Chapter 1
Reflections on Proof Complexity and Counting Principles

Noah Fleming and Toniann Pitassi

Abstract This paper surveys the development of propositional proof complexity and the seminal contributions of Alasdair Urquhart. We focus on the central role of counting principles, and in particular Tseitin's graph tautologies, to most of the key advances in lower bounds in proof complexity. We reflect on a couple of key ideas that Urquhart pioneered: (i) graph expansion as a tool for distinguishing between easy and hard principles, and (ii) "reductive" lower bound arguments, proving via a simulation theorem that an optimal proof cannot bypass the obvious (inefficient) one.

Key words: Theory of Computation, Complexity Theory, Propositional Proof Complexity, Counting Principles, Tseitin Tautologies

1.1 Introduction

One of the most basic questions of logic is the following: Given a universally true statement (tautology) what is the length of the shortest proof of the statement in some standard axiomatic proof system? The propositional logic version of this question is particularly important in computer science for both theorem proving and complexity theory. An important related algorithmic questions is whether there is an efficient algorithm that will produce a proof of any tautology? Such questions of theorem proving and complexity inspired Cook's seminal paper on NP-completeness notably entitled "The complexity of theorem-proving procedures" [17] and were

Noah Fleming
University of Toronto, e-mail: noahfleming@cs.toronto.edu
Toniann Pitassi
University of Toronto, e-mail: toni@cs.toronto.edu

contemplated even earlier by Gödel in his now well-known letter to von Neumann (see [61])[1].

These questions have fundamental implications for complexity theory. As formalized by Cook and Reckhow [17], there exists a propositional proof system giving rise to short (polynomial-size) proofs of all tautologies if and only if NP equals co-NP. Cook and Reckhow were the first to propose a program of research aimed at attacking the NP versus co-NP problem by systematically studying and proving strong lower bounds for standard proof systems of increasing complexity.

The second motivation concerns automated theorem proving. The main goal is to investigate the efficiency of heuristics for testing satisfiability, and to give some theoretical justification for them. The third and perhaps the most compelling reason to study the complexity of propositional proof systems is as a principled way to understand the limitations of current algorithmic approaches for solving NP-hard problems. Almost all algorithms that implement a deterministic or randomized procedure for solving an NP-hard optimization problem are based on a standard propositional proof system, and thus upper and lower bounds on these systems shed light on the inherent complexity of any theorem-proving system based upon it. The most striking example is Resolution on which almost all propositional theorem provers (and even first-order theorem provers) are based.

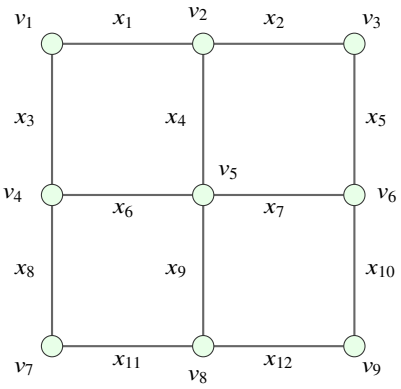
Cook and Reckhow's program has led to many beautiful results in the last twenty years, including strong connections to circuit lower bounds, and celebrated exponential lower bounds on proof size for a variety of important and well-studied proof systems (see the following surveys [6, 65, 57, 60]).

Most of these breakthroughs were established for counting principles such as the propositional pigeonhole principle, and a special family of mod p counting principles, introduced by Tseitin and called *Tseitin's graph tautologies*. Indeed, the pigeonhole principle and the Tseitin tautologies are the most well studied structured hard instances in propositional proof complexity. A Tseitin instance $\mathrm{TS}(G, l)$ is defined relative to an undirected graph $G = (V, E)$, and a labelling $l : V \to \{0, 1\}$ of the vertices of G. The variables correspond to the edges of G, and for each $v \in V$, we have a constraint $\bigoplus_{e:v\in e} x_e = l(v)$ asserting that the parity of the edge variables incident with vertex v must agree with the label $l(v)$. By the handshake principle, for any connected graph G, $\mathrm{TS}(G, l)$ is unsatisfiable if and only if the sum of all labels is odd (see Figure 1.1 for an example).

Tseitin was the first to study the optimal size of propositional proofs, and in particular the optimal size of Resolution proofs. In his landmark 1968 paper [63], Tseitin introduced his now famous Tseitin formulas, and proved lower bounds on the length of *regular* Resolution refutations of the Tseitin formulas on the grid graph.

Subsequently, these formulas have been central to nearly every result in propositional proof complexity. Superpolynomial lower bounds for the Tseitin formulas have been established for many well-studied proof systems, beginning with the seminal paper by Urquhart [64] who proved that Tseitin formulas require exponential-size Resolution refutations, building on Haken's [37] sub-exponential lower bound on the

[1] However, this letter was not discovered until after Cook's paper.



$v_1 : x_1 \oplus x_3 = 1$
$v_2 : x_1 \oplus x_2 \oplus e_4 = 1$
$v_3 : x_2 \oplus x_5 = 1$
$v_4 : x_3 \oplus x_6 \oplus x_8 = 1$
$v_5 : x_4 \oplus x_6 \oplus x_7 \oplus e_9 = 1$
$v_6 : x_5 \oplus x_7 \oplus x_{10} = 1$
$v_7 : x_8 \oplus x_{11} = 1$
$v_8 : x_{11} \oplus x_9 \oplus x_{12} = 1$
$v_9 : x_{12} \oplus x_{10} = 1$

Fig. 1.1 An unsatisfiable instance of the Tseitin formulas on a 3×3 grid graph with $l(v) = 1$ for all $v \in V$.

propositional pigeonhole principle. In the last thirty years, exponential lower bounds for Tseitin formulas were established for stronger proof systems, including including Nullstellensatz [35], the Polynomial Calculus [12], Sum-of-Squares [36, 58], and bounded-depth Frege [9, 49, 38].

In this paper, we survey the landscape of results on the proof complexity of Tseitin formulas and the important role they have played in our understanding of the proof complexity of stronger systems, as well as the complexity of large families of algorithms based on these proof systems. In Section 3, we present Urquhart's seminal result, proving truly exponential lower bound on the length of Resolution refutations for Tseitin formulas on any constant-degree expander graph. We highlight two central concepts that have remained quite important in nearly all subsequent lower bounds. The first is the role of graph expansion as the key combinatorial property underlying the lower bound. Over a highly expanding graph (which behaves like a random graph with respect to expansion), when viewing the graph *locally*, by looking at a small subset of the graph, there is a partial assignment to the edges of this subset which satisfies the Tseitin constraints on the vertices, whereas *globally* there is no satisfying assignment. Thus graph expansion is a crucial *pseudorandom* property used to show that weak proof systems that reason locally (such as Resolution) cannot reason about properties where there is a big distinction between the local versus global behaviour of the property. Secondly, we highlight the *reductive* nature of the lower bound: the proof not only rules out Resolution refutations of sub-exponential length, but it actually shows that *any* Resolution refutation must essentially mimic the obvious upper bound strategy that corresponds to Gaussian elimination.

In Section 4 we survey some of the subsequent important lower bounds in proof complexity. In these breakthrough results, we will see that the proofs, following Urquhart and Haken, show reductive lower bounds using graph expansion as the underlying pseudorandom property. Finally in Section 5 we present a new and very surprising result due to Dadush and Tiwari [19], who showed that Tseitin formulas are easy for Cutting Planes proofs, thereby refuting a widely believed conjecture.

We conclude in Section 6 with some open problems and potential barriers to future progress.

1.2 Preliminaries

1.2.1 Resolution

Using the standard reduction from SAT to 3-SAT, one can take an arbitrary propositional formula F and convert it to a CNF or 3-CNF formula in such a way that it has only polynomially larger size and is unsatisfiable iff the original formula was a tautology. To do this one adds new variables x_A to stand for each of its subformulas A and clauses to specify that the value at each connective is computed correctly, as well as one clause of the form $\neg x_F$. In this way, one can consider any sound and complete system that produces refutations for CNF formulas as a general propositional proof system.

A *literal* ℓ is a propositional variable x or its negation $\neg x$. A *clause* is a disjunction of literals. The Resolution refutation system has a single inference rule:

$$\frac{A \vee \ell,\ B \vee \neg\ell}{A \vee B}.$$

The Resolution rule says that if A and B are clauses and ℓ is a literal, then any assignment that satisfies both of the clauses $A \vee \ell$ and $B \vee \neg\ell$ also satisfies $A \vee B$. The clause $A \vee B$ is said to be a *resolvent* of the clauses $A \vee \ell$ and $B \vee \neg\ell$ derived by *resolving on* the variable ℓ. A *Resolution derivation* of a clause C from a CNF formula F consists of a sequence of clauses in which each clause is either a clause of F, or a resolvent of two previous clauses, and C is the last clause in the sequence; it is a *refutation* of F if C is the empty clause Λ. The *size* of a refutation is the number of resolvents in it. We can represent it as a directed acyclic graph (dag) where the nodes are the clauses in the refutation, each clause of F has out-degree 0, and any other clause has two incoming arcs from the two clauses that produced it. The arcs pointing from $C \vee \ell$ and $D \vee \neg\ell$ to $C \vee D$ are labeled with the literals ℓ and $\neg\ell$ respectively. It is well known that Resolution is a *sound* and *complete* propositional proof system, i.e., a formula F is unsatisfiable if and only if there is a Resolution refutation for F.

