover orders starting with 1?". One can extract from the lower bound of Dantchev et al. [5] that for $d = 2$, such coverings are impossible. Our proof of Theorem 1 implies that the answer to this question is no for any $d \leq n/100$. In our argument, we show that weakenings of some axioms of LOP can be modified to a normalized form without increasing their degree by much. Proving that our pseudo-expectation function is non-negative on these normalized weakenings is equivalent to a lower bound in the above covering problem.

Our linear lower bound on d is tight up to constant factors, as one can construct a covering with less than $d!$ collections for $d = n/2$.

Further Separations from Sherali-Adams Upper Bounds

Leveraging our Σ_2-pseudo-expectation, we show that if a $\mathsf{TF}\Sigma_2$ problem admits a Sherali-Adams upper bound then this problem is separated from the Linear Ordering Principle. Together with a simple Sherali-Adams upper bound for the LeastNumber we obtain Theorem 2. To prove this, we show that reductions can be split in two parts: first, a potentially hard to verify Σ_2-weakening step, followed by a *counter-example reduction*, a type of reduction introduced in [16, 23] that actually happens in TFNP^{dt}. This reveals how the two types of black-box reductions that have been studied in the literature for total functions in the polynomial hierarchy relate.

1.2 Related Works

Ghentiyala et al. [11] showed that LOP does not reduce to Avoid, even in the non-blackbox setting, assuming that $\mathsf{P}^{\mathsf{prMA}} \not\subset \mathsf{AM} \cap \mathsf{coAM}$. The only other known separation in $\mathsf{TF}\Sigma_2^{dt}$ was proven by Korten and Pitassi [18] who showed that the problem StrongAvoid $\notin \mathsf{LOP}^{dt}$. Their separation was obtained by proving a lower bound for depth-3 circuits via a switching lemma. This contrasts with our separation which uses proof complexity techniques.

Proof complexity techniques have been the main method for obtaining separations between classes in TFNP in the black-box setting. This has led to a complete understanding of the relationships between the major classes [1, 2, 8, 12, 13, 21]. Proof complexity has been useful in TFNP because of the fact that membership in (sufficiently uniform) TFNP^{dt} subclasses is equivalent to efficient provability in some associated proof system [3]. This has been used to show that, for example, the class PPADS^{dt} is equivalent to the unary Sherali-Adams proof system [12]. This connection was recently extended to subclasses of the total search problems in the polynomial hierarchy in [9], which is the basis for our work.

Recently, Thapen in [23] exhibited a new type of reductions, namely *counter-example reductions*. They, along with the notion of *herbrandization*, are used to find a TFNP^{dt} translation of what happens at the second level $\mathsf{TF}\Sigma_2^{dt}$ and they can also be used to define

TFNP^{dt} subclasses from $\mathsf{TF}\Sigma_2^{dt}$ problems. In Section 5, we add a bridge between this notion of counter-example reduction and the notion of Σ_2-weakening defined in [9] and show that they are sufficient ingredients to capture reductions.

2 Preliminaries on Black-Box $\mathsf{TF}\Sigma_2$

A *query search problem* is a sequence of relations $R_n \subseteq \{0,1\}^n \times \mathcal{O}_n$, one for each $n \in \mathbb{N}$. It is *total* if for every $x \in \{0,1\}^n$ there is an $o \in \mathcal{O}_n$ such that $(x,o) \in R_n$. We think of $x \in \{0,1\}^n$ as a bit string which can be accessed by querying individual bits, and we will measure the complexity of solving R_n as the number of bits that must be queried. Hence, an efficient algorithm for R_n will be one which finds a suitable o while making at most $\mathsf{polylog}(n)$-many queries to the input. We will not charge the algorithm for other computational steps, and therefore an efficient algorithm corresponds to a shallow decision tree. Total query search problems which can be computed by decision trees of depth $\mathsf{polylog}(n)$ belong to the class FP^{dt}, where dt indicates that it is a black-box class. While search problems are formally defined as sequences $R = (R_n)$, we will often want to speak about individual elements in the sequence. For readability, we will abuse notation and refer to elements R_n in the sequence as total search problems. In addition, we will often drop the subscript n, and rely on context to differentiate.

In this paper we will be considering total query search problems in the second level of the polynomial hierarchy $\mathsf{TF}\Sigma_2^{dt}$.

▶ **Definition 4.** *A total query search problem $R = (R_n)$, where $R_n \subseteq \{0,1\}^n \times \mathcal{O}_n$, belongs to $\mathsf{TF}\Sigma_2^{dt}$ if for every n, $o \in \mathcal{O}_n$ and $x \in \{0,1\}^n$*

$$R_n(x,o) \iff \forall z \in \{0,1\}^\ell, V_{o,z}(x) = 1,$$

where $\ell = \mathsf{polylog}(n)$ and for every $z \in \{0,1\}^\ell$, $V_{o,z}$ is a $\mathsf{polylog}(n)$-depth decision tree.

We can compare the complexity of total search problems by taking reductions between them. The following defines *decision tree reductions* (known as *formulations*) between total query search problems, the query analogue of polynomial-time reductions.

▶ **Definition 5.** *For total query search problems $R \subseteq \{0,1\}^n \times \mathcal{O}_n, S \subseteq \{0,1\}^m \times \mathcal{O}'_m$, there is an S-*formulation of R *if, for every $i \in [m]$ and $o \in \mathcal{O}'_m$, there are functions $f_i : \{0,1\}^n \to \{0,1\}$ and $g_o : \{0,1\}^n \to \mathcal{O}_n$ such that*

$$S(f(x),o) \implies R(x, g_o(x)),$$

where $f(x) = (f_1(x) \dots f_m(x))$. The depth *of the formulation is*

$$d := \max\left(\{\mathsf{depth}(f_i) : i \in [m]\} \cup \{\mathsf{depth}(g_o) : o \in \mathcal{O}'_m\}\right),$$

where $\mathsf{depth}(f)$ denotes the minimum depth of any decision tree which computes f. The size *of the formulation is m, the number of input bits to S. The* complexity *of the formulation is $\log m + d$. The* complexity of reducing R to S *is the minimum complexity of all possible S-formulations of R.*

We extend this definition to sequences in the natural way. If $S = (S_m)$ is a sequence and R_n is a single search problem, then the complexity of reducing R_n to S is the minimum over m of the complexity of reducing R_n to S_m. For two sequences of search problems $S = (S_m)$ and $R = (R_n)$, the complexity of reducing R to S is the complexity of reducing R_n to S, as a function of n. A reduction from R to S is efficient if its complexity is $\mathsf{polylog}(n)$; we denote this by $R \leq_{dt} S$.

Black-box $\mathsf{TF}\Sigma_2$ can be viewed as the study of the following family of total search problems. This will allow us to leverage a close connection between $\mathsf{TF}\Sigma_2^{dt}$ and proof complexity [9] to prove our main result.

▶ **Definition 6.** *Let $F = D_1 \wedge \ldots \wedge D_t$ be a formula in which each D_i is a DNF. The* false formula *search problem* $\mathrm{FF}_F \subseteq \{0,1\}^n \times [t]$ *is defined as*

$$\mathrm{FF}_F(x, o) \iff D_o(x) = 0.$$

▶ **Lemma 7** ([9])**.** *For any $R \in \mathsf{TF}\Sigma_2^{dt}$ there is an unsatisfiable formula $F_R = D_1 \wedge \ldots \wedge D_t$ where each D_i is a DNF of* $\mathsf{polylog}(n)$*-width, such that $R =_{dt} \mathrm{FF}_{F_R}$.*

This extends the well-known connection between TFNP^{dt} and the false clause search problem. The proof of this lemma proceeds by writing down the totality of R as a Σ_3-formula and then taking its negation to obtain F_R. Throughout this paper, we will abuse notation and use R to refer to F_R, using context to differentiate. Hence, $\mathrm{FF}_R = \mathrm{FF}_{F_R}$.

3 Proof Complexity and a Sufficient Condition for Separations

To prove Theorem 1 we make use of a connection between $\mathsf{TF}\Sigma_2^{dt}$ and proof complexity. Fleming, Imrek, and Marciot [9] showed that decision tree reductions between $\mathsf{TF}\Sigma_2^{dt}$ problems are equivalent to efficient proofs in certain proof systems equipped with the Σ_2-weakening rule from Definition 3. Using this connection they showed that proofs in the unary Σ_2-Sherali-Adams proof system are equivalent to reductions to StrongAvoid, which we describe next.

For any boolean formula F, we will assume without loss of generality that all negations occur at the leaves and let $\mathsf{Vars}^+(F)$ be the positive literals in F and $\mathsf{Vars}^-(F)$ be the negative literals. For any conjunct $t = \bigwedge_{x_i \in \mathsf{Vars}^+(t)} x_i \wedge \bigwedge_{x_j \in \mathsf{Vars}^-(t)} \neg x_j$, we associate the following polynomial $t(x) = \prod_{x_i \in \mathsf{Vars}^+(t)} x_i \prod_{x_j \in \mathsf{Vars}^-(t)} (1 - x_j)$. We refer to the polynomials $t(x)$ also as conjuncts and also denote them by t. We say that a *conical junta* is a sum of conjuncts $\mathcal{J} := \sum_i t_i$.

Let $D = \bigvee_{t \in D} t$ be any DNF. We can express D as a polynomial

$$D(x) = \sum_{t \in D} t(x) - 1. \tag{1}$$

The degree of this polynomial is $\deg(D) := \max_{t \in D} \deg(t)$. We refer to $\deg(D)$ as the *degree of the DNF D*. Observe that for any $x \in \{0,1\}^n$, the DNF $D(x)$ is true iff $\sum_{t \in D} t(x) - 1 \geq 0$. We will abuse notation and denote by D both the DNF and the associated polynomial, using context to differentiate.

Throughout we will work with *multi-linear arithmetic*, associating $x_i^2 = x_i$ for every variable x. This has the effect of restricting the underlying linear program to $\{0,1\}$-points.

