Rank Complexity Gap for Lovász-Schrijver and Sherali-Adams Proof Systems*

Stefan Dantchev
Department of Computer Science, Durham University
South Road, Durham, Co. Durham DH1 3LE, UK
s.s.dantchev@durham.ac.uk

ABSTRACT

We prove a dichotomy theorem for the rank of the uniformly generated (i.e. expressible in First-Order (FO) Logic) propositional tautologies in both the Lovász-Schrijver (LS) and Sherali-Adams (SA) proof systems. More precisely, we first show that the propositional translations of FO formulae that are universally true, i.e. hold in all finite and infinite models, have LS proofs whose rank is constant, independently from the size of the (finite) universe. In contrast to that, we prove that the propositional formulae that hold in all finite models but fail in some infinite structure require proofs whose SA rank grows poly-logarithmically with the size of the universe.

Up to now, this kind of so-called "Complexity Gap" theorems have been known for Tree-like Resolution and, in somehow restricted forms, for the Resolution and Nullstellensatz proof systems. As far as we are aware, this is the first time the Sherali-Adams lift-and-project method has been considered as a propositional proof system. An interesting feature of the SA proof system is that it is static and rank-preserving simulates LS, the Lovász-Schrijver proof system without semidefinite cuts.

Categories and Subject Descriptors

F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*Complexity of Proof Procedures*

General Terms

Theory

*This paper is dedicated to the memory of the late Mikhail Alekhnovich.

STOC'07, June 11–13, 2007, San Diego, California, USA.

Keywords

Propositional Proof Complexity, Lift and Project Methods, Lovász-Schrijver Proof System, Lower Bounds, Complexity Gap Theorems

1. INTRODUCTION

It is a trivial observation that the question as to whether a given propositional formula has a satisfying assignment can be reduced to a feasibility question for a certain Integer Linear Program (ILP), yet the easy reduction, when applied to propositional contradictions, gives rise to some very interesting

Propositional proof systems based on different methods for solving Integer Linear Programming

The two most important ILP-based proof systems are Cutting Planes (introduced as a general method for solving ILP in [3], and as a proof system in [2]) and Lovász-Schrijver (introduced as a general method for solving ILP in [6], and first considered as a proof system in [8]). A number of mixtures of ILP-based proof systems and algebraic proof systems are introduced and studied in [4].

Another method for solving ILP was proposed by Sherali and Adams in [11] but has not been explored as a propositional proof system up to now. The SA relaxation is interesting in that it is static and is stronger than LS, the Lovász-Schrijver relaxation without semidefinite cuts. More precisely, it is proven in [5] that rank k SA relaxation is tighter than rank k LS relaxation.

A number of lower bounds have been proven for the ILP-based proof systems. A non-comprehensive list of

Previous results

relevant to our work include the LS and LS_+ rank lower bounds for a number of specific tautologies from [1] as well as the LS rank lower bound for the Pigeon-Hole Principle (PHP) from [4]. No size lower bounds are known for LS, and it seems that the rank is the "right" complexity measure for LS in the same way that the degree is a good complexity measure for the algebraic proof systems.

All results in [1] and [4] are lower bounds for specific tautologies. The aim of this paper is to prove a very general LS rank lower bound that would apply to a large class of tautologies, namely those that can be expressed as FO sentences. Note that the Pigeon-Hole Principle as well as the Least Number Principle (stating that a finite order has a minimum element) are such tautologies, which have been much studied in the context of propositional proof complexity. Thus our motivation was to obtain a result, similar in spirit, to the so-called "Complexity Gap theorem" for Tree-like Resolution explicitly stated and proven by Riis in [10]:

THEOREM 1. *Given a FO sentence ψ which fails in all finite structures, consider its translation into a propositional CNF contradiction $\mathcal{C}_{\psi,n}$ where n is the size of the finite universe. Then either 1 or 2 holds:*
1. The sequence $\mathcal{C}_{\psi,n}$ has polynomial-size in n Tree-like Resolution refutation.
2. There exists a positive constant a such that for every n, every Tree-like Resolution refutation of $\mathcal{C}_{\psi,n}$ is of size at least 2^{an}.
Furthermore, 2 holds if and only if ψ has an infinite model.