History of the Complexity of Resolution Refutations.

Resolution was pre-dated by two systems known as Davis-Putnam procedures which are still the most widely used in propositional theorem proving. The general idea of these procedures is to convert a problem on n variables to problems on $n-1$ variables by eliminating all references to some variable. The former [21] which we call DP

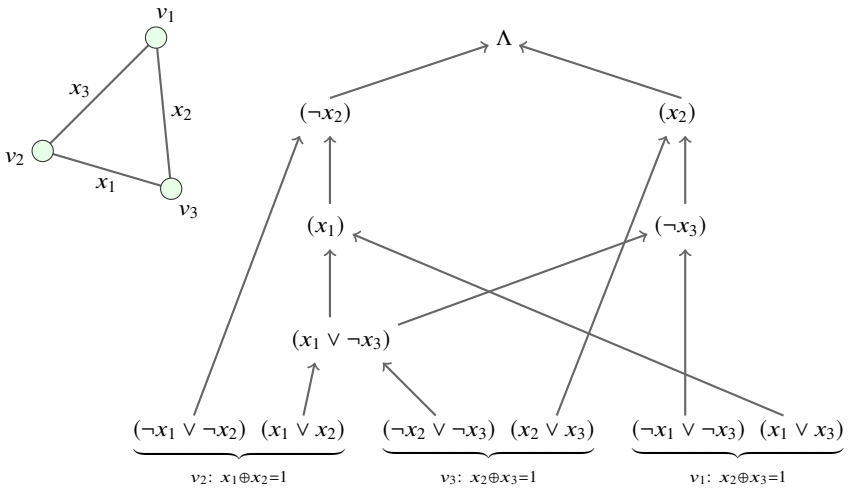


Fig. 1.2 A Resolution refutation of the Tseitin formula on a triangle with $l(v) = 1$ for all $v \in [3]$.

does this by applying all possible uses of the Resolution rule on a given variable to eliminate it. The latter [20], which we call DLL and is the form used today, branches based on the possible truth assignments to a given variable; although at first this does not look like Resolution, it is an easy argument to show that this second form is equivalent to the special class of tree-like Resolution proofs. As a proof system, Resolution is strictly stronger than DP [30] and DLL [65]. The reasons for DLL's popularity are related to its proof search properties which we discuss below.

A more general but still restricted form of Resolution is called *regular* Resolution, and was first introduced by Tseitin in a ground-breaking article [63], the published version of a talk given in 1966 at a Leningrad seminar. A regular Resolution refutation is a Resolution refutation whose underlying directed acyclic graph has the property that along each path from the root (empty clause) to a leaf (initial clause), each variable is resolved upon at most once. Observe that DP refutations are automatically regular. If refutations are represented as trees, rather than directed acyclic graphs, then minimal-size refutations are regular, as can be proved by a simple pruning argument [65, p. 436].

Tseitin [63] established the first super-polynomial lower bounds on the size of regular Resolution refutations of the Tseitin formulas. Interestingly, obtaining an improvement of this bound to an exponential one by Galil [28] was a driving force behind some of the early work in the development of the theory of expander graphs [25, 40].

There was a 15+ year gap before the first super-polynomial lower bound for proofs in general Resolution was obtained by Haken [37] who showed exponential lower bounds for the pigeonhole principle. Subsequently, Urquhart proved the first truly exponential exponential bounds for Resolution refutations of the Tseitin formulas [64].

Urquhart's original proof used the technique known as *bottleneck counting* due to Haken. In this method, one views the proof as a directed acyclic graph of clauses and views the truth assignments as flowing from the root of the directed acyclic graph to a leaf, where an assignment flows through a clause C if and only if: (i) it flows through the parent clause of C and (ii) the assignment falsifies C. Each assignment can be seen to flow through a unique path in any Resolution refutation. The idea is to show that for the formula in question, there must exist a large set of truth assignments with the property that each must pass through a wide clause (a clause containing many literals). Since a wide clause cannot be falsified by too many assignments, this implies that there must exist many wide clauses and hence the proof must be large.

An essential lemma in any bottleneck counting argument is to show that any Resolution refutation of F must involve a wide clause (a bottleneck). An important paper by Ben-Sasson and Wigderson [10], using ideas from [16], shows that this lemma is *sufficient*; namely, they prove that any Resolution refutation of small size can be converted into a refutation with no wide clauses. This result is important since it reduces the more difficult problem of proving Resolution *size* lower bounds to the easier problem of proving Resolution *width* lower bounds.

Lower bounds for the Tseitin formulas have paved the way to proving lower bounds for *random* unsatisfiable instances. Indeed, there is a strong link between Tseitin formulas and random CNF formulas. By varying the underlying odd-labelling, and d-regular graph, we get precisely a uniform distribution on d-XOR instances, where each variable occurs in exactly two equations. Because of this connection, lower bounds for Tseitin formulas has been a precursor to understanding lower bounds for random instances, such as random k-XOR and random k-SAT. Urquhart's lower bound for Tseitin was the precursor to Chvatal and Szemeredi's exponential lower bounds for Resolution refutations of random kCNF formulas [15], and similarly for other proof systems including Nullstellensatz, Polynomial Calculus and SOS [35, 36, 49, 38].

1.3 Urquhart's Resolution Lower Bound

In this section we present the main ideas behind Urquhart's exponential lower bound for Resolution refutations of the Tseitin formulas:

Theorem 1 *Let G be a d-regular odd-charged graph on n vertices with expansion $e(G) = \Omega(n)$. Then any resolution refutation of $TS(G)$ has size $2^{\Omega(n)}$.*

As previously mentioned, Urquhart's original proof used the bottleneck counting method due to Haken. Here we give a simpler presentation of his argument, using Ben-Sasson and Wigderson's size-width theorem for Resolution [10].

We start with some intuition behind the proof. Without loss of generality, we will consider a d-regular graph $G = (V, E)$ on n vertices, where n is odd and such that all charges are odd. The variables of $\mathrm{TS}(G)$ are $x_{i,j}$ where $(i, j) \in E$. Each vertex $i \in V$ corresponds to a constraint which says that the mod-2 sum of the edges incident to

v is odd:

$$\sum_{j,(i,j)\in E} x_{i,j} = 1 \ (mod\ 2).$$

The central intuition behind the proof is to relate the combinatorial notion of *expansion* of the underlying graph G to the complexity of refuting TS(G):

Definition 1 (Graph Expansion.) Let G be an undirected graph with n vertices. The expansion of G, $e(G)$, is $\min\{|E(V', V - V')| \ : \ V' \subseteq V, n/3 \leq |V'| \leq 2n/3\}$.

For any odd-charged graph G, the equations in TS(G) form a set of mod-2 constraints with the property that every variable occurs in exactly two constraints. Thus if G is a random d-regular graph with a random odd-charged labelling, then TS(G) can be viewed as a random XOR formula, where each mod-2 constraint contains $d - 1$ variables, and such that each variable occurs in exactly 2 equations. The most obvious way to obtain a contradiction is iteratively deriving new mod-2 constraints. For instance if $x_{1,2} + x_{1,3} + x_{1,4} = 1$ and $x_{1,3} + x_{2,5} = 1$ have been derived, then we can derive their sum mod-2: $x_{1,2} + x_{1,4} + x_{2,5} = 0$. This simple addition rule is sometimes called the Gaussian rule. A Gaussian refutation consists of a sequence of mod-2 equations where each equation is either an initial one or obtained by two previous equations by the Gaussian rule, and such that the final equation is $0 = 1$. For a Gaussian refutation Π of TS(G), let $width(\Pi)$ be the maximum number of variables that occur in any equation in Π. Now it follows fairly easily from the definition of expansion that if G is a connected expanding graph (i.e. $e(G) = \Omega(n)$), then any Gaussian refutation of G must have width $\Omega(n)$.

Let's see what happens when we try to mimic a Gaussian using Resolution. Since Resolution can only express *disjunctions* of literals, a mod-2 constraint involving k variables translates into an equivalent conjunction of 2^{k-1} clauses. Now if G is expanding, any Gaussian refutation has linear width, and therefore translating this refutation to a Resolution refutation will lead to a huge blowup in size – the proof will have size exponential in n.