▶ **Definition 8.** *Let $F = \{D_i\}_{i \in [m]}$ be an unsatisfiable collection of DNFs (a collection with no shared satisfying assignment). A* Σ_2-unary-Sherali-Adams *(denoted* uSA*) proof Π of F is a Σ_2-weakening $F' = \{D'_i\}_{i \in [m']}$ of F together with a list of conical juntas $\mathcal{J}_i, \mathcal{J}$, such that*

$$\sum_{i \in [m']} D'_i \mathcal{J}_i + \mathcal{J} = -1.$$

The degree $\deg(\Pi)$ *is the maximum degree among $D_i, D'_i\mathcal{J}_i$, and $\mathcal{J}$, and the* size $\mathsf{size}(\Pi)$ *is the number of monomials in $D_i, D'_i\mathcal{J}_i, \mathcal{J}$ counted with multiplicity (i.e., each monomial is counted again each time it appears). The* complexity *of the proof is given by* $\mathsf{uSA}(\Pi) := \deg(\Pi) + \log \mathsf{size}(\Pi)$*, and the complexity of proving F is* $\mathsf{uSA}(F) := \min_\Pi \mathsf{uSA}(\Pi)$*, where the minimum is taken over all* uSA *proofs Π of F.*

The only difference between the standard Sherali-Adams proof system and its unary variant is the measure of size; in particular, we cannot use large coefficients. When measuring degree, these systems are identical.

To prove Theorem 1 it suffices to show that there is no efficient Σ_2-uSA proof of LOP.

▶ **Theorem 9** ([9]). *For any $R \in \mathsf{TF}\Sigma_2^{dt}$ there exists a $\mathsf{polylog}(n)$-complexity Σ_2-uSA proof of R iff $R \in \mathsf{StrongAvoid}^{dt}$.*

To prove our lower bound for LOP we generalize the method of pseudo-expectations, which has been developed for Sherali-Adams (see [10] for a survey), to Σ_2-Sherali-Adams degree lower bounds.

▶ **Definition 10.** *For any collection of DNFs $F = \{D_i\}_{i\in[m]}$ a* degree-d pseudo-expectation *for F is a linear function $\tilde{\mathbb{E}} : \mathbb{R}[x] \to \mathbb{R}$ satisfying*
1. $\tilde{\mathbb{E}}[1] = 1$,
2. $\tilde{\mathbb{E}}[\mathcal{J}] \geq 0$ *for every conical junta $\mathcal{J}$ of degree at most d,*
3. $\tilde{\mathbb{E}}[D_i\mathcal{J}] \geq 0$ *for every $D_i \in F$ and conical junta J such that* $\deg(D_i\mathcal{J}) \leq d$.

*A degree-d Σ_2-*pseudo-expectation *for F is a family of degree-d pseudo-expectations $\{\tilde{\mathbb{E}}_G\}$, one for every Σ_2-weakening G of F.*

Note that a Σ_2-pseudo expectation for F is equivalent to a standard pseudo-expectation for the collection of DNFs F' which includes every Σ_2-weakening of the DNFs in F.

▶ **Lemma 11.** *There is a degree-d Σ_2-pseudo-expectation for F iff there is no degree-d Σ_2-uSA proof of F. Moreover, if F does not have a degree-d Σ_2-uSA, then there is a degree-d pseudo-expectation that works for all weakenings of F.*

Proof. The proof follows from the fact that for every unsatisfiable collection of DNFs (and indeed any collection of polynomials with no $\{0,1\}$-solutions), a degree-d pseudo-expectation exists iff a Sherali-Adams proof does not (see e.g. [10]). Applying this to each weakening completes the proof. However, since we will use the forward direction of this theorem we include a proof.

Let $\{\tilde{\mathbb{E}}_G\}$ be a degree-d Σ_2-pseudo-expectation and suppose for contradiction that there is also a degree-d Σ_2-Sherali-Adams proof (G, Π) of F, where $G = \{D'_i\}_{i\in[m']}$ is a Σ_2-weakening of F and Π is a list of conical juntas $\{\mathcal{J}_i\}, \mathcal{J}$ constituting a Sherali-Adams proof of G. Let $\tilde{\mathbb{E}}_G$ be the pseudo-expectation which corresponds to G, then

$$-1 = \tilde{\mathbb{E}}_G[-1] = \tilde{\mathbb{E}}\Big[\sum_{i\in[m']} D'_i\mathcal{J}_i + \mathcal{J}\Big] = \sum_{i\in[m']} \tilde{\mathbb{E}}[D'_i\mathcal{J}_i] + \tilde{\mathbb{E}}[\mathcal{J}] \geq 0.$$

Consider $\mathcal{G} = \bigcup_G G$, where the union is taken over every Σ_2-weakenings of F. Then $\tilde{\mathbb{E}}_{\mathcal{G}}$ must be a d-pseudo-expectation for all weakenings of F, as they are all subsets of $\mathcal{G}$. ◀

4 Separating LOP from Strong Avoid

In this section we prove the following.

▶ **Theorem 12.** *There exists Σ_2-pseudo-expectation for* LOP *of degree* $(\frac{n}{400} - 1)$.

From this, the separation LOP $\notin \mathsf{StrongAvoid}^{dt}$ follows immediately. Note that Theorem 12 is much stronger than we need in order to rule out a reduction to StrongAvoid, as works irrespective of the size of the uSA proofs, or size of the reduction.

Proof of Theorem 1. By Theorem 9 LOP $\in$ StrongAvoid^{dt} iff there exists a polylog(n) complexity Σ_2-uSA proof of the propositional encoding of LOP given by Lemma 7. In particular, such a proof must have polylog(n) degree. However, by Lemma 11, the existence of a degree $\Omega(n)$ Σ_2-pseudo-expectation precludes the existence of such a Σ_2-uSA proof. ◀

In the remainder of this section we prove Theorem 12. Applying Lemma 7 to LOP, we obtain an unsatisfiable formula which consists of the following constraints (axioms) over variables $x_{i,j}$, where $x_{i,j} = 1$ means "$i \prec j$":

- M_i: $\bigvee_{j \in [n] \setminus i} x_{j,i}$, for $i \in n$, *(Non-minimality)*
- R_i: $\neg x_{i,i}$, for $i \in [n]$, *(Irreflexivity)*
- $A_{i,j}$: $\neg x_{i,j} \vee \neg x_{j,i}$, for distinct $i, j \in [n]$, *(Asymmetry)*
- $T_{i,j,k}$: $x_{i,k} \vee \neg x_{i,j} \vee \neg x_{j,k}$, for distinct $i, j, k \in [n]$, *(Transitivity)*
- $O_{i,j}$: $x_{i,j} \vee x_{j,i}$, for distinct $i, j \in [n]$. *(Totality)*

Note that even though each of these axioms could be expressed as a clause, we consider them as DNFs of degree 1. We will use the same pseudo-expectation for each weakening of LOP, which will be the standard choice [5]: a uniform distribution over total orders. Let $\mathrm{Ord} = \{x \in [n]^n \mid \forall i \neq j, x_i \neq x_j\}$ be the set of total orders on $[n]$. We associate total orders on $[n]$ with both strings in $[n]^n$ (simply listing the elements in the given order) and bijective maps $[n] \to [n]$. For $z \in \mathrm{Ord}$, and $T \subseteq [n]$, we define $z{\restriction}T$ as the permutation on T, $\pi : |T| \to T$, that orders those elements in the same order as z does. We denote $z{\restriction}([n] \setminus T)$ by $z \setminus T$ for convenience. For $x \in \mathrm{Ord}$ and a monomial t we denote by $t(x)$ the monomial t evaluated over the variables $x_{i,j}$, where $x_{i,j} = 1$ means "$i \prec j$" (as above) in the order x.

▶ **Definition 13.** *We define our pseudo-expectation* $\tilde{\mathbb{E}}$ *on monomials* t *as*

$$\tilde{\mathbb{E}}\,[t] = \frac{|\{x \in \mathrm{Ord} \mid t(x) = 1\}|}{|\,\mathrm{Ord}\,|}.$$

The definition is extended to arbitrary polynomials by linearity.

By our encoding of DNFs as polynomials (Equation 1), for a DNF $D = \bigvee_i t_i$, we use the notation $\tilde{\mathbb{E}}\,[D]$ to mean $\tilde{\mathbb{E}}\,[\sum_i t_i - 1]$. Note that $\tilde{\mathbb{E}}\,[t + t']$ corresponds to the number of orders either t or t' accepts, plus the number of orders they both accept (all of this divided by $|\,\mathrm{Ord}\,|$). As an example,

$$\tilde{\mathbb{E}}\,[M_i] = \tilde{\mathbb{E}}\Big[\sum_{j \in [n] \setminus i} x_{j,i} - 1\Big] = \sum_{j \in [n] \setminus i} \tilde{\mathbb{E}}\,[x_{j,i}] - \tilde{\mathbb{E}}\,[1] = \sum_{j \in [n] \setminus i} \frac{1}{2} - 1 = \frac{n-1}{2} - 1 = \frac{n-3}{2},$$

as exactly half of the total orders are such that $j \prec i$.

When taking the pseudo-expectation of a sum of terms (or when considering a DNF), we will refer to total orders accepted by more than one term of the sum (or the DNF) as being *over-counted*. If a given order is accepted by $c > 1$ terms, we say that this order contributes $c - 1$ *extra occurrences*. In particular, in the above example, a total order which has i as the 10th element of its order will be accepted by 9 terms, and is thus over-counted. Moreover, 8 of those 9 terms provide extra occurrences of the order.

We will show that our function $\tilde{\mathbb{E}}$ gives a degree-d Σ_2-pseudo-expectation for the set of axioms of LOP for any $d \leq \frac{n}{400} - 1$. We proceed by showing that it satisfies the conditions of Definition 10.

Condition 1. Note that the constant 1 is treated as a monomial such that $1(x) = 1$ for every x, and thus $\tilde{\mathbb{E}}\,[1] = 1$.

Condition 2. Notice that conjuncts containing $x_{i,i}$ are never satisfied by any total order. We can thus always delete those conjuncts without changing the pseudo-expectation. On the other hand, the negated variables $x_{i,i}$ are always true in any total order by the axioms R_i, thus we can simply remove (or replace by constant 1) any negated $x_{i,i}$ without changing the pseudo-expectation.

Additionally we can rewrite any conjunct t over the LOP variables $x_{i,j}$ for $i \neq j$ to an equivalent conjunct t' without using negations, simply replacing any negated variable $x_{i,j}$ by the variable $x_{j,i}$. To see this, notice that for any total order on LOP variables, if $i \neq j$, $x_{j,i}$ holds iff $x_{i,j}$ does not hold, and so the total orders which satisfy t are exactly those that satisfy t'. Thus, $\tilde{\mathbb{E}}[t] = \tilde{\mathbb{E}}[t']$. Observe that the polynomial corresponding to t' is simply a monomial, and $\tilde{\mathbb{E}}[t] = \tilde{\mathbb{E}}[t'] \geq 0$ by definition. Since any conical junta is a sum of conjuncts, this implies that $\tilde{\mathbb{E}}[\mathcal{J}] \geq 0$ for any conical junta $\mathcal{J}$.