We are able to prove similar result for LS. In its strongest form,

Our result

can be stated as follows.

THEOREM 2. *Given a FO sentence ψ which fails in all finite structures, consider its translation into a propositional CNF contradiction $\mathcal{C}_{\psi,n}$ where n is the size of the finite universe. Then either 1 or 2 holds:*
1. There exists a constant r such that $\mathcal{C}_{\psi,n}$ has rank-r LS refutation for every n.
2. There exists a positive constant a such that for every n, every SA refutation of $\mathcal{C}_{\psi,n}$ is of rank at least $(\log n)^a$.
Furthermore, 2 holds if and only if ψ has an infinite model.

The rest of the paper is organised as follows. In section 2, we define the two proof systems LS and SA, and explain the translation from a FO sentence into a family of finite propositional contradictions. The main part of the paper, section 3, contains the proof of Theorem 2. It is divided into two - we first prove the "easy", constant LS rank, case in subsection 3.1 and then move onto the "hard" non-constant lower bound for the SA rank in subsection 3.2. We finally discuss some open questions.

2. PRELIMINARIES

The Lovasz-Schrijver (LS) proof system

is a lift-and-project proof system: it operates on linear inequalities over continuous variables in $[0,1]$ by first "lifting" them into quadratic inequalities via multiplication by certain linear terms and then "projecting" these back into linear inequalities by taking linear combinations that in which the quadratic terms cancel out. Formally, we introduce two continuous $[0,1]$ variables, p_v and $p_{\neg v}$, for every propositional variable v of the original CNF formula φ, with the intention that $p_v = 1$ if $v = \top$ and $p_v = 0$ if $v = \bot$. We introduce the equations

$$p_v + p_{\neg v} - 1 = 0$$

for every propositional variable v as well as the inequalities

$$p_v \geq 0 \text{ and } p_{\neg v} \geq 0.$$

. We encode a clause $\bigvee_{j \in J} l_j$ of φ by the inequality

$$\sum_{j \in J} p_{l_j} - 1 \geq 0.$$

There are three kinds of derivation rules.

1. Multiply a linear (in)equality by a variable p_l where l is a literal in order to get a quadratic (in)equality. If the original (in)equality contained a term $p_{\neg l}$, the new quadratic term $p_l p_{\neg l}$ automatically vanishes, i.e. does not appear in the result. If the original (in)equality contained a term p_l, the new quadratic term p_l^2 automatically reduces to p_l.

2. Multiply any equation (either linear or quadratic) by a constant (real number) or multiply any inequality by a positive constant.

3. Add any two (in)equalities.

An LS derivation of an (in)equality from a set of (in)equalities (often called axioms) can be represented as a tree, whose leaves are labelled by axioms, and such that every internal node is labelled by an (in)equality that can be derived in a single step from the (in)equalities that label the children of the node. The root of the tree is labelled by the (in)equality that is finally derived. The *rank* of a LS derivation is the maximal number of derivation steps of the first kind, i.e. multiplications of a linear (in)equality by a variable, over all branches (paths form the root to a leaf) of the derivation tree. The rank of an (in)equality with respect to a set of axioms is the minimal rank over all possible derivations of the (in)equality from the axioms. Finally, the LS rank of an unsatisfiable CNF φ is the rank of the inequality $-1 \geq 0$ with respect to the axioms.

The Sherali-Adams (SA) proof system

is a static proof system, so we shall define SA proofs of rank k for every k, $0 \leq k < n$. More specifically, we shall encode a CNF formula φ over n propositional variables as a linear program $\mathcal{L}_k$ - a system of linear equations and inequalities over $\sum_{d=0}^{k+1} \binom{n}{d} 2^d$ continuous variables in the interval $[0,1]$.