The difficult step in proving Resolution lower bounds for TS(G) is therefore to prove that Resolution can do no better – that is, that the optimal size Resolution refutation for TS(G) is that obtained by mimicking a Gaussian refutation.

We now proceed to the proof. An *assignment* for a formula F (sometimes we call it also a *restriction*) is a Boolean assignment to some of the variables in the formula; the assignment is *total* if all the variables in the formula are assigned values. If C is a clause, and σ an assignment, then we write $C\lceil\sigma$ for the result of applying the assignment to C, that is, $C\lceil\sigma = 1$ if $\sigma(l) = 1$ for some literal l in C, otherwise, $C\lceil\sigma$ is the result of removing all literals set to 0 by σ from C (with the convention that the empty clause is identified with the Boolean value 0). If F is a CNF formula, then $F\lceil\sigma$ is the conjunction of all the clauses $C\lceil\sigma$, C a clause in F. If $R = C_1, \ldots, C_k$ is a Resolution derivation from a formula F, and σ an assignment to the variables in F, then we write $R\lceil\sigma$ for the sequence $C_1\lceil\sigma, \ldots, C_k\lceil\sigma$.

Ben-Sasson and Wigderson [10] proved the following relationship between Resolution width and Resolution proof size.

Lemma 1 (Size-Width Lemma) *Let F be an unsatisfiable k-CNF formula over n variables, with a Resolution refutation of size s. Then F also has a refutation of width $O(\sqrt{n \log s}) + k$.*

Thus, sufficiently strong lower bounds on *width* imply superpolynomial or even exponential lower bounds on proof *size*. For tree-like Resolution proofs, they obtained a similar result, proving that tree-like refutations of size S imply refutations of width $O(\log S)$.

We first reproduce Ben-Sasson and Widerson's proof of the Size-Width Lemma, which uses the following Lemma.

Lemma 2 (Lemma 3.2 in [10])

Let F be an unsatisfiable k-CNF formula, and let ℓ be a literal appearing in F. If $F\lceil(\ell = 1)$ has a refutation of width $w - 1$ and $F\lceil(\ell = 0)$ have a refutation of width at most w, then F has a width-w refutation.

Proof (Proof sketch) Let $\pi\lceil(\ell = 1)$ be a width $w - 1$ refutation of $F\lceil(\ell = 1)$. By adding ℓ and $\neg\ell$ back to all initial clauses of F, $\pi\lceil(\ell = 1)$ becomes a derivation of the clause $\neg\ell$ of width w. We can then resolve the derived clause $\neg\ell$ with all clauses in F to derive $F\lceil(\ell = 0)$ in width w. Then by assumption that $F\lceil(\ell = 0)$ has a width w refutation, we can refute F in width w. □

Proof (Proof of Size-Width Lemma) Let F be an unsatisfiable k-CNF over n variables, and let π be a size s resolution refutation of F. Let π^* denote the set of wide clauses in π, where a wide clause is one that contains at least $w := \sqrt{2n \log s}$ literals. We prove by induction on b and n that if $|\pi^*| < a^b$ then F has a refutation of width at most $w + k + b$, where $a = (1 - w/(2n))^{-1}$. The base case ($b = 0$) is trivially true. For the induction step, by an averaging argument, there must exist a literal, say ℓ, appearing in at least the average number of wide clauses, which is $|\pi^*| \cdot w/(2n)$. Restricting the entire proof π by setting $\ell = 1$ gives a refutation of $F\lceil(\ell = 1)$ with at most $(1 - w/2n)|\pi^*| < a^{b-1}$ wide clauses, which by induction on b has a refutation of width at most $w + k + b - 1$. On the other hand, setting $\ell = 0$ gives a refutation of $F\lceil(\ell = 0)$ of width at most $w + k + b$, by induction on n. Applying Lemma 2 completes the proof. □

The next lemma shows that as long as G is a connected graph with good expansion, then any resolution refutation of $\mathrm{TS}(G)$ must have linear width. This combined with the Size-Width Lemma completes the proof of 1.

Lemma 3 (Tseitin Width Lower Bound) *Let G be a connected d-regular odd-charged graph with n vertices, and linear expansion, i.e. $e(G) = \Omega(n)$. Then any Resolution refutation of $TS(G)$ requires width $\Omega(n)$.*

Proof Let π be a Resolution refutation of $\mathrm{TS}(G)$. Let A be the set of clauses of $\mathrm{TS}(G)$, let $A(v)$ denote the clauses associated with vertex v, and for $V^* \subseteq V$ let $A(V^*) := \cup_{v \in V^*} A(v)$. We define the following complexity measure $\mu(C)$ on clauses C over the variables of $\mathrm{TS}(G)$.

$$\mu(C) = min\{|V'| \mid V' \subseteq V,\ A(V') \implies C^*\}.$$

That is, $\mu(C)$ is the size of the minimal set of vertices $V' \subseteq V$ such that the clauses associated with V' imply C. Since Resolution is a sound procedure, μ is subadditive. That is, if C is derived from the Resolution rule applied to clauses C_1 and C_2, then $\mu(C) \leq \mu(C_1) + \mu(C_2)$. Note that for an initial clause, $C \in A$, $\mu(C) = 1$, and for the final empty clause of Π, $\mu(\Lambda) = |A|$, because if one of the clauses of A is left out then A becomes satisfiable. Therefore by subadditivity, there exists a clause C^* in π such that $|A|/3 \leq \mu(C^*) \leq 2|A|/3$.

Let V' denote a minimal set of vertices such that $A(V') \implies C$. By expansion of G, the size of the boundary $|E(V', V \setminus V')| = \Omega(n)$. We will argue that $width(C) = \Omega(n)$ since all literals associated with the edges in $E(V', V \setminus V')$ must occur in C. Let $x_i \in E(V', V \setminus V')$ and suppose that x_i does not occur in C. We will construct an assignment that falsifies C but satisfies $A(V')$, contradicting that $A(V') \implies C$. Because $x_i \in E(V', V \setminus V')$, there is exactly one vertex $v \in V'$ incident to x_i. Pick an assignment α that falsifies $A(v)$, falsifies C, and satisfies $A(v')$ for all $v' \in V' \setminus \{v\}$; the existence of such an assignment follows because $A(V') \implies C$ and V' is minimal. Let $\alpha^{\oplus i}$ be the assignment obtained from α by flipping the ith bit, i.e. $\alpha_j^{\oplus i} = \alpha_j$ for $j \neq i$ and $\alpha_i^{\oplus i} = 1 - \alpha_i$. Then $\alpha^{\oplus i}$ falsifies $A(v')$ and therefore $A(V')$, however it satisfies C because C does not depend on x_i; this contradicts that $A(V') \implies C$. Therefore C must depend on all of the variables in $|E(V', V \setminus V')|$. Altogether, we have shown that any Resolution refutation of TS(G) has width $\Omega(n)$. □

1.4 Subsequent Lower Bounds for Tseitin Formulas

1.4.1 Bounded-Depth Frege

A Frege system is a propositional proof system where the underlying lines in a proof are Boolean formulas over the basis $\wedge$, $\vee$ and $\neg$. There are a large number of axiomatizations of Frege systems, and by the foundational results of Cook and Reckhow [17], they are all known to be polynomially equivalent, meaning that the minimum proof length of any tautology remains the same to within a polynomial factor. Obtaining even super-linear lower bounds for Frege proofs for any family of tautologies remains one of the most important open problem in proof complexity.

A well-known restricted proof system is *bounded-depth* Frege, where the rules remain the same, but we impose the restriction that every formula in the proof has depth at most d. (In order for this to remain a complete system, we measure depth of a formula as the number of alternations of $\vee$ and $\wedge$ connectives in the formula, or equivalently, we can generalize the connectives $\wedge$ and $\vee$ and associated rules to have unbounded-fanin, and then the depth is simply the number of alternations of unbounded-fanin $\wedge$ and $\vee$ gates.

In a breakthrough result, Ajtai [1] established super-polynomial lower bounds on the length of bounded-depth Frege proofs of the propositional pigeon hole principle. In this tour-de-force paper, he actually proved the existence of a nonstandard model for an axiomatic system of arithmetic ($I\Delta_0$) that corresponds to bounded-depth Frege, where the pigeonhole principle is false. His ingenious construction of a nonstandard model is obtained by showing, through a combinatorial switching lemma, that any small bounded-depth Frege proof of the pigeonhole principle actually implies the existence of a small, simpler Resolution proof. Thus the proof again establishes a reductive lower bound by reducing a more complex bounded-depth Frege proof to a simpler Resolution proof. Subsequently, [8] obtained somewhat stronger bounds, via a purely combinatorial analysis of Ajtai's proof. Exponential lower bounds for bounded-depth Frege proofs of the pigeonhole principle were established in [48, 47].