Condition 3. It remains to show that condition 3 holds for our function $\tilde{\mathbb{E}}$. Our goal is thus to show that for all conical juntas $\mathcal{J}$, and any weakening D' of any axiom of LOP, $\tilde{\mathbb{E}}[D'\mathcal{J}] \geq 0$, if $\deg(D'\mathcal{J})$ is small enough. First we observe that this easily follows for the axioms other than type M_i. We will handle axioms of type M_i in Section 4.1.

The axioms of LOP that are not of type M_i must hold for every total order. For any such axiom $D = \bigvee_{t \in D} t$ we have $\cup_{t \in D}\{x \in \mathrm{Ord} \,|\, t(x)\} = \mathrm{Ord}$. Furthermore, this holds for any weakening W of D, as W must be satisfied by at least the same assignments as D. Hence, for any conical junta $\mathcal{J}$, we can deduce that

$$\tilde{\mathbb{E}}[W\mathcal{J}] = \left(\sum_{j \in J}\sum_{t \in W} \lambda_j \frac{|\{x \in \mathrm{Ord} \mid t(x)j(x) = 1\}|}{|\,\mathrm{Ord}\,|}\right) - \sum_{j \in J} \lambda_j \frac{|\{x \in \mathrm{Ord} \mid j(x) = 1\}|}{|\,\mathrm{Ord}\,|} \geq 0,$$

where j sums over every conjunct in $\mathcal{J}$, and λ_j is the number of times j appears in $\mathcal{J}$. As every order x must satisfy at least one term t in W, the last inequality holds. This proves that the pseudo-expectation is non-negative for any conical junta and any weakening of an axiom that is not of type M_i, regardless of the degree.

4.1 The No-Minimal-Element Axioms

By symmetry, it suffices to consider weakenings of M_1, weakenings of M_i can be handled analogously. It thus suffices, without loss of generality, to prove that for every low-degree conical junta $\mathcal{J}$, and every low-degree weakening W of M_1, $\tilde{\mathbb{E}}[W\mathcal{J}] \geq 0$. The main difficulty is to handle the diversity of the weakenings of M_1.

We begin by fixing some notation. It will be convenient to go back and forth between total orders and canonical terms representing them. For a set $S \subseteq [n]$ with $|S| \geq 2$, consider an ordering $\pi : [|S|] \to S$ of its elements. We will denote by $[\![\pi(1)\dots\pi(|S|)]\!]$, or $[\![S]\!]_\pi$, the term $x_{\pi(1),\pi(2)}x_{\pi(2),\pi(3)}\cdots x_{\pi(|S|-1),\pi(|S|)}$.

We say that a term *mentions* i if it contains one of the variables $x_{i,j}$ or $x_{j,i}$ for some $j \neq i$. For example, $x_{1,2}x_{2,3}$ mentions $\{1,2,3\}$. We refer to the set of elements that a term mentions as its *support* and the number of such elements as the *support size* of the term.

To simplify our argument, we will first show that we can assume that each DNF has several convenient properties.

▶ **Definition 14.** *We say that a DNF D is* normalized *if*

1. *Every term t of D is of the form $t = [\![S]\!]_\pi$ for some set S which includes 1, and*
2. *All terms have the same support size, which we refer to as the* support size of D.

▶ **Lemma 15.** *Let W be a DNF of degree d. There exists a normalized DNF N of support size $\leq 2d+1$ such that W accepts exactly the same total orders as N does, and $\tilde{\mathbb{E}}[W] = \tilde{\mathbb{E}}[N]$.*

Proof. Let $t \not\equiv 0$ be one of the terms of W. We first transform t into a term t' with no negated literals as follows: we replace each negated literal $\neg x_{i,j}$ in t with $x_{j,i}$, noting that for any total order, $x_{j,i}$ holds iff $x_{i,j}$ does not, and so the total orders satisfying t are exactly those satisfying t'. Hence, $\tilde{\mathbb{E}}[t] = \tilde{\mathbb{E}}[t']$.

Next, we replace t by a collection of terms of the form $[\![S]\!]_\pi$ for some $S \subseteq [n]$ containing 1 (defined below), and a collection of orderings π on S. Let T be the support of t and note that $|T| \leq 2d$ as each variable in t can mention at most two elements. We modify T to define S as follows: Include the element 1 if it was not included yet in T, and pad T with additional elements as necessary, so that the resulting set S has size $|S| = 2d+1$. We use the resulting set S as the support for the collection of terms replacing t.

Let Π be the set of total orders on S that are consistent with t. We claim that for every total order z, $t(z) \Leftrightarrow \bigvee_{\pi\in\Pi}[\![S]\!]_\pi(z)$, and $\tilde{\mathbb{E}}[t] = \tilde{\mathbb{E}}\left[\sum_{\pi\in\Pi}[\![S]\!]_\pi\right]$. In other words, $\bigvee_{\pi\in\Pi}[\![S]\!]_\pi$ accepts exactly the same total orders t does, and each total order accepted by t is consistent with exactly one $\pi \in \Pi$ (thus total orders accepted by t are accepted by a unique term in $\bigvee_{\pi\in\Pi}[\![S]\!]_\pi$). We set N to be the resulting DNF $\bigvee_{t\in W}\bigvee_{\pi\in\Pi_t}[\![S_t]\!]_\pi$. ◀

Note that the transformation in the proof of Lemma 15 can highly increase the number of terms in W, but it does not change its pseudo-expectation, and it at most doubles the degree: Notice that for a normalized DNF with support size k its degree is exactly $k-1$. The support size of the resulting DNF N is at most $2d+1$, thus the degree of N is at most $2d$.

Suppose that M_1 has a weakening W of degree d such that $\tilde{\mathbb{E}}[W] < 0$. This lemma implies that there is also a normalized weakening N of M_1 with $\tilde{\mathbb{E}}[N] < 0$ and of not much larger degree.

▶ **Lemma 16.** *For any normalized DNF N of support size k, every term t of N has $\tilde{\mathbb{E}}[t] = 1/k!$.*

Proof. Let $t = [\![T]\!]_\pi$ be one of the terms of N. As a reminder, $\tilde{\mathbb{E}}[t] = |\{x \in \mathrm{Ord} \mid t(x) = 1\}|/|\mathrm{Ord}|$ where Ord is the set of total orders. Since $t = [\![T]\!]_\pi$, $\tilde{\mathbb{E}}[t] = |\{x \in \mathrm{Ord} \mid x{\restriction}T = \pi\}|/|\mathrm{Ord}|$. To count the number of orders consistent with t, one can simply select the positions of the elements in T, set them in the right order, and then fill in the rest any possible way, thus,

$$|\{x \in \mathrm{Ord} \mid x{\restriction}T = \pi\}| = \binom{n}{k}(n-k)! = \frac{n!}{k!}.$$ ◀

▶ **Lemma 17.** *Let N be normalized DNF with support size k that is a weakening of M_1, and let R be the set of total orders rejected by N. Then the following holds:*

1. *$\tilde{\mathbb{E}}[N] \geq 0$ if and only if N has at least $k!$ terms,*
2. *Let S_i be the set of total orders accepted by exactly i terms in N, then $\tilde{\mathbb{E}}[N] \geq 0$ if and only if $|R| \leq \sum_{i=2}^{\infty}(i-1)|S_i|$.*

We remark that the equivalence with the covering problem presented in the introduction comes from the first point in this lemma. Finding a covering of Ord^{*1} with less than $d!$ sets (that correspond here to normalized terms) is equivalent to finding a normalized weakening N with support size d and negative pseudo-expectation.

Proof. For the first point, using Lemma 16,

$$\tilde{\mathbb{E}}[N] = \tilde{\mathbb{E}}\left[\sum_{[\![T]\!]_\pi\in N}[\![T]\!]_\pi - 1\right] = \sum_{[\![T]\!]_\pi\in N}\tilde{\mathbb{E}}[[\![T]\!]_\pi] - 1 = \frac{\#\text{of terms in } N}{k!} - 1.$$

This is non-negative if and only if there are at least $k!$ terms in N. For the second point, recall that for terms of the form $t = [\![T]\!]_\pi$, we have $\tilde{\mathbb{E}}[t] = |\{x \in \mathrm{Ord} \mid x{\restriction}T = \pi\}|/|\mathrm{Ord}|$. Thus

$$\begin{aligned}
\tilde{\mathbb{E}}[N] &= \frac{1}{|\mathrm{Ord}|}\left(\sum_{[\![T]\!]_\pi \in N} |\{x \in \mathrm{Ord} \mid x{\restriction}T = \pi\}|\right) - 1 \\
&= \frac{1}{|\mathrm{Ord}|}\left(\sum_{x \in \mathrm{Ord}} (|\{[\![T]\!]_\pi \in N \mid x{\restriction}T = \pi\}|\right) - \frac{|\mathrm{Ord}|}{|\mathrm{Ord}|} \\
&= \frac{1}{|\mathrm{Ord}|}\left(\sum_{i=1}^{\infty} i|S_i|\right) - \frac{1}{|\mathrm{Ord}|}\left(|R| + \sum_{i=1}^{\infty}|S_i|\right) \\
&= \frac{1}{|\mathrm{Ord}|}\left(\sum_{i=2}^{\infty}(i-1)|S_i| - |R|\right).
\end{aligned}$$

◀

The proof of Theorem 1 uses the second point of this lemma: we show that if the degree of a normalized weakening is low enough, then a large number of orders must be accepted by several terms of $W\mathcal{J}$.

4.1.1 A Warmup

We begin by considering the case when $\mathcal{J} = 1$. As a warmup, we show a simplified (but not sufficient) argument, where we only focus on orders accepted by at least two terms. Later we present the stronger argument considering orders accepted by at least 3 terms which is sufficient to prove our theorem. We do this for the general case of arbitrary conical juntas in Section 4.1.3.

We use the notation $1z$ to denote a total order on $[n]$ where the element 1 is first and z is a total order of the remaining $n-1$ elements. The notation $i1z$ is similarly defined: it denotes orders where i is first, 1 is second, and z is a total order of the remaining $n-2$ elements.