We first introduce variables p_C for every conjunct $C = \bigwedge_{i \in I} l_i$ of no more than k variable-distinct literals l_i, $|I| \leq k$ (we shall often write $|C| \leq k$ instead). Ideally, we would like to have $p_C = 1$ if $C = \top$ and $p_C = 0$ otherwise. However, what we can express in linear programming is the inequalities

$$p_C \geq 0 \tag{1}$$

as well as the equations

$$p_{C \wedge v} + p_{C \wedge \neg v} = p_C \tag{2}$$

for every conjunct C with $|C| \leq k$ and every variable v. We also add the obvious equation

$$p_E = 1 \tag{3}$$

where E is the empty conjunct (of size 0, i.e. $E = \top$). Note that these equations do not depend on the initial CNF φ but only on the rank k. As for the clauses (disjuncts) of φ, we encode any such clause $D \equiv \bigvee_{j\in J} l_j$ by the following set of linear inequalities

$$\sum_{j\in J} p_{l_j \wedge C} \geq p_C \tag{4}$$

for every conjunct C with $|C| \leq k$. It is important to note that when writing indices of the form $l \wedge C$, the variable $p_{l\wedge C}$ vanishes whenever $\neg l$ is present in C.

Finally, we say that the CNF φ has an SA refutation of rank k if k is the smallest number for which the linear system $\mathcal{L}_k$, consisting of equations (2), (3) and inequalities (4), (1), is inconsistent. Thus, the system $\mathcal{L}_k$ itself serves as a refutation of φ that can be verified in polynomial (in its size) time by some poly-time linear programming algorithm. On the other hand, in order to establish a rank lower bound k for an SA refutation, we need to produce a valuation of the variables p_C with $|C| \leq k+1$ that satisfies the linear system $\mathcal{L}_k$.

It is not hard to see that

SA simulates LS.

More precisely, a rank k refutation in LS can be transformed into a rank k refutation in SA. We omit the proof here and refer to [5] instead.

Translation of FO sentences into propositional CNF formulae.

We use the language of FO logic with equality but without function symbols, i.e. we only allow relation symbols as well as constants. We assume that the FO sentence is in prenex normal form. The purely universal case is easy - a formula of the form

$$\forall x_1, x_2, \ldots x_k\ \mathcal{F}(x_1, x_2, \ldots x_k),$$

where $\mathcal{F}$ is quantifier-free, is translated into propositional CNF as follows. Let us first consider $\mathcal{F}(x_1, x_2, \ldots x_k)$ as a propositional formula over propositional variables of two different kinds: $R\left(x_{i_1}, x_{i_2}, \ldots x_{i_p}\right)$, where R is a p-ary predicate symbol, and $(x_i = x_j)$. We transform $\mathcal{F}$ into CNF and then take the union of all such CNF formulae for $x_1, x_2, \ldots x_k$ ranging over $[n]^k$ (assuming the finite universe is $[n] = \{1, 2, \ldots n\}$). The variables of the form $(x_i = x_j)$ evaluate to either true or false, and we are left with variables of the form $R\left(x_{i_1}, x_{i_2}, \ldots x_{i_p}\right)$ only. We map the constant symbols $a_1, a_2, \ldots a_m$ into the first m elements of the universe $[n]$. Note that this assumes - without loss of generality as far as any reasonable proof system is concerned - that all constants in the original FO formula are interpreted by distinct elements.

The general case - a formula of the form

$$\forall x_1 \exists y_1 \ldots \forall x_k \exists y_k\ \mathcal{F}(\overline{x}, \overline{y})$$

can be reduced to the previous case by Skolemisation. We introduce Skolem relations $S_i(x_1, x_2, \ldots x_i, y_i)$ for $1 \leq i \leq k$. $S_i(x_1, x_2, \ldots x_i, y_i)$ witnesses y_i for any given $x_1, x_2, \ldots x_i$, so we need to add clauses stating that such a witness always exists, i.e.

$$\bigvee_{y_i=1}^{n} S_i(x_1, x_2, \ldots x_i, y_i) \tag{5}$$

for all $(x_1, x_2, \ldots x_i) \in [n]^i$. The original formula can be transformed into the following purely universal one

$$\forall \overline{x}, \overline{y} \bigwedge_{i=1}^{k} S_i(x_1, \ldots x_i, y_i) \rightarrow \mathcal{F}(\overline{x}, \overline{y}). \tag{6}$$

We shall call clause (5) a "big" (or Skolem) clause, and a clause that results as in the translation of (6), a "small" clause, in order to emphasise the fact that the former contain n literals while the latter contains constant number of literals independent from n.