Urquhart and Fu [67] gave a more streamlined presentation of the lower bounds mentioned above, and extended the argument to prove super-polynomial lower bounds for Tseitin's formulas on the complete graph. Shortly thereafter and independently, Ben-Sasson [9] managed to prove lower bounds for the Tseitin formulas over any expander graph by a clever reduction to the pigeonhole lower bound.

The above lower bounds left open the question of proving *optimal* lower bounds. Due to the difficulty of proving a switching lemma in this context, the state-of-the-art lower bounds for depth-d Frege proofs of the pigeonhole principle (as well as for Tseitin) are exponential in $n^{1/2^d}$, whereas the best possible upper bounds are exponential in $n^{1/d}$, leaving a large gap. Improving the lower bound to matching the upper bound appeared to be much more difficult, and closing this gap remained open for over twenty years. The first progress was made by Pitassi, Rossman, Servedio and Tan [49] who were able to close this gap but only for a certain range of parameters, by establishing a switching lemma following some of the high-level ideas in a related breakthrough result [39]. In a remarkable recent paper, Håstad [38] finally obtained near-optimal lower bounds for Tseitin formulas, for the grid graph (the original graph used by Tseitin to prove super-polynomial lower bounds on regular Resolution proofs). Recently, Galesi et al. [26] extended Håstad result to every graph, obtaining a lower bound of $2^{tw(G)^{\Omega(1/d)}}$ where $tw(G)$ is the *tree-width* of G. Furthermore, they show that this is tight up to a multiplicative constant in the top exponent.

1.4.2 Algebraic Proof Systems

Algebraic proof systems, first defined in [5], are aimed at proving the unsolvability of a family of polynomial equalities or inequalities over an underlying field, and as a special case they are refutation systems for unsatisfiable CNF formulas. Given an unsatisfiable k-CNF formula over n variables, by a standard translation we can convert the formula into a family of degree-k polynomial equations $\mathcal{P} = \{p_1 = 0, \ldots, p_m = 0\}$ over variables $x_1, \ldots, x_n$ such that the polynomial equations are satisfiable over $\{0, 1\}$ if and only if the formula is satisfiable. A Nullstellensatz refutation is a set of polynomials $q_1, \ldots, q_m$ such that $\sum_i p_i q_i = 1$ – that is, the

polynomials q_i witness the fact that 1 is in the ideal generated by the p_i's and therefore they are not simultaneously satisfiable. The *degree* of a refutation is the maximal degree of the q_i's. The Polynomial Calculus (PC) is a dynamic version of the Nullstellensatz refutation system allowing proofs of potentially lower degree. The *semi-algebraic* systems Sherali-Adams (SA) and Sum-of-Squares (SoS) are further extension which refute families of polynomial inequalities.

There is a long history of degree lower bounds for all of these systems, again with the Tseitin formulas being the prototypical hard instance. The sequence of papers [35, 12, 36, 58] culminated in linear degree lower bound for Tseitin formulas over expander graphs. In [7], and later [34], lower bounds for more general *lifted* Tseitin formulas were proven by a reduction to communication complexity lower bounds, which in turn implies lower bounds for more general dynamic SoS systems as well as certain extensions of them. (See e.g., [23] (Chapter 5) for a simplified presentation of the SoS lower bound, as well as a survey of related results.)

By fairly standard low-degree reductions, lower bounds as well as integrality gaps have been proven for a variety of other problems. An integrality gap is aimed at proving lower bounds for approximation algorithms for NP-hard optimization algorithms, which is more general than proving lower bounds for solving the problem exactly. At a high level, these SoS integrality gap lower bounds show that no efficient SoS-based algorithm can approximate Max-3SAT any better than the trivial algorithm that achieves a $7/8$-approximation factor. SoS-based algorithms capture a natural family of linear and semidefinite programming algorithms, and thus SoS lower bounds rule out a large and natural family of algorithms for approximating NP-hard optimization problems.

Extended Formulations of Linear Programs.

In a beautiful line of work, the above-mentioned SoS lower bounds have been central to proving strong lower bounds for *extended formulations* of linear programs. More specifically, Sherali-Adams degree bounds for Tseitin formulas were used to prove lower bounds for LP (linear programming) extended formulations [13, 45], and similarly SoS bounds were shown to imply lower bounds for SDP (semidefinite programming) extended formulations. These lower bounds on extension complexity are again reductive – they show that, given an extended formulation for the optimization problem, the algorithm implies a much simpler SA-based algorithm for the same problem. That is, they prove in a very constructive sense that an algorithm coming from the larger family of polynomial-size extended formulations, actually implies the much simpler SA-based algorithm. Göös, Jain and Watson [33] proved the first truly exponential lower bounds on extension complexity by a reduction to a simpler lower bound for the Tseitin formulas.

We cannot begin to do justice to this fascinating topic and developments, but refer the reader to [23] for a comprehensive treatment.

1.5 Cutting Planes and Tseitin Formulas

The method of using cutting planes for inference in the study of polytopes in integer programming was first described by Gomory [32], modified and shown to be complete by Chvátal [14], and first analyzed for its efficiency as a proof system in [18].

Cutting Planes proofs manipulate integer linear inequalities. A CNF formula $F = (C_1, \ldots, C_m)$ is translated into a system of linear inequalities as follows: for each variable x_i add the inequalities $x_i \geq 0$ and $x_i \leq 1$. For each clause $C_i = \bigvee_{i \in P} x_i \vee \bigvee_{i \in N} \neg x_i$, add

$$\sum_{i \in P} x_i + \sum_{i \in N} (1 - x_i) \geq 1.$$

It can be checked that F is satisfiable if and only if there exists an assignment in $\{0,1\}^n$ satisfying this system of inequalities.

Given a system of linear inequalities $Ax \geq b$, a Cutting Planes (CP) derivation of $Ax \geq b$ is a sequence of inequalities $\{c_i x \geq d_i\}_{i \in [t]}$, where $c_i \in \mathbb{Z}^n$ is a vector and $d_i \in \mathbb{Z}$, such that every $c_i x \geq d_i$ either belongs to $Ax \geq b$, or is obtained from earlier inequalities by a *Chvátal-Gomory cut* (CG cut). A CG cut consists of two steps

- *Linear Combination*: From previously derived inequalities $(c_{i_1} x \geq d_{i_1}), \ldots, (c_{i_k} x \geq d_{i_k})$ and $\lambda_{i_1}, \ldots, \lambda_{i_k} \geq 0$, let $ax \geq b$ be such that $a = \sum_{j=1}^{k} \lambda_{i_j} c_{i_j}$ and $b = \sum_{j=1}^{k} \lambda_{i_j} d_{i_j}$.
- *division*: Derive $ax \geq \lceil b \rceil$.

A CP derivation is a *refutation* if the final inequality is $0 \geq 1$.

The *size* of the refutation is the number of lines (inequalities) in the refutation (which is polynomially related to the bit-size complexity [18]). Associated with any CP refutation is a directed acyclic graph labelled with the inequalities in the refutation, such that (i) each leaf is an inequality from $Ax \geq b$; (ii) intermediate nodes follow from their children by a CG cut; (iii) the root is $0 \geq 1$. A CP refutation is *tree-like* if the graph is a tree.

History of the Complexity of Cutting Planes Refutations.

There is a long history of lower bounds for CP, beginning with a paper by Impagliazzo, Pitassi and Urquhart [43] who proved lower bounds on tree-like CP proofs by a reduction to the communication complexity of an associated search problem. The first lower bounds for general CP were established by Pudlak [50] and independently by Bonet, Pitassi and Raz [11] (for the case of bounded coefficients) using the method of feasible interpolation [46]. However, the formulas for which these lower bounds were obtained had to be specially tailored to the method of feasible interpolation, and it remained a longstanding open problem to resolve the complexity of the Tseitin formulas, as well as random instances, for CP.

Recently, [24, 41] proved super-polynomial lower bounds on the size of Cutting Planes refutations for random k-CNF formulas, for $k = O(\log n)$. This is the first example of a proof system for which lower bounds on random formulas did not follow from lower bounds for the Tseitin formulas. Following this, Garg et al. [29] showed that Urquhart's lower bound for Resolution implied a lower bounds on the size of CP refutation of *lifted* Tseitin formulas by establishing a general lifting theorem from Resolution lower bounds to CP lower bounds. Recently, Dadush and Tiwari [19] showed that CP has quasi-polynomial size proofs of the Tseitin formulas.

Lower Bounds via Communication Complexity.

Here we sketch the main idea behind all of the aforementioned lower bounds. At their core, all of these lower bounds are reductions to the communication complexity[2] of an associated search problem.