▶ **Lemma 18.** *For a normalized DNF,*

1. *Total orders starting with 1 can only be accepted by terms $[\![T]\!]_\pi$ with $\pi(1) = 1$.*
2. *Total orders with first element i and second element 1 can only be accepted by terms $[\![T]\!]_\pi$ with either $\pi(1) = 1$ and $i \notin T$, or $\pi(1) = i$ and $\pi(2) = 1$.*

Proof. For the first point, since 1 is in the support T for each term in a normalized DNF, if $\pi(1) \neq 1$ then π is not equal to $1z{\restriction}T$ for any $1z \in \mathrm{Ord}$. For the second, consider the following two cases: If $\pi(1) = 1$ and $i \in T$, then $i1z{\restriction}T \neq \pi$ for any $i1z \in \mathrm{Ord}$. If $\pi(2) = 1$ and $\pi(1) \neq i$, then $i1z{\restriction}T \neq \pi$. Otherwise, if $\pi^{-1}(1) > 2$, then $i1z{\restriction}T \neq \pi$. ◀

The following definitions help to describe our combinatorial argument. To simplify notation, we denote $z\backslash\{i\}$ by $z\backslash i$, and $[n]\backslash\{1\}$ by $[n]\backslash 1$.

▶ **Definition 19.** *Let z be an order on $[n]\backslash 1$. We say that a term $[\![T]\!]_\pi$ is z-*good *for i if $i \in T$, $1 \in T$, $\pi(1) = i, \pi(2) = 1$, and $i1(z\backslash i){\restriction}T = \pi$. That is, the order π on T places i first, 1 second, and is consistent with the total order $i1(z\backslash i)$.*

Notice that if a term $[\![T]\!]_\pi$ is z-good for i then it accepts the order $i1(z\backslash i)$. Moreover, it also accepts the orders $ij1(z\backslash\{i,j\})$ and $ji1(z\backslash\{i,j\})$ for any j that is not in its support.

▶ **Definition 20.** *Let z be an order on $[n]\backslash 1$. We call an unordered pair of distinct elements $(i,j) \in ([n]\backslash 1)^2$ a* hitting pair *for z if the following conditions hold:*

- *N has a term $[\![T_1]\!]_{\pi_1}$ such that $\pi_1(1) = i, \pi_1(2) = 1, j \notin T_1$ and $i1(z\backslash i)\restriction T_1 = \pi_1$.*
- *N has a term $[\![T_2]\!]_{\pi_2}$ such that $\pi_2(1) = j, \pi_2(2) = 1, i \notin T_2$ and $j1(z\backslash j)\restriction T_2 = \pi_2$.*

In other words, a pair (i,j) is a hitting pair for z, if there is some term that is z-good for i but its support does not contain j, and there is some term that is z-good for j but its support does not contain i.

Notice that if (i,j) is a hitting pair for z, then *both* $ij1(z\backslash\{i,j\})$ and $ji1(z\backslash\{i,j\})$ are accepted by *both* $[\![T_1]\!]_{\pi_1}$ and $[\![T_2]\!]_{\pi_2}$. Therefore, we can get a lower bound on the number of orders accepted by several terms of N by getting a lower bound on the number of hitting pairs.

▶ **Lemma 21.** *Let N be a normalized DNF with support size k that is a weakening of M_1. Let $R \subseteq$ Ord be the set of orders rejected by N. If $k \leq n/10$, there are at least $4|R|/5$ total orders accepted by more than one term of N.*

Proof. If $R = \emptyset$, we are done, so suppose otherwise. Recall that we assume for normalized DNFs that each term contains the element 1, so we do not explicitly state this condition.

Given an order z on $[n]\backslash 1$, we define a matrix H_z as follows. Let H_z be an $(n-1)\times(n-1)$ matrix with rows and columns indexed by $i,j \in [n]\setminus 1$, respectively, where $H_z(i,j) = 1$ iff $i \neq j$, and N has a term $[\![T]\!]_\pi$ such that $\pi(1) = i$, $\pi(2) = 1$, $j \notin T$, and $i1(z\backslash i)\restriction T = \pi$. That is, the (i,j)th entry is 1 iff there is a term that is z-good for i but its support does not contain j. Thus, (i,j) is a hitting pair iff both $H_z(i,j) = 1$ and $H_z(j,i) = 1$.

Let $1z \in R$ be an order rejected by N. Note that since N is a weakening of M_1, every order it rejects must start with 1. Moreover, for every $i \in [n]\backslash 1$, N must accept the total order $i1(z\backslash i)$, as it is accepted by M_1.

If N had any terms $[\![T]\!]_\pi$ with $\pi(1) = 1$ and $1z\restriction T = \pi$, $1z$ would not be rejected. Thus, by point 2 of Lemma 18, for each $i \in [n]\backslash 1$, the order $i1(z\backslash i)$ must be accepted by a term $[\![T]\!]_\pi$ of N with $\pi(1) = i, \pi(2) = 1$.

That is, for each $i \in [n]\backslash 1$, and any z such that $1z$ is rejected by N, there is a term of N that is z-good for i. Since N is normalized, the support of each term has size k, and so row i of H_z must have at least $n-k$ entries with value 1 (there are $n-k$ elements j not included in the support of any particular term). The matrix thus contains at least $(n-1)(n-k)$ entries which are 1.

Next, observe that the largest number of 1-entries in a matrix without a hitting pair is at most $\binom{n-1}{2}$: by definition, any such matrix has a 1 in at most one of the entries $H_z(i,j)$ or $H_z(j,i)$, for each unordered pair (i,j). Moreover, the following holds.

▶ **Observation 22.** *Any $(n-1)\times(n-1)$ $\{0,1\}$-matrix with $q > \binom{n-1}{2}$ 1-entries outside of the diagonal must contain at least $q - \binom{n-1}{2}$ hitting pairs.*

Proof. Let H be an arbitrary $(n-1)\times(n-1)$ $\{0,1\}$-matrix with $q > \binom{n-1}{2}$ 1-entries, none of which are on the diagonal. Since the number of 1-entries is more than $\binom{n-1}{2}$, H must contain at least one hitting pair (i,j). Replace one of the entries corresponding to this pair (either the entry $H(i,j)$ or the entry $H(j,i)$) by 0. We can repeat this step at least $q - \binom{n-1}{2}$ times, accounting for at least $q - \binom{n-1}{2}$ hitting pairs. ◀

Thus, in the matrix H_z there are at least $(n-1)(n-k) - \binom{n-1}{2}$ hitting pairs for z. This holds for each z such that $1z \in R$.

Each of the hitting pairs in H_z causes (at least) two orders to be accepted twice, in particular, as discussed above, if (i,j) is a hitting pair for z, then *both* $ij1(z\backslash\{i,j\})$ and $ji1(z\backslash\{i,j\})$ are accepted twice. Notice that the total orders we count twice this way due to (i,j) are the same for z and z' iff $z\backslash\{i,j\} = z'\backslash\{i,j\}$. For any given z, the number of z' such that $z\backslash\{i,j\} = z'\backslash\{i,j\}$ is at most $(n-1)(n-2)$.

Thus, we can conclude that there must be at least

$$\frac{2\left((n-1)(n-k) - \binom{n-1}{2}\right)}{(n-1)(n-2)}|R|$$

orders accepted by at least two terms of N. This value is always smaller than $|R|$. However, if $k \leq n/10$, it is at least $4|R|/5$. ◀

4.1.2 Normalization with Respect to Juntas

The above argument shows that focusing on hitting pairs is almost enough to conclude that $\tilde{\mathbb{E}}[W] \geq 0$ for weakenings W, even if they have a fairly high degree. Our final proof for the general case is similar to this argument, but focuses on hitting triples. Before stating and proving this result, we need to deal with conical juntas. We extend the normalization done in Lemma 15 to handle the additional inclusion of a conical junta. This is captured by the following lemmas.

▶ **Lemma 23.** *Let the DNF W be a weakening of M_1 and $\mathcal{J}$ be a conical junta. If $\tilde{\mathbb{E}}[W\mathcal{J}] < 0$ there exists N and t such that*

1. *N is the normalized form of W given by Lemma 15,*
2. *$t = [\![T]\!]_\pi$ such that $1 \in T$, $\pi(1) = 1$, and the support size of t is $\leq 2\deg(\mathcal{J}) + 1$*
3. *$\deg(Nt) \leq 2\deg(W) + 2\deg(\mathcal{J})$,*
4. *$\tilde{\mathbb{E}}[Nt] < 0$.*

Proof. Let $\mathcal{J} = \sum_{t\in\mathcal{J}} t$ and suppose that $0 > \tilde{\mathbb{E}}[W\mathcal{J}] = \sum_{t\in\mathcal{J}} \tilde{\mathbb{E}}[Wt]$, where the equality follows by linearity. As the sum is negative, there must be at least one term t^* such that $\tilde{\mathbb{E}}[Wt^*] < 0$. Let T be the support of t^*. Note that without loss of generality we can assume that 1 belongs to T: If $1 \notin T$ then let Π be the set of permutations on $T \cup \{1\}$ consistent with t^*. As shown in the proof of Lemma 15, t^* is equivalent to $\bigvee_{\pi\in\Pi}[\![T\cup\{1\}]\!]_\pi$, and $\tilde{\mathbb{E}}[t^*] = \tilde{\mathbb{E}}\left[\sum_{\pi\in\Pi}[\![T\cup\{1\}]\!]_\pi\right]$. In particular, every order accepted by t^* is accepted by $[\![T\cup\{1\}]\!]_\pi$ for exactly one $\pi \in \Pi$. Thus, $\tilde{\mathbb{E}}[Wt^*] = \sum_{\pi\in\Pi}\tilde{\mathbb{E}}[W[\![T\cup\{1\}]\!]_\pi]$, and if $\tilde{\mathbb{E}}[Wt^*] < 0$ then there must exist $\pi \in \Pi$ such that $\tilde{\mathbb{E}}[W[\![T\cup\{1\}]\!]_\pi] < 0$.

Suppose that $\pi(1) \neq 1$, then $W \wedge [\![T]\!]_\pi$ does not accept any order that starts with 1, since $[\![T]\!]_\pi$ does not. As W is a weakening of M_1, it must accept every order that does not start with 1. Thus, $W \wedge [\![T]\!]_\pi$ is equivalent to $[\![T]\!]_\pi$ and $\sum_{w\in W}\tilde{\mathbb{E}}[w[\![T]\!]_\pi] \geq \tilde{\mathbb{E}}[[\![T]\!]_\pi]$, as every order accepted by $[\![T]\!]_\pi$ must be accepted by at least one term w of W. This implies that

$$\tilde{\mathbb{E}}\big[W[\![T]\!]_\pi\big] = \tilde{\mathbb{E}}\left[\Big(\sum_{w\in W} w - 1\Big)[\![T]\!]_\pi\right] = \sum_{w\in W}\tilde{\mathbb{E}}[w[\![T]\!]_\pi] - \tilde{\mathbb{E}}[[\![T]\!]_\pi] \geq \tilde{\mathbb{E}}[[\![T]\!]_\pi] - \tilde{\mathbb{E}}[[\![T]\!]_\pi] = 0.$$

Combining this with the previous argument, we get that if there is some conical junta such that $\tilde{\mathbb{E}}[W\mathcal{J}] < 0$ then there is also a term $t = [\![T]\!]_\pi$, with $1 \in T$ and $\pi(1) = 1$, such that $\tilde{\mathbb{E}}[Wt] < 0$. Note that the support size of $t = [\![T]\!]_\pi$ is at most $2\deg(\mathcal{J}) + 1$, and its degree is at most $2\deg(\mathcal{J})$.