For a given FO sentence ψ, we denote its CNF propositional translation obtained as explained above by $\mathcal{C}_{\psi,n}$ where n is the size of the (finite) model. We also consider the (infinite) case $n = \omega$, which is an infinite (but countable) propositional CNF that has the same set of countable models as the FO sentence ψ, except for the Skolem relations $S_i(x_1, x_2, \ldots x_i, y_i)$ that are made explicit in $\mathcal{C}_{\psi,\omega}$.

Given a (propositional) variable of the form $R_i(c_1, c_2, \ldots c_p)$ or $S_j(c_1, c_2, \ldots c_p, x)$, we call $c_1, c_2, \ldots c_p$ *arguments* of R_i or S_j, respectively. We call x the *witness* of S_j. We also call $c_1, c_2, \ldots c_p$ and x the *elements* of R_i or S_j, respectively. Two propositional formulae, built upon R-variables and S-variables are isomorphic iff there is a bijection between the elements of the two that induces a bijection between the variables that in turn induces an isomorphisms between the formulae. Given a propositional formula φ, built upon R-variables and S-variables, we call *instances* of φ all formulae that are isomorphic to φ.

3. OUR RESULT

3.1 First-Order Contradictions have Constant Rank Lovasz-Schrijver refutations

The FO sentences that have no models, either finite or infinite, are universally false, so they have (finite) refutations in any sound and complete proof systems for FO logic. We shall first introduce such a proof system, which is in fact FO Resolution but presented in a tableau-style manner. We shall then show how to translate a proof of FO contradiction ψ into constant-rank LS proof of $\mathcal{C}_{\psi,n}$ ("constant" here and thereafter means being independent from the size of the finite model n).

The refutation of ψ is a decision tree $\mathcal{T}_\psi$ that tries to build a model of $\mathcal{C}_{\psi,\omega}$ as follows. It starts with the constants of ψ, $\{1, 2, \ldots m\}$ and witnesses new constants whenever necessary. Every internal node of $\mathcal{T}_\psi$ makes one of the following two kinds of queries:

1. A Boolean query of the form $R_i(c_1, c_2, \ldots c_p)$ where R_i is a p-ary predicate symbol from ψ, and $c_1, c_2, \ldots c_p$ are constants that have already been witnessed along the path from the root of $\mathcal{T}_\psi$ to the current node. The tree then branches on the two possible answers, $\bot$ and $\top$.

2. A Skolem query of the form $S_j(c_1, c_2, \ldots c_q, x)$ where S_j is a q-ary Skolem relation witnessing a variable x for some already existing constants $c_1, c_2, \ldots c_q$. There are finitely many possible answers to such a query - x is either one of the r constants witnessed along the path from the root to the current node, $\{1, 2, \ldots r\}$, or a new constant, which takes the next available "name" - $r + 1$.

Every node u of the tree $\mathcal{T}_\psi$ can naturally be labelled by the conjunction C_u of all answers to the queries made along the path from the root to u. A branch is closed, or equivalently its end-node v is a leaf of the tree, iff the conjunction C_v contradicts to one of the small clauses of $\mathcal{C}_{\psi,\omega}$.

The set of initial constants, which are known at the root of, is simply the set of constant symbols of ψ (whose "names" are $1, 2, \dots m$); if ψ does not contain any constant symbols at all. we instantiate a single constant and name it by 1. The order of variables in which the decision tree $\mathcal{T}_\psi$ makes queries is as follows. Given the set of constants $U = \{1, 2, \dots r\}$, known at a certain node u of the tree, any R-variable (with arguments within U) comes before any S-variable (with arguments within U). The order of $S_j(c_1, c_2, \dots c_q, x)$-variables is lexicographically-ascending on the tuples $\left(\sum_{k=1}^{q} c_k + j, \sum_{k=2}^{q} c_k + j, \dots, c_q + j, j\right)$; as a matter of fact any order that eventually lists every possible S-variable is adequate for our purpose.