Definition 2 (Canonical Search Problem) Let $F = (C_1, \ldots, C_m)$ be a CNF formula and (X, Y) be a partition of its variables. The associated search problem $Search_{X,Y}(F) \subseteq \{0,1\}^X \times \{0,1\}^Y \times [m]$ asks, given $(x, y) \in \{0,1\}^X \times \{0,1\}^Y$ to find the index of a clause $i \in [m]$ that is violated by (x, y), i.e. $C_i(x, y) = 0$.

The use of communication complexity to obtain proof complexity lower bounds was pioneered in the work of Impagliazzo, Pitassi, and Urquhart [43]. We illustrate their main technique in Lemma 4 for the case of *low-weight* CP refutations in which the sum of the magnitude of the coefficients of each line require at most $t = O(\log n)$ bits to express.

Lemma 4 (Impagliazzo et al. [43]) *Let F be an unsatisfiable formula and (X, Y) be any partition of the variables. If there is a tree-like CP refutation of F of size s in which every line can be expressed in t bits, then the communication complexity of solving $Search_{X,Y}(F)$ is $O(t \log s)$.*

Proof First note that with at most a polynomial increase in the size, we can assume that the graph of the refutation has fan-in at most 2; let s' be the size of the fan-in 2 proof. Let the input to Alice and Bob be $x \in \{0,1\}^X$ and $y \in \{0,1\}^Y$ respectively. The proof is by induction on s'. Viewing the refutation π as a tree, there exists an intermediate node l such that the number of nodes above l is between $s'/3$ and $2s'/3$. Denote by π_1 the subtree rooted at l, and let the remainder $\pi \setminus \pi_1$ be denoted by π_0.

Alice and Bob will evaluate $l := ax + by \geq d$. To do so, Alice evaluates $ax - d$ and sends the result to Bob in $O(\log n)$ bits, who can then evaluate whether $ax + by \geq d$. If l is falsified under (x, y) then Alice and Bob proceed on the subtree π_1. Otherwise, they recurse on π_0. Both π_0 and π_1 have size at most $2s'/3$ and thus by induction they have communication protocols of size $t \log(2s'/3)$. Therefore, in total the number of bits communicated by the protocol solving the search problem on π is at most $t + t \log(2s'/3) \leq t \log s' = O(t \log s)$.

[2] We refer the reader to the excellent book [51] for definitions and a rigorous treatment of communication complexity.

The correctness of the protocol follows by the soundness of the refutation: if l is falsified by (x, y) then at least one child of l must also falsified by (x, y). Therefore, this procedure is guaranteed to arrive at at a clause of F which is falsified by (x, y).□

Using this reduction together with the monotone formula lower bounds of Raz and Wigderson [52], Impagliazzo, Pitassi, and Urquhart [43] obtained lower bounds for tree-like CP. The issue of low coefficients is avoided by using randomized or real models of communication, both of which have short protocols for computing integer linear inequalities. Lower bounds on general CP size can be obtained by switching to an appropriate *dag-like* model of communication [53, 62].

1.5.1 A Warmup to an Upper Bounds on the Tseitin Formulas

Reductions to communication complexity form the backbone of all of the known lower bounds for CP. This presented a significant barrier against obtaining lower bounds on the Tseitin formulas: the search problem associated with the Tseitin formulas has a short communication protocol. We will first give a general strategy for finding a violated constraint in $\mathrm{TS}(G, l)$ which will be useful in the following section. In Lemma 5 we will show that this strategy can be implemented with a short communication protocol and in Section 1.5.2 we will describe how this upper bound strategy can be implemented in CP.

Fix an assignment $(x, y) \in X \times Y$. For $U \subset V$ let $E[U] := E[U, V \setminus U]$ and let $l(U) := \bigoplus_{u \in U} l(u)$ be the parity of the total labelling on U. At each recursive round we will maintain a subset of $U \subseteq V$ and a value $\delta \in \{0, 1\}$ such that $\delta \neq l(U)$ (initially $U = V$, $\delta = 0$) such that we have determined that $\sum_{e \in E[U]} x_e = \delta \pmod 2$. This will ensure that U contains a vertex whose constraint is falsified by (x, y). Indeed, suppose the constraints of U are satisfiable, then

$$l(U) = \sum_{u \in U} \sum_{e : u \in e} x_e \pmod 2 = \sum_{e=(u,v) : u,v \in U} 2x_e + \sum_{e \in E[U]} x_e \pmod 2 = \delta,$$

which contradicts $\delta \neq l(U)$.
At each round, perform the following:

1. Partition U into two halves, U_1 and U_2.
2. Determine $\delta_1 = \sum_{e \in E[U_1]} x_e \pmod 2$ and $\delta_2 = \sum_{e \in [U_2]} x_e \pmod 2$.
3. Recurse on U_1 if $\delta_1 \neq l(U_1)$ and otherwise on U_2 when $\delta_2 \neq l(U_2)$.

The recursion halts when $|U| = 1$, at which point we have found a violated constraint.

Lemma 5 *Fix a graph G, an odd labelling l, and let (X, Y) be any partition of the variables of $TS(G, l)$. There is a $O(\log n)$ communication protocol solving $Seach_{X,Y}(TS(G, l))$.*

Proof Given $x \in X$ and $y \in Y$ respectively, Alice and Bob will implement the aforementioned upper bound strategy. To do so, we must show that they are able to

perform step (2) while communicating $O(1)$ bits. Under the partition (X, Y) of the edges E, the first sum in (2) can be written as $\sum_{e \in E[U_1], e \in X} x_e + \sum_{e \in E[U_1], e \in Y} y_e$. Alice evaluates $\sum_{e \in E[U_1], e \in Y} y_e$ and sends the answer (a single bit) to Bob, who is then able to determine δ_1 and send the answer to Alice. Similarly, they are able to compute δ_2 in 2 bits of communication.

Each recursive round halves the size of the set U. Thus, there are $\log n$ rounds, each costing 4 bits of communication. □

1.5.2 Tseitin Formulas are Easy For Cutting Planes

In a surprising breakthrough result, Dadush and Tiwari [19] refuted the conjecture that the Tseitin formulas are hard for CP.

Theorem 2 (Dadush-Tiwari [19]) *For any graph G and odd labelling l there is a quasi-polynomial size CP refutation of $TS(G, l)$.*

Their proof shows that a known refutation of the Tseitin formulas in the stronger Stabbing Planes proof system (introduced by Beame et al. [4]) could be simulated in CP. In the remainder we will describe the proof of this upper bound.

Definition 3 (Stabbing Planes) Let F be an unsatisfiable system of linear inequalities. A Stabbing Planes (SP) refutation of F is a directed binary tree in which each edge is labelled with a linear integral inequality satisfying the following conditions:

- *Internal Nodes*: For any internal node u, if the right outgoing edge is labelled with $ax \geq b$, then the left outgoing edge is labelled with $ax \leq b - 1$.
- *Leaves*: Each leaf node u is labelled with a non-negative linear combination of inequalities in F with inequalities along the path leading to u that yield $0 \geq 1$.

Associated with every node u in the SP proof is a polytope P_u formed by the intersection of the inequalities in F together with the inequalities labelling the root-to-u path. The polytopes labelling the leaves are empty. The pair of inequalities $(ax \leq b - 1,\ ax \geq b)$ is called the *query* corresponding to the node. The *slab* corresponding to the query is $\{x \in \mathbb{R}^n \mid b - 1 < ax < b\}$. The *size* of a refutation is the number of queries in the tree, which is polynomially equivalent to the bit-length needed to encode a description of the entire proof tree [19].

Lemma 6 (Beame et al. [4]) *For any G and odd labelling l there is a quasi-polynomial size SP refutation of $TS(G, l)$.*

Proof (Proof Sketch) We show that the upper bound strategy from the previous section can be implemented in SP. However, this implementation will be lossy because SP cannot reason about mod 2 equations directly. Instead of maintaining that we have some $\delta \in \{0, 1\}$ such that $\sum_{e \in E[U]} x_e = \delta$ and $\delta \neq l(U)$, we will strengthen our invariant to require that we have determined $\sum_{e \in E[U]} x_e$ exactly. That is, we have determined that $\sum_{e \in E[U]} x_e = \delta$ for $\delta \in \{0, \ldots, |E[U]|\}$.

At each round we will perform the following:

1. Partition U into two halves, U_1 and U_2.
2. Determine $\delta_1 = \sum_{e \in E[U_1]} x_e$ and $\delta_2 = \sum_{e \in E[U_2]} x_e$.
3. Recurse on U_1 if $l(U_1) \neq \delta_1 \pmod 2$ and otherwise on U_2 when $l(U_2) \neq \delta_2 \pmod 2$. If neither holds then we have a contradiction to our inductive assumption and we can derive $0 \geq 1$.

The recursion stops when $|U| = 1$, at which point we obtain a contradiction between δ and the constraint of the single vertex in U.