Next, we apply Lemma 15 to the DNF W to obtain a normalized DNF N. Recall that the degree of N will be at most $2\deg(W)$. Since both the degree of N and t might have at most doubled, and $\deg(Nt) \leq \deg(N) + \deg(t)$, we get that $\deg(Nt) \leq 2\deg(W) + 2\deg(\mathcal{J})$, concluding (3).

We now argue that $\tilde{\mathbb{E}}[Wt] < 0$ implies that $\tilde{\mathbb{E}}[Nt] < 0$. Expanding Nt, we have

$$\tilde{\mathbb{E}}[Nt] = \tilde{\mathbb{E}}\Big[\Big(\sum_{v\in N} v - 1\Big)t\Big] = \sum_{v\in N} \tilde{\mathbb{E}}[vt] - \tilde{\mathbb{E}}[t].$$

Since N is the normalization of W, the terms of N can be partitioned into groups, one for each term of W, such that the term of W accepts a given order iff a unique term in its corresponding group accepts that order, and the term of W rejects a given order iff all terms in the corresponding group reject it. Thus, $\sum_{v\in N} \tilde{\mathbb{E}}[vt] = \sum_{w\in W} \tilde{\mathbb{E}}[wt]$, and so $\tilde{\mathbb{E}}[Nt] = \tilde{\mathbb{E}}[Wt] < 0$. ◀

Note that in the above lemma, both N and t are normalized, but the DNF corresponding to $N \wedge t$ (obtained by taking the conjunction of each term of N with t) is not. We address this issue next.

▶ **Lemma 24.** *Let $t = [\![T]\!]_{\pi^*}$ be a term of support size ℓ such that $1 \in T$ and $\pi^*(1) = 1$, and let N be a normalized DNF of support size h that is a weakening of M_1. Then there is a DNF N' such that*

1. *For every total order x, $N'(x) \Leftrightarrow N(x) \wedge t(x)$. Hence, N' is a weakening of $M_1 \wedge t$,*
2. *N' is normalized with support size at most $\ell + h - 1$, and degree at most $\ell + h - 2 \leq \deg(N) + \deg(t)$,*
3. *$\tilde{\mathbb{E}}[Nt] = \tilde{\mathbb{E}}[N' + 1 - t]$.*

Coming back to our analogy with the covering problem presented in the introduction, proving that $\tilde{\mathbb{E}}[N' + 1 - t]$ is non-negative for all such N' of small degree (for a fixed t), is equivalent to saying that, even if we restrict the sets Ord and Ord^{*1} to their orders consistent with t, it is impossible to get a covering of $(\mathrm{Ord} \wedge t)^{*1}$ with fewer sets than the number needed to cover $\mathrm{Ord} \wedge t$, if the support size of the sets is small.

Proof of Lemma 24. We construct N' by "merging" terms of N and t. Formally, let $[\![S]\!]_\sigma$ be a term of N, and let $\Pi_{S,\sigma}$ be the set of permutations on $S \cup T$ consistent with both σ and π^*; note that this set might be empty. Define $N' := \bigvee_{[\![S]\!]_\sigma \in N} \bigvee_{\pi\in\Pi_{S,\sigma}} [\![S \cup T]\!]_\pi$ and observe that this DNF accepts the same total orders Nt does. Indeed, an order is consistent with a term $[\![S]\!]_\sigma$ of N and t iff there is an ordering in $\Pi_{S,\sigma}$ consistent with it. Every term in N' is normalized, but since for different terms $[\![S]\!]_\sigma, [\![S']\!]_\sigma$ of N, the sizes of $S \cup T$ and $S' \cup T$ might be different, terms in N' might be of different sizes. As in the previous arguments, one can simply add elements to the support of smaller terms to get them to a fixed size. The maximal support size of a term in N' is the maximal size of $S \cup T$, where S is the support of some term in N. Since N and t are both normalized, and since they both contain 1, this value is at most $\ell + h - 1$. This also implies that the degree $\deg(N') \leq \ell + h - 2 \leq \deg(N) + \deg(t)$.

It remains to prove that $\tilde{\mathbb{E}}[Nt] = \tilde{\mathbb{E}}[N' + 1 - t]$. For a single term $[\![S]\!]_\sigma$ of N it holds that $\tilde{\mathbb{E}}[[\![S]\!]_\sigma t] = \tilde{\mathbb{E}}\left[\sum_{\pi\in\Pi_{S,\sigma}} [\![S \cup T]\!]_\pi\right]$, as every total order accepted by $[\![S]\!]_\sigma t$ is consistent with a unique $\pi \in \Pi_{S,\sigma}$. Thus,

$$\begin{aligned}
\tilde{\mathbb{E}}[Nt] &= \tilde{\mathbb{E}}\Big[\sum_{[\![S]\!]_\sigma\in N} [\![S]\!]_\sigma t - t\Big] = \sum_{[\![S]\!]_\sigma\in N} \tilde{\mathbb{E}}\big[[\![S]\!]_\sigma t\big] - \tilde{\mathbb{E}}[t] \\
&= \sum_{[\![S]\!]_\sigma\in N} \tilde{\mathbb{E}}\Big[\sum_{\pi\in\Pi_{S,\sigma}} [\![S \cup T]\!]_\pi\Big] - \tilde{\mathbb{E}}[t] = \tilde{\mathbb{E}}\Big[\sum_{[\![S]\!]_\sigma\in N} \sum_{\pi\in\Pi_{S,\sigma}} [\![S \cup T]\!]_\pi\Big] - \tilde{\mathbb{E}}[t] \\
&= \tilde{\mathbb{E}}[N' + 1] - \tilde{\mathbb{E}}[t] = \tilde{\mathbb{E}}[N' + 1 - t].
\end{aligned}$$

◀

4.1.3 The General Case

We are now ready to state and prove our main theorem about weakenings of M_1.

▶ **Theorem 25.** *Let $t = [\![T^*]\!]_{\pi^*}$ be a term of support size ℓ such that $\pi^*(1) = 1$, and let N be a normalized DNF of support size h that is a weakening of M_1. If $\ell + h \leq n/100$, then $\tilde{\mathbb{E}}[Nt] \geq 0$.*

In Section 4.1.1 we presented a warmup argument for the case when the conical junta, $\mathcal{J}$, was assumed to be the trivial junta $\mathcal{J} = 1$. This simpler argument (considering only "pairs") is not sufficient to get a good bound, even for the case when $\mathcal{J} = 1$, but it illustrates our approach well. To get a good enough bound, we extend this argument to triples. To prove Theorem 25, we allow $\mathcal{J}$ to be an arbitrary conical junta. This also includes the case when $\mathcal{J}$ could be the trivial junta $\mathcal{J} = 1$, in which case we define its support to be $T^* = \{1\}$.

Proof of Theorem 25. Let N' be the weakening of $M_1 \wedge t$ given by Lemma 24; that is $\tilde{\mathbb{E}}[Nt] = \tilde{\mathbb{E}}[N' + 1 - t]$. We will prove that $\tilde{\mathbb{E}}[N' + 1 - t] \geq 0$. Let $\ell + g$ be the support size of N', and recall that this is at most $\ell + h - 1 \leq n/100 - 1$. Let R be the set of total orders that are consistent with t and are rejected by N', and hence also by $N \wedge t$. Recall also that by Lemma 24 N' only accepts orders consistent with t. We will prove that the over-counting in the pseudo-expectation that is due to orders that are accepted by several terms of N' is sufficiently large compared to the number of orders rejected by N' that are consistent with t. A sufficiently large "over-count" will be enough to show that $\tilde{\mathbb{E}}[N' + 1 - t] \geq 0$, since $\tilde{\mathbb{E}}[N' + 1 - t]$ is exactly the sum of the number of orders that each term of N' accepts minus the number of orders consistent with t, divided by $n!$. $\tilde{\mathbb{E}}[N' + 1 - t]$ is thus non-negative, if and only if the number of extra occurrences of orders accepted by more than one term of N' is at least $|R|$ (Recall that $|R|$ is the number of orders consistent with t that are not accepted by any term of N').

Given an order z on $[n] \backslash 1$, we call an unordered triple of distinct elements $(i, j, k) \in ([n] \backslash T^*)^3$ a *hitting triple for z* if the following hold:

1. N' has a term $[\![T_1]\!]_{\pi_1}$ such that $\pi_1(1) = i, \pi_1(2) = 1, j, k \notin T_1$ and $i1(z \backslash i) \restriction T_1 = \pi_1$.
2. N' has a term $[\![T_2]\!]_{\pi_2}$ such that $\pi_2(1) = j, \pi_2(2) = 1, i, k \notin T_2$ and $j1(z \backslash j) \restriction T_2 = \pi_2$.
3. N' has a term $[\![T_3]\!]_{\pi_3}$ such that $\pi_3(1) = k, \pi_3(2) = 1, i, j \notin T_3$ and $k1(z \backslash k) \restriction T_3 = \pi_3$.

In other words, (i, j, k) is a hitting triple for z, if there is some term that is z-good for i but j and k are missing from its support, there is some term that is z-good for j but i and k are missing from its support, and there is some term that is z-good for k but i and j are missing from its support. Analogous to hitting pairs, the presence of hitting triples implies that there are orders accepted by several terms. Indeed, if (i, j, k) is a hitting triple for z, then for all six permutations σ of $\{i, j, k\}$, the order $\sigma(1)\sigma(2)\sigma(3)1(z \setminus \{i, j, k\})$ is accepted by at least three terms.