In other words, when starting from u with the set of known constants $U = \{1, 2, \dots r\}$, the decision tree $\mathcal{T}_\psi$ first expands a subtree rooted at u that makes all Boolean R-queries with arguments in U. Any leaf of the subtree then picks the first S-variable that has not yet been queried, and branches on it. If the answer was within U, the next S-variable is picked up and queried and so on; if the answer was a new constant (whose name is now $r + 1$), the respective node expands a subtree that queries all R-variables with at least one argument set to $r + 1$. Any leaf of the subtree then picks the next unqueried S-variable and so on. Of course, one has to bear in mind that a branch is closed, i.e. the respective node becomes a leaf of the decision tree, as soon as the information gathered by the queries along the branch is a direct contradiction to one of the small clauses of $\mathcal{C}_{\psi,\omega}$.

It is not hard to see that the procedure described above is a sound and complete proof system (or rather refutation system) for FO logic.

Theorem 3. *The decision tree $\mathcal{T}_\psi$ is finite if and only if $\mathcal{C}_{\psi,\omega}$ is a propositional contradiction, which is equivalent to ψ being a FO contradiction.*

Proof. Indeed, by expanding the decision tree $\mathcal{T}_\psi$, one attempts to create all at most countable models of $\mathcal{C}_{\psi,\omega}$, both finite and infinite. If the tree is finite, i.e. all branches have been closed, it follows that ψ has no models, i.e. it is a FO contradiction.

It is not hard to see that an infinite branch is in fact an infinite model as it never violates a small clause and eventually satisfies any big clause by finding a witness of the infinite disjunction. Suppose now that ψ is a FO contradiction ($\mathcal{C}_{\psi,\omega}$ is a propositional contradiction) but the tree $\mathcal{T}_\psi$ is infinite. As the branching factor of every internal node is finite, by König's lemma, there must be an infinite branch which constitutes an infinite model of ψ - a contradiction. $\square$

Before we explain how to turn a finite decision tree tree $\mathcal{T}_\psi$ into constant rank LS refutation of $\mathcal{C}_{\psi,n}$, we need the following technical lemma.

Lemma 4. *The inequality $\sum_{j=1}^{d} p_{l_j} - 1 \geq 0$ has an LS derivation of rank at most d from the inequalities*

$$\begin{aligned} \sum_{j=1}^{d} \alpha_j p_{l_j} - \beta &\geq 0 \\ p_{l_j} + p_{\neg l_j} - 1 &= 0 \\ p_{l_j}, p_{\neg l_j} &\geq 0 \end{aligned}$$

where $\alpha_j > 0$ for every j, and $\beta > 0$.

The lemma follows trivially from the fact that the LS-rank of a valid inequality is bounded from above by the number of variables. We will however give a concrete derivation for the sake of completeness.

Proof. We shall prove by induction that the inequality

$$\beta \sum_{j=1}^{i} p_{l_j} + \sum_{j=i+1}^{d} \alpha_j p_{l_j} - \beta \geq 0 \tag{7}$$

has a rank i derivation. The basis case $i = 0$ is trivial. As for the inductive step, we multiply the inequality (7) by $p_{\neg l_{i+1}}$, and get

$$\beta \sum_{j=1}^{i} p_{l_j} p_{\neg l_{i+1}} + \sum_{j=i+2}^{d} \alpha_j p_{l_j} p_{\neg l_{i+1}} - \beta p_{\neg l_{i+1}} \geq 0. \tag{8}$$

For every $j \neq i+1$, we add the equation $p_{l_{i+1}} + p_{\neg l_{i+1}} - 1 = 0$ multiplied by $-p_{i_j}$ to the inequality $p_{l_{i+1}} \geq 0$ multiplied by p_{l_j} and we then multiply the result by either $-\beta$ if $j \leq i$ or by $-\alpha_j$ if $j \geq i + 2$ and add it to (8) in order to transform any term of the form $p_{i_j} p_{\neg l_{i+1}}$ into the term p_{l_j}. We finally multiply the equation $p_{l_{i+1}} + p_{\neg l_{i+1}} - 1 = 0$ by β and add it to the transformed inequality (8). The final result then is

$$\beta \sum_{j=1}^{i+1} p_{l_j} + \sum_{j=i+2}^{d} \alpha_j p_{l_j} - \beta \geq 0, \tag{9}$$

which completes the inductive step. In the end we multiply the final inequality (8) for $i = d$ by $\frac{1}{\beta}$ in order to get the desired result $\sum_{j=1}^{d} p_{l_j} - 1 \geq 0$. Each inductive step increased the rank by at most 1, so the total rank of the derivation is at most d as claimed. $\square$

We are now ready to state and prove our main lemma in the "easy" case.