It remains to show how to perform step (2) in SP: First, query

$$\Big(\sum_{e \in E[U_1]} x_e \leq \delta_1 - 1, \quad \sum_{e \in E[U_1]} x_e \geq \delta_1 \Big) \qquad \text{for } \delta_1 = 1, \ldots, |E[U_1]|,$$

where the ith query is attached to the right child of the $(i-1)$st query (see Figure 1.3). Each leaf of this tree has determined that $\sum_{e \in E[U_1]} x_e = \delta_1$ (for the edge cases $\delta_1 = 0, \delta_1 = |E[U_1|$ we use the axioms $x_i \geq 0$ and $x_i \leq 1$). At each leaf of this tree we query

$$\Big(\sum_{e \in E[U_2]} x_e \leq \delta_2 - 1, \quad \sum_{e \in E[U_2]} x_e \geq \delta_2 \Big) \qquad \text{for } \delta_2 = 1, \ldots, |E[U_2]|,$$

where again the ith query is attached to the right child of the $(i-1)$st. This completes the simulation of (2).

Each recursive step halves the size of the set U under consideration and converges in $O(\log n)$ recursive steps. Each recursive step can be implemented in $O(|E|^2)$ queries. Thus, the total proof size is quasi-polynomial. □

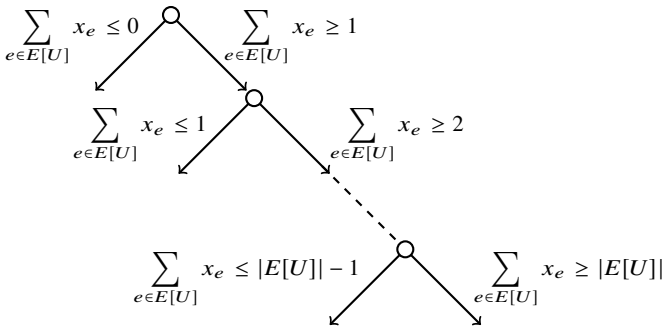


Fig. 1.3 The first tree of SP queries in a recursive step.

Simulating the Tseitin Refutation in CP.

Dadush and Tiwari showed that Lemma 6 could be converted into a CP refutation. Let us first recall some basic facts about polytopes and the geometry of CP proofs.

A hyperplane $ax \geq b$ is *valid* for a polytope P if $ax \geq b$ for every $x \in P$. $F := P \cap \{x : ax = b\}$ is *face* of P if at least one of $ax \geq b$, $ax \leq b$ is valid for P. We can view a CP refutation of a set of linear inequalities F as a sequence of polytopes $F = P_0, P_1, \ldots, P_t = \emptyset$ where P_i is derived from P_{i-1} by a Chvátal-Gomory (CG) cut, which corresponds to taking a hyperplane $ax \geq b$ that is valid for P_{i-1} and shifting it to the nearest integral point (see Figure 1.5). The principal difference between CP and SP is that SP can cut *within* the current polytope by removing a slab from it, while CP can only cut on the boundary of the current polytope.

The key observation is that the SP refutation from Lemma 6 works from the boundary of the polytope $P = \mathrm{TS}(G, l)$ inwards. Indeed, if the vertices U are partitioned into U_1 and U_2, then the first branch (see Figure 1.3) corresponds to refuting the face $\sum_{e \in E[U_1]} x_e = 0$, the second to refuting $\sum_{e \in E[U_1]} x_e = 1$, and so on (see Figure 1.4). More generally, for every query $(ax \leq b, ax \geq b + 1)$ corresponding to some node u in the SP refutation of $\mathrm{TS}(G, l)$, $ax \geq b$ will be valid for the current polytope P_u:

- If $b = 0$ this follows from the axioms $x_i \geq 0$.
- If $b > 0$ then this follows because the SP proof queries $(ax \leq b, ax \geq b + 1)$ sequentially from $b = 1, 2, \ldots$ and so $ax \geq b$ is one of the defining inequalities for P.

Therefore, the SP proof refutes $TS(G, l)$ by recursively cutting away at the sides of the polytope.

A second important observation is that if we have a CP derivation of a polytope P' from P such that $P' \cap \{x : ax = b\} = \emptyset$ (i.e. we have refuted the face $\{x : ax = b\}$) then CP can drive $ax \geq b + 1$ from P'. Indeed, there must be some $\varepsilon \in (0, 1)$ such that $ax \geq b + \varepsilon$ is valid for P' and so $ax \geq \lceil b + \varepsilon \rceil$ is a GC cut from P'.

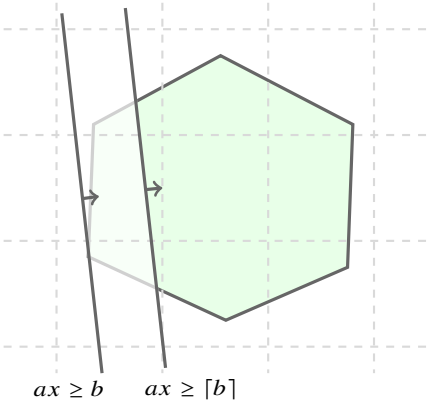


Fig. 1.4 A CG cut $ax \geq \lceil b \rceil$, the shift of a valid halfspace $ax \geq b$ to the nearest integer point. The grid intersection points represent integer points.

Together, these observations show that each recursive round of the SP proof looks locally like a CP proof. The challenge is to show that we can unroll the recursion. We would like to show that if we have CP refutations of $P \cap \{x : \sum_{e \in E[U]} x_e = \delta\}$ for all $\delta = 0, \ldots, |E[U]|$ then we can glue these refutations together to form a refutation P. The following Lemma due to Shrijver [59] will allow us to do this.

Lemma 7 (Shrijver [59]) *Let P be a polytope and F be a face of P. If F' is obtained from F by a CG cut then there is a polytope P' obtained by a CG cut from P such that $P' \cap F \subseteq F'$.*

The high-level idea of the proof is as follows: Let $P = \{Ax \leq b\}$, $F = \{A_0 x = b_0, A_1 x \leq b_1\}$ and $ax \geq b$ be the CG cut which obtains F' from F. Observe that shifting $ax \geq b$ by factors of $A_0 x \leq b_0$ does not change its effect on F. Since we care only about the resulting polytope when restricted to the face F, we can shift $ax \geq b$ by $A_0 x \leq b_0$ so that it no longer depends on $A_0 x \geq b_0$ and is therefore a valid cut from P.

Repeated application of this lemma allows us to simulate a refutation of a face F on P itself.

Corollary 1 *Suppose that $ax \geq b$ is valid for P, $a \in \mathbb{Z}^n, b \in \mathbb{Z}$, and let $F := P \cap \{x : ax = b\}$. Let $F = F_0, \ldots, F_k = \emptyset$ be a CP refutation of F, then there is a CP derivation $P = P_0, \ldots, P_k, P_{k+1}$ such that $P_{k+1} \subseteq P \cap \{x : ax \geq b + 1\}$.*

Proof For $i = 1, \ldots, k$, apply Lemma 7 to obtain P_i from F_i and P_{i-1} such that $P_i \cap F \subset F_i$ and therefore $P_k \cap F = \emptyset$. Because $ax \geq b$ is valid for P and $P_k \cap \{x : ax = b\} = \emptyset$ it follow that there exists $0 < \varepsilon \leq 1$ such that $ax \geq b + \varepsilon$ is valid for P. Therefore, $P_{k+1} := P_k \cap \{x : ax \geq b + 1\}$ is a CG cut from P_k. □

We now sketch the proof of the CP refutation of the Tseitin formulas by Dadush and Tiwari.

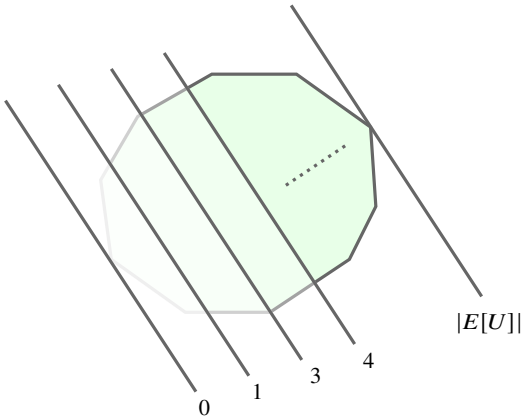


Fig. 1.5 The faces $\sum_{e \in E[U]} x_e = \delta$ for $\delta \in [E[U]|]$ of the TS(G, l) polytope refuted sequentially in the SP refutation.