Given an order z on $[n] \backslash 1$, define an array $H_z : ([n] \backslash T^*)^3 \to \{0, 1\}$, such that $H_z(i, j, k) = 1$ iff i, j, k are distinct and not in T^*, and N' has a term that is z-good for i but j and k are missing from its support (that is N' has a term $[\![S]\!]_\pi$ with $\pi(1) = i, \pi(2) = 1$, $j, k \notin S$ and $i1(z \backslash i) \restriction S = \pi$). Note that by this definition, $H_z(i, j, k) = H_z(i, k, j)$. Furthermore, notice that (i, j, k) is a hitting triple for z iff for all six permutations π of $\{i, j, k\}$, $H_z(\pi(1), \pi(2), \pi(3)) = 1$. In other words, 1-entries now come in pairs in the array, and one needs six 1-entries to get a hitting-triple. Thus, there can be at most $4\binom{n-\ell}{3}$ 1-entries in an array H_z without any hitting triples. The $-\ell$ comes from the fact that we do not consider elements of T^* in the triples. Similarly to Observation 22, we get the following.

▶ **Observation 26.** *Let H be an $(n-\ell)\times(n-\ell)\times(n-\ell)$ {0,1}-array with $q > 4\binom{n-\ell}{3}$ 1-entries, where all 1 entries are in positions with distinct coordinates i,j,k, such that $H(i,j,k) = H(i,k,j)$ for all i,j,k. Then H must contain at least $\frac{1}{2}\left(q - 4\binom{n-\ell}{3}\right)$ hitting triples.*

Proof. Let H be an $(n-\ell)\times(n-\ell)\times(n-\ell)$ {0,1}-array with $q > 4\binom{n-\ell}{3}$ 1-entries, where all 1-entries are in positions with distinct coordinates i,j,k, such that $H(i,j,k) = H(i,k,j)$ for all i,j,k. Since the 1-entries come in pairs, we have $q \geq 4\binom{n-\ell}{3} + 2$ entries that are 1, which implies that for at least one triple with distinct i,j,k, all 6 corresponding entries must be 1. Thus H contains at least one hitting triple. Replace one pair of entries corresponding to this triple, say $H(i,j,k)$ and $H(i,k,j)$, by 0. We can repeat this step at least $\frac{1}{2}\left(q - 4\binom{n-\ell}{3}\right)$ times, accounting for at least $\frac{1}{2}\left(q - 4\binom{n-\ell}{3}\right)$ hitting triples. ◀

Let $1z \in R$ be a rejected order that is consistent with t. By a similar argument to the proof of Lemma 21, for every $i \in [n]\backslash T^*$, there is a term of N' that is z-good for i (and thus this term accepts the order $i1(z\backslash i)$). This only holds for i that are not in T^* (the support of t), since N' only accepts total orders that are accepted by t, that is only total orders that are consistent with π^* on T^*, and moving any element in T^* before 1 would contradict π^*.

Since for all $i \notin T^*$, the order $i1(z\backslash i)$ is accepted by a term of support size $\ell + g$, with $\pi(1) = i, \pi(2) = 1$, there must be at least $(n-\ell-g)(n-\ell-g-1)$ 1-entries in H_z at positions starting with i. Hence, there are at least $\frac{1}{2}\left((n-\ell)(n-\ell-g)(n-\ell-g-1) - 4\binom{n-\ell}{3}\right)$ hitting triples for z. Each hitting triple gives (at least) six orders accepted by (at least) three terms: specifically, as noted earlier, for all six permutations σ of $\{i,j,k\}$, the order $\sigma(1)\sigma(2)\sigma(3)1(z\backslash\{i,j,k\})$ is accepted by at least three terms. Thus each hitting triple gives (at least) 12 total extra occurences of over-counted orders. By a similar reasoning to the proof of Lemma 21, each of these specific orders could come from hitting triples of at most $(n-1)(n-2)(n-3)$ different z-s, since for any given order z on $[n]\backslash 1$ consistent with t, the number of orders z' on $[n]\backslash 1$ consistent with t such that $z\backslash\{i,j,k\} = z'\backslash\{i,j,k\}$ is at most $(n-1)(n-2)(n-3)$.

Thus, we get that there are at least

$$\frac{6\left((n-\ell)(n-\ell-g)(n-\ell-g-1) - 4\binom{n-\ell}{3}\right)}{(n-1)(n-2)(n-3)}|R|$$

total extra occurrences of over-counted orders.

Plugging in $\ell + g + 1 \leq n/100$, we get

$$\begin{aligned}
&\frac{6\left((n-\ell)(n-\ell-g)(n-\ell-g-1) - 4\binom{n-\ell}{3}\right)}{(n-1)(n-2)(n-3)}|R| \\
&= \left(\frac{6(n-\ell)(n-\ell-g)(n-\ell-g-1)}{(n-1)(n-2)(n-3)} - \frac{24\binom{n-\ell}{3}}{(n-1)(n-2)(n-3)}\right)|R| \\
&\geq \left(\frac{6\left((99n/100)^3\right)}{n^3} - \frac{4(n-\ell)(n-\ell-1)(n-\ell-2)}{(n-1)(n-2)(n-3)}\right)|R| \\
&\geq \left(\frac{6\left((99n/100)^3\right)}{n^3} - 4\right)|R| = \left(6\frac{99^3}{100^3} - 4\right)|R| \geq |R|
\end{aligned}$$ ◀

We can now conclude the proof of the main theorem of this section.

Proof of Theorem 12. As we argued above, it suffices to consider weakenings of M_1. By Lemma 23, if $\tilde{\mathbb{E}}[W\mathcal{J}] < 0$ for a weakening W of M_1 and conical junta $\mathcal{J}$, both of degree at most d, then there is a normalized weakening N of support size $h \leq 2d+1$ and a normalized term t of support size $\ell \leq 2d+1$ such that $\tilde{\mathbb{E}}[Nt] < 0$. Notice that if $d \leq (\frac{n}{400} - 1)$, then $\ell + h \leq 4d + 2 \leq n/100$. By Theorem 25, this cannot happen. Thus, there is a Σ_2-pseudo-expectation for LOP of degree $(\frac{n}{400} - 1)$. ◀

Note that the $n/100$ bound was picked for convenience, one could get a better bound by counting more than the hitting triples, and by being more careful in the computation.

5 A Criterion for Non-Reducibility of LOP

In this section, we show that our pseudo-expectation for LOP (Lemma 11) implies a general criterion for the non-reducibility of LOP to other problems in $\mathsf{TF}\Sigma_2^{dt}$.

▶ **Corollary 27.** *Suppose that $R \in \mathsf{TF}\Sigma_2^{dt}$ and R has a $\mathsf{polylog}(n)$-degree uSA-refutation. Then* LOP $\not\leq_{dt} R$.

Using this, we show that LOP $\notin \mathsf{LeastNumber}^{dt}$.

Proof of Theorem 2. The axioms of the formula of $\textsc{LeastNumber}_n$ are:

- $\bigvee_{i\in[n]} x_i$, giving the inequality $\sum_{i\in[n]} x_i - 1 \geq 0$; *(Not all 0)*
- $\overline{x_i} \vee \bigvee_{j<i} x_j$, giving the inequality $(1-x_i) + \sum_{j<i} x_j - 1 = -x_i + \sum_{j<i} x_j \geq 0$ for each $i \in [n]$. *(i is the least index taking value 1)*

This admits the following simple uSA-refutation,

$$\sum_{i=1}^{n} x_i - 1 + \sum_{i=0}^{n-1} 2^i \Big(- x_{n-i} + \sum_{j<n-i} x_j \Big) = -1 + \sum_{i=1}^{n} \Big(1 + \sum_{j=0}^{n-i-1} 2^j - 2^{n-i} \Big) x_i = -1.$$

Although, the refutation is of exponential size, it has constant degree. By Corollary 27, LOP $\notin \mathsf{LeastNumber}^{dt}$. ◀

We now prove Corollary 27. To do so, in Lemma 32 we show that we can factorize any reduction between $\mathsf{TF}\Sigma_2^{dt}$ problems into two parts (i) a weakening step, akin to Σ_2-weakening, and (ii) a *counter-example reduction* [16,23] which takes place in TFNP^{dt}. This elucidates how the two types of reductions between $\mathsf{TF}\Sigma_2^{dt}$ problems that have been considered (formulations and counter-example reductions) relate. Using this, our proof proceeds as follows:

1. Consider an R-formulation of LOP where R admits a low-degree uSA-refutation.
2. Factorize the R-formulation into two parts using Lemma 32: a weakening from LOP to a problem P, and a counter-example reduction from P to R.
3. Argue that we can derive an uSA-refutation of P from an uSA-refutation of R while roughly maintaining the degree (Proposition 34).

5.1 Factoring Reductions

To improve readability, we will use a slightly different notation for $\mathsf{TF}\Sigma_2^{dt}$ problems in this section.

▶ **Notation 28.** *A $\mathsf{TF}\Sigma_2^{dt}$ search problem R is a sequence of search problems $R_m \subseteq \{0,1\}^m \times \mathcal{O}_m^R \times \mathcal{W}_m^R$ for each $m \in \mathbb{N}$. $\mathcal{O}_m^R$ is the set of possible R_m-outputs and $\mathcal{W}_m^R$ is the set of R_m-witnesses. We think of $b \in \mathcal{O}_m^R$ as being a solution to $a \in \{0,1\}^m$ for R_m if $R_m(a,b;c)$ holds for all $c \in \mathcal{W}_m^R$.*

A weaker notion of reducibility between $\mathsf{TF}\Sigma_2$ problems has been considered, known as *counter-example reducibility* [23]. Such a reduction from a search problem Q to a search problem R can be interpreted as the following game: Suppose Alice wants to solve Q on input x. She heard that Bob is actually able to solve R using some solver S, and thus, using the reduction, she sends $f(x)$ to Bob. In turns, Bob uses his solver $b := S\big(f(x)\big)$ and Alice transforms b into the output $y := g(x, b)$. Now, if y is indeed an output for x, Alice got what she wanted. But let us say that Alice does not trust the solution and, after some searching, is able to find z such that $Q(x, y; z)$ does not hold (and hence y is not a correct output for x). She confronts Bob about the issue, claiming that Bob's solver is not correct. Bob, who is very proud of his solver, dismisses her by asking for a proof. Then by computing $c := h(x, b, z)$, Alice is capable of finding a counter-example to b being a correct output for $f(x)$, and hence proving that S is incorrect.

▶ **Definition 29.** *Let Q and R be two $\mathsf{TF}\Sigma_2^{dt}$ problems. A* counter-example reduction from Q to R *or a* counter-example R-formulation of Q *is an R-formulation (f, g) along with, for each $b \in \mathcal{O}^R_{s(n)}$ and $z \in \mathcal{W}^Q_n$, a function $h_{b,z} : \{0,1\}^n \to \mathcal{W}^R_{s(n)}$ such that*

$$R_{s(n)}\big(f(x), b; h_{b,z}(x)\big) \implies Q_n\big(x, g_b(x); z\big)$$

where $s(n)$ is the size of the reduction. The depth, and complexity of the formulation are defined as in Definition 5.