Lemma 5. *Whenever a node u in the tree $\mathcal{T}_\psi$ is labelled by a conjunction $\wedge_{j=1}^{d} l_j$, there are LS derivations of all instances of the inequality $\sum_{j=1}^{d} p_{\neg l_j} - 1 \geq 0$ of rank at most $h_u h$ where h is the height of $\mathcal{T}_\psi$ and h_u is the height of the subtree rooted at u.*

Proof. We shall proceed by induction on h_u.

The basis case $h_u = 0$ is easy: u is a leaf of the tree, so it is labelled by a direct contradiction to a small clause. More formally, the information gathered along the path from the root to u is a conjunction of the form $\bigwedge_{i \in C} \neg l_i \wedge \bigwedge_{j \in D} l_j$ where the disjunction $\bigvee_{i \in C} l_i$ is a small clause from $\mathcal{C}_{\psi,n}$. Recall now that the LS encoding of that clause is $\sum_{i \in C} p_{l_i} - 1 \geq 0$, which when added to the LS axioms $p_{\neg l_j} \geq 0$ for all $j \in D$ gives the desired result $\sum_{i \in C} p_{l_i} + \sum_{j \in D} p_{\neg l_j} - 1 \geq 0$, and note that this derivation is of rank 0.

As for the inductive step in case $h_u > 0$, we need to consider the type of query, which the internal node u makes.

Let us first denote the conjunction label of u by $\bigwedge_{i\in C} l_i$, where each l_i is a literal built upon either R-variable or S-variable.

1. The query at u is a Boolean one, i.e. of the form $R_i(c_1, c_2, \ldots c_p)$ (we shorten this notation to $R_i(\bar{c})$). The two successors of u are then labelled by $\bigwedge_{i\in C} l_i \wedge R_i(\bar{c})$ and $\bigwedge_{i\in C} l_i \wedge \neg R_i(\bar{c})$, respectively which, by the inductive hypothesis, implies that both
$$\sum_{i\in C} p_{\neg l_i} + p_{\neg R_i(\bar{c})} - 1 \geq 0$$
and
$$\sum_{i\in C} p_{\neg l_i} + p_{R_i(\bar{c})} - 1 \geq 0,$$
have LS derivations of rank at most $h(h_u - 1)$. Adding these two plus the LS axiom $p_{R_i(\bar{c})} + p_{\neg R_i(\bar{c})} - 1 = 0$ multiplied by -1 yields
$$2\sum_{i\in C} p_{\neg l_i} - 1 \geq 0.$$
An application of Lemma 4 with $\alpha_i = 2$, $\beta = 1$, and $d = |C| \leq h$ gives the desired inequality with an LS derivation of rank at most h.

2. The query at u is a Skolem one, i.e. of the form $S_j(c_1, c_2, \ldots c_q, x)$ (we shorten this to $S_j(\bar{c}, x)$). Denoting the set of constants, known at the node u, by $U = \{1, 2, \ldots r\}$, there are $r+1$ successors of u in the tree. We shall consider two sub-cases:

   (a) x is a constant already known, i.e. $x \in U$. By the inductive hypothesis the inequalities
   $$\sum_{i\in C} p_{\neg l_i} + p_{\neg S_j(\bar{c},x)} - 1 \geq 0 \quad \text{for} \quad x \in U$$
   have LS derivations of rank at most $h(h_u - 1)$.