Proof (Proof Sketch of Theorem 2.) The proof is a post-order traversal of the SP refutation of TS(G, l). Consider some node in the SP refutation corresponding to a query $(ax \leq b,\ ax \geq b+1)$, and let P be the polytope associated with this node. Suppose that we have CP refutations of the left and right children $P \cap \{x : ax = b\}$ and $P \cap \{x : ax = b+1\}$. We will construct a CP refutation of P as follows:

1. Apply Corollary 1 to the refutation of the left child in order to obtain a CP derivation $P = P_0, \ldots, P_{t+1}$ such that $P_{t+1} \subseteq P \cap \{x : ax \geq b+1\}$.
2. Append the CP refutation of the right child in order to refute P_{t+1}.

Since the root of the SP refutation corresponds to the polytope $TS(G, l)$, this procedure will produce a CP refutation of $TS(G, l)$. Observe that this simulation preserves the size of the SP refutation. □

In a followup work Fleming et al. [22] gave an alternative proof of Theorem 2 and extended it to hold for any unsatisfiable system of linear equations over a finite field. To do this, they showed that CP can quasi-polynomially simulate any SP proof provided that the coefficients of the proof are quasi-polynomially bounded.

1.6 Concluding Remarks

We end with several related open problems.

Are Optimal Resolution Refutations of Tseitin Formulas Regular?

In his paper [63], Tseitin makes the following remarks about the heuristic interpretation of the regularity restriction:

> The regularity condition can be interpreted as a requirement for not proving intermediate results in a form stronger than that in which they are later used (if A and B are disjunctions such that $A \subseteq B$, then A may be considered to be the stronger assertion of the two); if the derivation of a disjunction containing a variable ξ involves the annihilation of the latter, then we can avoid this annihilation, some of the disjunctions in the derivation being replaced by "weaker" disjunctions containing ξ.

These heuristic remarks of Tseitin suggest that there is always a regular Resolution refutation of minimal size, as in the case of tree Resolution. Consequently, some authors tried to extend Tseitin's results to general Resolution by showing that regular Resolution can simulate general Resolution efficiently. The results of [42, 31, 2, 66] show that these attempts were doomed to failure. Despite this negative result showing that regular Resolution can be substantially weaker than general Resolution in the worst case, it remains open for natural examples such as counting principles. In particular, Urquhart [64] conjectured that the minimal-size Resolution refutations for Tseitin are always regular.

There has been significant progress towards resolving this conjecture for constant-degree graphs. For constant-degree graphs, In this case, the results of [27, 3] imply a

regular Resolution proof of size $2^{O(tw(G))} \log |V|$, where $tw(G)$ is the tree-width of G. As observed by [44], the technique of Galesi et al. [26] implies a $2^{\Omega(tw(G)^{1/10})}$ lower bound for Resolution. Recently [44] proved near-optimal $2^{\tilde{\Omega}(tw(G))}$ lower bounds for regular Resolution.

Further Lower Bounds for Bounded-depth Frege.

It is still an open to prove optimal lower bounds for the propositional pigeonhole principle. As mentioned above, the best known lower bound is exponential in $n^{1/2^{O(d)}}$ for depth-d Frege proofs, whereas the best known upper bound is exponential in $n^{1/d}$. Another longstanding open problem is to prove lower bounds for bounded-depth Frege proofs for k-CNF random formulas; in this case there are no nontrivial lower bounds known for $d > 2$. Lower bounds for the Tseitin formulas have typically paved the way for obtaining lower bounds on random formulas. Indeed, k-XOR instances can be viewed as k-CNF formulas with additional clauses, and thus random k-XOR lower bounds imply lower bounds on random k-CNFs. While near-optimal lower bounds on the Tseitin formulas are known, these lower bounds rely on certain structure of the underlying graph that is not available for random k-XOR instances.

Finally, the most longstanding open problem concerning bounded-depth Frege systems is to prove lower bounds for bounded-depth Frege systems over the basis which includes mod-p gates, for any prime $p \geq 2$. It is conjectured that the Tseitin formulas (mod 2) should be hard for bounded-depth Frege systems over the basis which includes mod-p gates, for any $p \neq 2$.

Unprovability of $\mathsf{P} \neq \mathsf{NP}$.

In [53] Razborov showed that if one assumes the existence of strong pseudo-random generators, then certain systems of Bounded Arithmetic cannot prove circuit lower bounds, thus ruling out any approach for proving $\mathsf{NP} \not\subseteq \mathsf{P/poly}$ that could be implemented in these systems. As well, this implies the same result for any propositional proof system which admits feasible interpolation by monotone circuits (equivalently, any proof system that can be simulated by the RCC_1 [24] proof system). In several followup works Razborov established unconditionally that the Polynomial Calculus [54], Resolution [55], and Resolution over $o(\log n)$-DNFs [56] do not possess short proofs of $\mathsf{NP} \not\subseteq \mathsf{P/poly}$. It remains an open problem to extend these lower bounds to stronger systems such as bounded-depth Frege.

1.7 Acknowledgments

Toniann Pitassi would like to express her gratitude to Alasdair Urquhart, and acknowledge his enormous influence on her academic and intellectual development. He is a true scholar, and has been a source of great ideas and inspiration.

Both authors would like to thank Paul Beame for reviewing this chapter.

References

1. Miklós Ajtai. The complexity of the pigeonhole principle. In *29th Annual Symposium on Foundations of Computer Science, White Plains, New York, USA, 24-26 October 1988*, pages 346–355. IEEE Computer Society, 1988.
2. Michael Alekhnovich, Jan Johannsen, Toniann Pitassi, and Alasdair Urquhart. An exponential separation between regular and general resolution. *Theory of Computing*, 3(1):81–102, 2007.
3. Michael Alekhnovich and Alexander A. Razborov. Satisfiability, branch-width and tseitin tautologies. *Comput. Complex.*, 20(4):649–678, 2011.
4. Paul Beame, Noah Fleming, Russell Impagliazzo, Antonina Kolokolova, Denis Pankratov, Toniann Pitassi, and Robert Robere. Stabbing planes. In *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, pages 10:1–10:20, 2018.
5. Paul Beame, Russell Impagliazzo, Jan Krajícek, Toniann Pitassi, and Pavel Pudlák. Lower bound on hilbert's nullstellensatz and propositional proofs. In *35th Annual Symposium on Foundations of Computer Science, Santa Fe, New Mexico, USA, 20-22 November 1994*, pages 794–806. IEEE Computer Society, 1994.
6. Paul Beame and Toniann Pitassi. Propositional proof complexity: Past, present and future. *Electronic Colloquium on Computational Complexity (ECCC)*, 5(67), 1998.
7. Paul Beame, Toniann Pitassi, and Nathan Segerlind. Lower bounds for lov[a-acute]sz–schrijver systems and beyond follow from multiparty communication complexity. *SIAM J. Comput.*, 37(3):845–869, 2007.
8. Stephen Bellantoni, Toniann Pitassi, and Alasdair Urquhart. Approximation and small depth frege proofs. In *Proceedings of the Sixth Annual Structure in Complexity Theory Conference, Chicago, Illinois, USA, June 30 - July 3, 1991*, pages 367–390. IEEE Computer Society, 1991.
9. Eli Ben-Sasson. Hard examples for the bounded depth frege proof system. *Comput. Complex.*, 11(3-4):109–136, 2002.
10. Eli Ben-Sasson and Avi Wigderson. Short proofs are narrow - resolution made simple. *J. ACM*, 48(2):149–169, 2001.
11. Maria Luisa Bonet, Toniann Pitassi, and Ran Raz. Lower bounds for cutting planes proofs with small coefficients. *J. Symb. Log.*, 62(3):708–728, 1997.
12. Samuel R. Buss, Dima Grigoriev, Russell Impagliazzo, and Toniann Pitassi. Linear gaps between degrees for the polynomial calculus modulo distinct primes. *J. Comput. Syst. Sci.*, 62(2):267–289, 2001.
13. Siu On Chan, James R. Lee, Prasad Raghavendra, and David Steurer. Approximate constraint satisfaction requires large LP relaxations. *J. ACM*, 63(4):34:1–34:22, 2016.
14. Vasek Chvátal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discret. Math.*, 4(4):305–337, 1973.
15. Vasek Chvátal and Endre Szemerédi. Many hard examples for resolution. *J. ACM*, 35(4):759–768, 1988.
16. Matthew Clegg, Jeff Edmonds, and Russell Impagliazzo. Using the groebner basis algorithm to find proofs of unsatisfiability. In Gary L. Miller, editor, *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA, May 22-24, 1996*, pages 174–183. ACM, 1996.