By abuse of notation, an R-formulation of a problem Q will also be said to be a counter-example reduction if there exists a function h such that the triplet (f, g, h) is a counter-example reduction. We also define another type of reduction that is the translation of the Σ_2-weakenings.

▶ **Definition 30.** *Let Q and R be two $\mathsf{TF}\Sigma_2^{dt}$ search problems and let (f, g) be an R-formulation of Q. We say that (f, g) is a* weakening *if*

- *f is the identity,*
- *for each b, g_b is a constant function.*

In particular, weakenings are formulations that are always of size n and depth 1.

▶ **Proposition 31.** *Let Q and R be two $\mathsf{TF}\Sigma_2^{dt}$ search problems. Then the following are equivalent:*

- *There exists a weakening from Q to R,*
- *F_{R_n} is a Σ_2-weakening of F_{Q_n} for every n.*

For a decision tree $T(x)$ whose leaves are labeled with 0 and 1, we will abuse notation and also denote by $T(x)$ the DNF obtained from taking the OR over the terms corresponding to its accepting paths. We also write $\overline{T}(x)$ for the function $1 - T(x)$. Its decision tree is the same as T but with inverted labels and its DNF is obtained from taking the OR over the terms corresponding to the rejecting paths of T.

Proof of Proposition 31. Writing $s := s(n)$, say we have a weakening R-formulation (f, g) of Q. Then for each $b \in \mathcal{O}^R_s$ the axiom $\bigvee_c \overline{R}_n(x, b; c)$ is a Σ_2-weakening of $\bigvee_z \overline{Q}_n(x, g_b; z)$ by the correctness of the reduction, where c ranges over $\mathcal{W}^R_s$ and z over $\mathcal{W}^Q_n$. If F_{R_n} is a Σ_2-weakening of F_{Q_n}, then set $g_b = y_b$ such that $\bigvee_c \overline{R}_n(x, b; c)$ is a weakening of $\bigvee_z \overline{Q}_n(x, y_b, z)$. ◀

Considering only weakenings and counter-example reductions seems very restrictive, as the correctness of counter-example reductions is efficiently verifiable (since they actually occur at the TFNP level as highlighted in [23]) and weakenings are somewhat trivial. Surprisingly enough, they seem to capture reductions in their entirety in the sense of the following lemma.

▶ **Lemma 32.** *Let Q and R be two $\mathsf{TF}\Sigma_2^{dt}$ problems and let (f, g) be an R-formulation of Q. Then there exists $P \in \mathsf{TF}\Sigma_2^{dt}$ such that (f, g) decomposes into two reductions $Q \to P \to R$ and:*

- *the reduction $Q \to P$ is a weakening,*
- *the reduction $P \to R$ is a counter-example reduction.*

This decomposition lemma mirrors the characterization given by Theorem 9: One part corresponds to taking a weakening of a formula while the other has a validity that can be efficiently verified (on one hand uSA-refutations and on the other counter-example reductions – See more about this in Section 6). In order to prove this lemma, we first need to introduce the *reduced problem* associated with a reduction, as it will be our choice for P.

For a search problem $R_m(a, b; c)$, we write $R_{m,b,c}(a)$ for the decision tree computing $R_m(a, b; c)$ to put an emphasis on the fact that its only variable is a.

▶ **Definition 33.** *Let Q and R be two $\mathsf{TF}\Sigma_2^{dt}$ search problems. For (f, g) an R-formulation of Q of size $s := s(n)$, the* reduced problem *$R(f, g)$ is the sequence of ternary relations $R(f, g)_n \subseteq \{0,1\}^n \times (\mathcal{O}_n^Q \times \mathcal{O}_s^R) \times \mathcal{W}_s^R$ given by*

$$R(f,g)_n\big(x,(y,b);c\big) \iff g_b(x) = y \text{ and } R_s\big(f(x),b;c\big).$$

The formula associated to the reduced problem $F_{R(f,g)_n}$ has axioms

$$\overline{G}_{b,y}(x) \vee \bigvee_c \overline{R}_{s,b,c} \circ f(x)$$

and $G_{b,y}(x)$ is the indicator function for the event $g_b(x) = y$. We see $G_{b,y}(x)$ as the decision tree (and hence can also be thought of as a DNF) obtained from the decision tree computing g_b and replacing the labels of its leaf with 1 if the leaf is labeled with y and 0 otherwise.

The formula associated to the reduced problem $R(f, g)$ states "the reduction given by (f, g) is not correct", and for the tuple (b, y), the meaning of the corresponding axiom is "either $g_b(x) \neq y$, or there is a counter-example c to b being an R-output for $f(x)$". The formula encodes the following procedure: Say we are given a Q-input x and a purported output y. To verify that y is a Q-output for x, instead of computing $Q(x, y; z)$ for each z, one can do the following: compute $f(x)$ and, given b an R-output for $f(x)$, compute $R(f(x), b; c)$ for each c and verify that $g_b(x) = y$.

Proof of Lemma 32. Let Q, R, and (f, g) be as in the statement of the lemma, set P to be $R(f, g)$ and denote by $s := s(n)$ the size of the reduction. The formula $F_{R(f,g)_n}$ has an axiom $\overline{G}_{b,y}(x) \vee \bigvee_c \overline{R}_{s,b,c} \circ f(x)$ for every pair y, b, where the variable c ranges over $\mathcal{W}_s^R$, and F_{Q_n} has an axiom $\bigvee_z \overline{Q}_{n,y,z}(x)$ for each y where the variable z ranges over $\mathcal{W}_n^Q$. By the correctness of the reduction (if $y = g_b(x)$ and b is an R-output for $f(x)$, then y is a Q-output for x), we get that the formula

$$\bigvee_z \overline{Q}_{n,y,z}(x) \implies \overline{G}_{b,y}(x) \vee \bigvee_c \overline{R}_{s,b,c} \circ f(x)$$

is a tautology and $F_{R(f,g)_n}$ is indeed a Σ_2-weakening of F_{Q_n}.

Now, setting $f'(x) = f(x)$ for all x, $g'_b(x) = \big(g_b(x), b\big)$ for all b, and $h'_{b,c}(x) = c$ for each pair b, c, the triplet (f', g', h') forms a counter-example R-formulation of $R(f, g)$ since for every x, y, b, and c we get

$$R_s\big(f'(x), b; h'_{b,c}(x)\big) = R_{s,b,c} \circ f(x) = R(f,g)_n\big(x,(b,g_b(x));c\big) = R(f,g)_n\big(x,g'_b(x);c\big). \quad \blacktriangleleft$$

5.2 Proving Corollary 27

▶ **Proposition 34.** *Let Q and R be two $\mathsf{TF}\Sigma_2^{dt}$ search problems and let (f, g) be an R-formulation of Q of size $s(n)$ and depth $d(n)$. If F_{R_m} admits an uSA-refutation of degree $d'(m)$, then $F_{R(f,g)_n}$ admits one of degree $d(n)(d'(s(n)) + 2)$.*

In the following proof, for a polynomial $p(a) = p(a_1, \dots, a_m)$ over m variables and a collection of m decision trees $f(x) = (f_i(x))_{i\in[m]}$, the polynomial $p \circ f(x)$ is $p\big(f_1(x), \dots, f_m(x)\big)$ where each f_i is seen as the sum of the monomials representing its accepting paths. Note that when p is a junta, the resulting polynomial can also be represented by a junta. Indeed, for a term $t(a)$, one can compute $t \circ f(x)$ by querying f_i if the variable a_i appears in t. This gives a decision tree that can in turn be transformed into a conical junta. The degree of $p \circ f(x)$ is equal to $\deg(p) \cdot \mathsf{depth}(f)$. We also define $T \circ f(x)$ a similar way when T is a decision tree and its depth is equal to $\mathsf{depth}(T) \cdot \mathsf{depth}(f)$.

Proof. Writing $s := s(n)$, let us first recall the axioms (in their polynomial inequality form) of the two formulas.

- Axioms of F_{R_s}: $\sum_c \overline{R}_{s,b,c}(a) - 1 \geq 0$ for each $b \in \mathcal{O}_s^R$ with c ranging over $\mathcal{W}_s^R$;
- Axioms of $F_{R(f,g)_n}$: $\overline{G}_{b,y}(x) + \sum_c \overline{R}_{s,b,c} \circ f(x) - 1 \geq 0$ for each pair $b \in \mathcal{O}_s^R$ and $y \in \mathcal{O}_n^Q$ with c ranging over $\mathcal{W}_s^R$. Recall that each axiom encodes the fact that either $g_b(x) \neq y$ or c witnesses that b is not an R-output of $f(x)$ and their conjunction contradicts the correctness of the reduction.

As the axioms of $F_{R(f,g)_n}$ closely resemble the ones of F_{R_s}, our goal is to transform the uSA-refutation of the later formula to one of the former. The idea is pretty simple: compose every polynomial appearing in the refutation of R_s with the function f and the function computed by the sum should still be -1. Let $\sum_b J_b(a)(\sum_c \overline{R}_{s,b,c}(a) - 1) + J(a) = -1$ be an uSA-refutation of R_s. We explain how to deduce an uSA-refutation step by step.

1. For each axiom $\overline{G}_{b,y}(x) + \sum_c \overline{R}_{s,b,c} \circ f(x) - 1$, multiply by $G_{b,y}(x)$. Since $T(x) \cdot \overline{T}(x) = 0$ for any decision tree, we get $G_{b,y}(x) \cdot (\sum_c \overline{R}_{s,b,c} \circ f(x) - 1)$ for each tuple (b, y).
2. Using the fact that $\sum_y G_{b,y} = 1$ (as we sum over the indicator variables of all paths of the decision tree g_b, and every assignment must be consistent with at least one of those paths), summing over y and multiplying by $J_b \circ f$ gives us
$$J_b \circ f(x)\Big(\sum_c \overline{R}_{s,b,c} \circ f(x) - 1\Big).$$
3. Finally, summing over all b and summing $J \circ f(x)$, the result becomes
$$\sum_b J_b \circ f(x)\Big(\sum_c \overline{R}_{s,b,c} \circ f(x) - 1\Big) + J \circ f(x)$$

This actually gives a valid uSA-refutation since we (1) Multiply our axioms by conical juntas, (2) Sum all the results, and (3) Add an extra junta to the sum. Looking at this polynomial as a function over the boolean cube, we get the following: let χ be an assignment of x and α be the resulting assignment $f(\chi)$ of a. Then

$$\sum_b J_b \circ f(\chi)\Big(\sum_c \overline{R}_{s,b,c} \circ f(\chi) - 1\Big) + J \circ f(\chi) = \sum_b J_b(\alpha)\Big(\sum_c \overline{R}_{s,b,c}(\alpha) - 1\Big) + J(\alpha) = -1.$$

Since functional equality is the same as polynomial equality in multi-linear arithmetic we get the equality $\sum_b J_b \circ f(x)\Big(\sum_c \overline{R}_{s,b,c} \circ f(x) - 1\Big) + J \circ f(x) = -1$. Hence, setting $J'_{b,y}(x) = J_b \circ f(x) \cdot G_{b,y}(x)$ and $J'(x) = J \circ f(x)$, we get the uSA-refutation

$$\sum_{b,y} J'_{b,y}(x)\Big(\overline{G}_{b,y}(x) + \sum_c \overline{R}_{s,b,c} \circ f(x) - 1\Big) + J'(x) = -1$$

of $F_{R(f,g)_n}$. The degree of the refutation is $d(n)(d'(s(n)) + 2)$. ◀

Using this last result, we are able to prove Corollary 27.