17. Stephen A. Cook and Robert A. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 6:169–184, 1979.
18. William J. Cook, Collette R. Coullard, and György Turán. On the complexity of cutting-plane proofs. *Discrete Applied Mathematics*, 18(1):25–38, 1987.
19. Daniel Dadush and Samarth Tiwari. On the complexity of branching proofs. *CoRR*, abs/2006.04124, 2020.
20. Martin Davis, G. Logemann, and D. Loveland. A machine program for theorem proving. *Communications of the Association for Computing Machinery*, 5:394–397, 1962.
21. Martin Davis and Hilary Putnam. A computing procedure for quantification theory. *J. ACM*, 7(3):201–215, 1960.
22. Noah Fleming, Mika Göös, Russell Impagliazzo, Toniann Pitassi, Robert Robere, Li-Yang Tan, and Avi Wigderson. On the power and limitations of branch and cut. *Electron. Colloquium Comput. Complex.*, 28:12, 2021.
23. Noah Fleming, Pravesh Kothari, and Toniann Pitassi. Semialgebraic proofs and efficient algorithm design. *Foundations and Trends in Theoretical Computer Science*, 14(1-2):1–221, 2019.
24. Noah Fleming, Denis Pankratov, Toniann Pitassi, and Robert Robere. Random (log n)-cnfs are hard for cutting planes. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 109–120, 2017.
25. Ofer Gabber and Zvi Galil. Explicit constructions of linear-sized superconcentrators. *J. Comput. Syst. Sci.*, 22(3):407–420, 1981.
26. Nicola Galesi, Dmitry Itsykson, Artur Riazanov, and Anastasia Sofronova. Bounded-depth frege complexity of tseitin formulas for all graphs. In Peter Rossmanith, Pinar Heggernes, and Joost-Pieter Katoen, editors, *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen, Germany*, volume 138 of *LIPIcs*, pages 49:1–49:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.
27. Nicola Galesi, Navid Talebanfard, and Jacobo Torán. Cops-robber games and the resolution of tseitin formulas. In Olaf Beyersdorff and Christoph M. Wintersteiger, editors, *Theory and Applications of Satisfiability Testing - SAT 2018 - 21st International Conference, SAT 2018, Held as Part of the Federated Logic Conference, FloC 2018, Oxford, UK, July 9-12, 2018, Proceedings*, volume 10929 of *Lecture Notes in Computer Science*, pages 311–326. Springer, 2018.
28. Zvi Galil. On the complexity of regular resolution and the davis-putnam procedure. *Theor. Comput. Sci.*, 4(1):23–46, 1977.
29. Ankit Garg, Mika Göös, Pritish Kamath, and Dmitry Sokolov. Monotone circuit lower bounds from resolution. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 902–911. ACM, 2018.
30. A. Goerdt. Davis-Putnam resolution versus unrestricted resolution. *Annals of Mathematics and Artificial Intelligence*, 6:1–3, 1992.
31. Andreas Goerdt. Regular resolution versus unrestricted resolution. *SIAM J. Comput.*, 22(4):661–683, 1993.
32. Ralph E. Gomory. Outline of an algorithm for integer solutions to linear programs *and* an algorithm for the mixed integer problem. In Michael Jünger, Thomas M. Liebling, Denis Naddef, George L. Nemhauser, William R. Pulleyblank, Gerhard Reinelt, Giovanni Rinaldi, and Laurence A. Wolsey, editors, *50 Years of Integer Programming 1958-2008 - From the Early Years to the State-of-the-Art*, pages 77–103. Springer, 2010.
33. Mika Göös, Rahul Jain, and Thomas Watson. Extension complexity of independent set polytopes. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:70, 2016.
34. Mika Göös and Toniann Pitassi. Communication lower bounds via critical block sensitivity. In David B. Shmoys, editor, *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 847–856. ACM, 2014.
35. Dima Grigoriev. Tseitin's tautologies and lower bounds for nullstellensatz proofs. In *39th Annual Symposium on Foundations of Computer Science, FOCS '98, November 8-11, 1998, Palo Alto, California, USA*, pages 648–652. IEEE Computer Society, 1998.

36. Dima Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theor. Comput. Sci.*, 259(1-2):613–622, 2001.
37. Armin Haken. The intractability of resolution. *Theor. Comput. Sci.*, 39:297–308, 1985.
38. Johan Håstad. On small-depth frege proofs for tseitin for grids. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 97–108, 2017.
39. Johan Håstad, Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. An average-case depth hierarchy theorem for boolean circuits. *J. ACM*, 64(5):35:1–35:27, 2017.
40. Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.
41. Pavel Hrubes and Pavel Pudlák. Random formulas, monotone circuits, and interpolation. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 121–131, 2017.
42. Wenqui Huang and Xiangdong Yu. A DNF without regular shortest consensus path. *SIAM J. Comput.*, 16(5):836–840, 1987.
43. Russell Impagliazzo, Toniann Pitassi, and Alasdair Urquhart. Upper and lower bounds for tree-like cutting planes proofs. In *Proceedings of the Ninth Annual Symposium on Logic in Computer Science (LICS '94), Paris, France, July 4-7, 1994*, pages 220–228. IEEE Computer Society, 1994.
44. Dmitry Itsykson, Artur Riazanov, Danil Sagunov, and Petr Smirnov. Almost tight lower bounds on regular resolution refutations of tseitin formulas for all constant-degree graphs. *Electronic Colloquium on Computational Complexity (ECCC)*, 26:178, 2019.
45. Pravesh K. Kothari, Raghu Meka, and Prasad Raghavendra. Approximating rectangles by juntas and weakly-exponential lower bounds for LP relaxations of csps. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 590–603. ACM, 2017.
46. Jan Krajícek. Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic. *J. Symb. Log.*, 62(2):457–486, 1997.
47. Jan Krajícek, Pavel Pudlák, and Alan R. Woods. An exponential lower bound to the size of bounded depth frege proofs of the pigeonhole principle. *Electronic Colloquium on Computational Complexity (ECCC)*, 1(18), 1994.
48. Toniann Pitassi, Paul Beame, and Russell Impagliazzo. Exponential lower bounds for the pigeonhole principle. *Comput. Complex.*, 3:97–140, 1993.
49. Toniann Pitassi, Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. Poly-logarithmic frege depth lower bounds via an expander switching lemma. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 644–657, 2016.
50. Pavel Pudlák. Lower bounds for resolution and cutting plane proofs and monotone computations. *J. Symb. Log.*, 62(3):981–998, 1997.
51. Anup Rao and Amir Yehudayoff. *Communication Complexity: and Applications*. Cambridge University Press, 2020.
52. Ran Raz and Avi Wigderson. Monotone circuits for matching require linear depth. In Harriet Ortiz, editor, *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, May 13-17, 1990, Baltimore, Maryland, USA*, pages 287–292. ACM, 1990.
53. Alexander Razborov. Unprovability of lower bounds on circuit size in certain fragments of bounded arithmetic. *Izvestiya Mathematics*, 59(1):205–227, 1995.
54. Alexander A. Razborov. Lower bounds for the polynomial calculus. *Comput. Complex.*, 7(4):291–324, 1998.
55. Alexander A. Razborov. Resolution lower bounds for perfect matching principles. *J. Comput. Syst. Sci.*, 69(1):3–27, 2004.
56. Alexander A Razborov. Pseudorandom generators hard for k-dnf resolution and polynomial calculus resolution. *Annals of Mathematics*, pages 415–472, 2015.
57. Alexander A. Razborov. Guest column: Proof complexity and beyond. *SIGACT News*, 47(2):66–86, 2016.

58. Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 593–602. IEEE Computer Society, 2008.
59. A. Schrijver. On cutting planes. In Peter L. Hammer, editor, *Combinatorics 79*, volume 9 of *Annals of Discrete Mathematics*, pages 291 – 296. Elsevier, 1980.
60. Nathan Segerlind. The complexity of propositional proofs. *Bulletin of Symbolic Logic*, 13(4):417–481, 2007.
61. Michael Sipser. The history and status of the P versus NP question. In S. Rao Kosaraju, Mike Fellows, Avi Wigderson, and John A. Ellis, editors, *Proceedings of the 24th Annual ACM Symposium on Theory of Computing, May 4-6, 1992, Victoria, British Columbia, Canada*, pages 603–618. ACM, 1992.
62. Dmitry Sokolov. Dag-like communication and its applications. *ECCC TR16-202*, 2017.
63. G.S. Tseitin. On the complexity of derivation in propositional calculus. *Studies in Constructive Mathematics and Mathematical Logic, Part 2, Consultants Bureau, New York-London*, pages 115–125, 1968.
64. Alasdair Urquhart. Hard examples for resolution. *J. ACM*, 34(1):209–219, 1987.
65. Alasdair Urquhart. The complexity of propositional proofs. *Bulletin of Symbolic Logic*, 1(4):425–467, 1995.
66. Alasdair Urquhart. A near-optimal separation of regular and general resolution. *SIAM J. Comput.*, 40(1):107–121, 2011.
67. Alasdair Urquhart and Xudong Fu. Simplified lower bounds for propositional proofs. *Notre Dame J. Formal Log.*, 37(4):523–544, 1996.