Proof of Corollary 27. Let R be a $\mathsf{TF}\Sigma_2^{dt}$ search problem such that R_m admits a uSA-refutation of poly-logarithmic degree $d'(m)$ and let (f, g) be an R-formulation of LOP of size $s(n)$ and degree $d(n)$. Then by Lemma 32, we have an R-formulation of similar size and depth of $R(f, g)$ and by the same lemma as well as Proposition 34 and the proof of Theorem 1, we have

- $R(f, g)_n$ is a weakening of LOP_n for each n and hence admits a super-poly-logarithmic lower bound on the degree of its uSA-refutations;
- $R(f, g)_n$ admits an uSA-refutation of degree $d(n)(d'(s(n)) + 2)$.

We then conclude that either $s(n)$ is super quasi-polynomial or either $d(n)$ is super poly-logarithmic since the sum $d(n)(d'(s(n)) + 2)$ is super poly-logarithmic and d' is poly-logarithmic and hence the R-formulation is not efficient. ◀

6 Open Problems

Separating LeastNumber from StrongAvoid

Can we show that LeastNumber $\not\leq_{dt}$ StrongAvoid? While LeastNumber has a low-degree uSA proof, the size of this proof is exponential, and therefore this does not imply a reduction to StrongAvoid. The existence of a low-degree uSA proof precludes us from using the techniques in this paper to obtain this separation. Furthermore, to our knowledge, all of the size lower bounds that have been proven for Sherali-Adams proceed by first proving a degree lower bound via pseudo-expectations, and then applying the size-degree tradeoff [22]. The only lower bounds which do not proceed via this strategy apply strictly to uSA [6, 12], they do not apply to the general Sherali-Adams proof system. Can these techniques be used in order to separate LeastNumber from StrongAvoid?

Counter-Example Reductions and Herbrandization

Thapen [23] gave a notion of *herbrandization* for $\mathsf{TF}\Sigma_2^{dt}$ problems, which associates a $\mathsf{TF}\Sigma_2$ problem R with a TFNP problem Checkable(R) satisfying the following property.

▶ **Lemma 35** ([23]). *Let $Q, R \in \mathsf{TF}\Sigma_2^{dt}$. If Q admits low-complexity counter-example R-formulation, then* Checkable(Q) *admits a low-complexity* Checkable(R)*-formulation.*

This previous lemma, coupled with our decomposition (Lemma 32), implies the following.

▶ **Lemma 36.** *If for every weakening P of $Q \in \mathsf{TF}\Sigma_2^{dt}$,* Checkable($P$) $\not\leq_{dt}$ Checkable(R)*, then $Q \not\leq_{dt} R$.*

In light of this, if P is a weakening of Q, then it is natural to ask whether there is anything that we can infer about the relationship between Checkable(Q) and Checkable(P)? Heuristically, to obtain intuition as to whether Q reduces to R it has been useful to look at the relationship between Checkable(Q) and Checkable(R) in TFNP^{dt}. This question asks to what degree this heuristic can be formalized.

Decision-tree reductions between problems in TFNP and proofs in certain polynomial-time verifiable proof systems are tightly connected. Counter-example reductions are a polynomial-time verifiable way to relate search problems in higher levels of the polynomial hierarchy. We ask whether one can also obtain characterizations of counter-example reductions to $\mathsf{TF}\Sigma_2^{dt}$ problems by proof systems. Furthermore, if this is the case, how do these proof systems compare to the ones obtained by first herbrandizing the $\mathsf{TF}\Sigma_2$ problem R to a TFNP problem Checkable(R) and then taking its corresponding proof system via [3].

References

1 Paul Beame, Stephen A. Cook, Jeff Edmonds, Russell Impagliazzo, and Toniann Pitassi. The relative complexity of NP search problems. *J. Comput. Syst. Sci.*, 57(1):3–19, 1998. doi:10.1006/jcss.1998.1575.

2 Josh Buresh-Oppenheim and Tsuyoshi Morioka. Relativized NP search problems and propositional proof systems. In *19th Annual IEEE Conference on Computational Complexity (CCC 2004), 21-24 June 2004, Amherst, MA, USA*, pages 54–67. IEEE Computer Society, 2004. doi:10.1109/CCC.2004.1313795.

3 Sam Buss, Noah Fleming, and Russell Impagliazzo. TFNP characterizations of proof systems and monotone circuits. In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA*, volume 251 of *LIPIcs*, pages 30:1–30:40. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.ITCS.2023.30.

4 Lijie Chen, Shuichi Hirahara, and Hanlin Ren. Symmetric exponential time requires near-maximum circuit size. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 1990–1999. ACM, 2024. doi:10.1145/3618260.3649624.

5 Stefan S. Dantchev, Barnaby Martin, and Mark Nicholas Charles Rhodes. Tight rank lower bounds for the sherali-adams proof system. *Theor. Comput. Sci.*, 410(21-23):2054–2063, 2009. doi:10.1016/J.TCS.2009.01.002.

6 Susanna F. de Rezende, Aaron Potechin, and Kilian Risse. Clique is hard on average for unary sherali-adams. In *64th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2023, Santa Cruz, CA, USA, November 6-9, 2023*, pages 12–25. IEEE, 2023. doi:10.1109/FOCS57990.2023.00008.

7 Noah Fleming, Stefan Grosser, Siddhartha Jain, Jiawei Li, Hanlin Ren, Morgan Shirley, and Weiqiang Yuan. Total search problems in ZPP. In Shubhangi Saraf, editor, *17th Innovations in Theoretical Computer Science Conference, ITCS 2026, Bocconi University, Milan, Italy, January 27-30, 2026*, volume 362 of *LIPIcs*, pages 60:1–60:26. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026. doi:10.4230/LIPIcs.ITCS.2026.60.

8 Noah Fleming, Stefan Grosser, Toniann Pitassi, and Robert Robere. Black-box PPP is not turing-closed. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 1405–1414. ACM, 2024. doi:10.1145/3618260.3649769.

9 Noah Fleming, Deniz Imrek, and Christophe Marciot. Provably Total Functions in the Polynomial Hierarchy. In Srikanth Srinivasan, editor, *40th Computational Complexity Conference (CCC 2025)*, volume 339 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 28:1–28:40, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2025.28.

10 Noah Fleming, Pravesh Kothari, and Toniann Pitassi. Semialgebraic proofs and efficient algorithm design. *Found. Trends Theor. Comput. Sci.*, 14(1-2):1–221, 2019. doi:10.1561/0400000086.

11 S. Ghenntiyala, L. Zeyong, and N. Stephens-Davidowitz. Range avoidance, arthur-merlin, and tfnp. *In Submission*, 20225.

12 Mika Göös, Alexandros Hollender, Siddhartha Jain, Gilbert Maystre, William Pires, Robert Robere, and Ran Tao. Separations in proof complexity and TFNP. *CoRR*, abs/2205.02168, 2022. doi:10.48550/arXiv.2205.02168.

13 Mika Göös, Pritish Kamath, Robert Robere, and Dmitry Sokolov. Adventures in monotone complexity and TFNP. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124 of *LIPIcs*, pages 38:1–38:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.ITCS.2019.38.

14 Venkatesan Guruswami, Xin Lyu, and Xiuhan Wang. Range avoidance for low-depth circuits and connections to pseudorandomness. *ACM Trans. Comput. Theory*, 17(2):14:1–14:23, 2025. doi:10.1145/3718745.

15 Robert Kleinberg, Oliver Korten, Daniel Mitropolsky, and Christos H. Papadimitriou. Total functions in the polynomial hierarchy. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference, ITCS 2021, January 6-8, 2021, Virtual Conference*, volume 185 of *LIPIcs*, pages 44:1–44:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.ITCS.2021.44.

16 Leszek Aleksander Kolodziejczyk and Neil Thapen. Approximate counting and NP search problems. *J. Math. Log.*, 22(3):2250012:1–2250012:31, 2022. doi:10.1142/S021906132250012X.

17 Oliver Korten. The hardest explicit construction. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 433–444. IEEE, 2021. doi:10.1109/FOCS52979.2021.00051.

18 Oliver Korten and Toniann Pitassi. Strong vs. weak range avoidance and the linear ordering principle. In *65th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2024, Chicago, IL, USA, October 27-30, 2024*, pages 1388–1407. IEEE, 2024. doi:10.1109/FOCS61266.2024.00089.

19 Oliver Korten, Toniann Pitassi, and Russell Impagliazzo. Stronger cell probe lower bounds via local prgs. *Electron. Colloquium Comput. Complex.*, TR25-030, 2025. URL: https://eccc.weizmann.ac.il/report/2025/030.

20 Zeyong Li. Symmetric exponential time requires near-maximum circuit size: Simplified, truly uniform. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 2000–2007. ACM, 2024. doi:10.1145/3618260.3649615.

21 Tsuyoshi Morioka. Classification of search problems and their definability in bounded arithmetic. *Electron. Colloquium Comput. Complex.*, TR01-082, 2001. URL: https://eccc.weizmann.ac.il/eccc-reports/2001/TR01-082/index.html.

22 Toniann Pitassi and Nathan Segerlind. Exponential lower bounds and integrality gaps for tree-like lovász-schrijver procedures. *SIAM J. Comput.*, 41(1):128–159, 2012. doi:10.1137/100816833.

23 Neil Thapen. How to fit large complexity classes into TFNP. *arXiv:2412.09984 [cs.CC]*, 2024. doi:10.48550/arXiv.2412.09984.