   (b) x is a new constant, i.e. $x = r+1$. As the set of known constants U is contiguous at any node of the decision tree, i.e. $x \notin U$ is equivalent to $x \notin \mathrm{Elms}\left(\bigwedge_{i\in C} l_i\right)$, i.e. by the inductive hypothesis, we can derive all instances of $\bigwedge_{i\in C} l_i \wedge S_j(\bar{c}, x)$ where $x \notin \mathrm{Elms}\left(\bigwedge_{i\in C} l_i\right)$. (Here $\mathrm{Elms}(C)$ denotes the set of all elements mentioned by the conjunct C.) Thus we can derive in LS the inequalities
   $$\sum_{i\in C} p_{\neg l_i} + p_{\neg S_j(\bar{c},x)} - 1 \geq 0 \quad \text{for} \quad x \notin U$$
   by derivations of rank at most $h(h_u - 1)$.

   Adding together the inequalities obtained in the two cases yield
   $$n\sum\nolimits_{i\in C} p_{\neg l_i} + \sum_{x\in[n]} p_{\neg S_j(\bar{c},x)} - n \quad \geq \quad 0.$$
   We now add the inequality above to the big clause $\sum_{i\in[n]} p_{S_j(\bar{c},x)} - 1 \geq 0$ together with the LS axioms $p_{S_j(\bar{c},x)} + p_{\neg S_j(\bar{c},x)} - 1 = 0$ multiplied by -1 for every $x \in [n]$ in order to get
   $$n\sum\nolimits_{i\in C} p_{\neg l_i} - 1 \quad \geq \quad 0.$$
   Finally, an application of Lemma 4 with $\alpha_i = n$, $\beta = 1$, and $d = |C| \leq h$ gives the desired inequality with an LS derivation of rank at most h.

□

In the end, we can derive and state the main theorem as an easy consequence of Lemma 5. Indeed, applying the lemma to the root of the decision tree $\mathcal{T}_\psi$, we realise that there is an LS derivation of the inequality $-1 \geq 0$ of rank at most h^2, thus proving the following:

THEOREM 6. *Given a First-Order contradiction ψ, its standard translation into propositional Conjunctive Normal Form over a finite universe of size n $\mathcal{C}_{\psi,n}$ has an LS refutation of constant rank that depends on the formula ψ but does not depend on n.*

3.2 Infinite model implies non-constant Sherali-Adams rank

We shall prove prove that if a FO sentence ψ has an infinite model, the rank of the SA refutation of its propositional translation $\mathcal{C}_{\psi,n}$ grows with n. As an SA refutation is simply an inconsistent linear program, we shall show that for every fixed k there is a big enough $n_0 = n(k)$ such that for every $n \geq n_0$ the rank k SA linear program for $\mathcal{C}_{\psi,n}$ is consistent. This can be done by establishing a specific valuation of the variables of the SA system via counting (or probabilistic) argument over all finite segments of any class of infinite models of ψ (or more precisely, $\mathcal{C}_{\psi,\omega}$). If we consider the class of all countable models of $\mathcal{C}_{\psi,\omega}$, the lower bound on the SA rank k as a function of the size of the model n is poly-logarithmic, i.e. $\Omega\left((\log n)^\alpha\right)$ for some constant α, $0 < \alpha \leq 1$, that depends only on the FO sentence ψ.

We start by recalling the structure of the rank k SA system (linear program) for a FO sentence ψ, which we denote by $\mathcal{S}_{\psi,k,n}$. It is built upon real variables of the form
$$p_{\bigwedge_{j\in C} l_j}$$
where the index $\bigwedge_{j\in C} l_j$ is a conjunction of no more than k literals l_j, each of which is made up either of an R-variable or an S-variable. $\mathcal{S}_{\psi,k,n}$ consists of the following equations and inequalities:
$$p_\top = 1, \tag{10}$$
which takes care of the empty conjunct $\top$;
$$p_{\bigwedge_{j\in C} l_j \wedge l} + p_{\bigwedge_{j\in C} l_j \wedge \neg l} = p_{\bigwedge_{j\in C} l_j} \tag{11}$$
for every $|C| \leq k$ and every literal l whose variable is different from each of the variables of l_j;
$$p_{\bigwedge_{j\in C} l_j} \geq 0 \tag{12}$$
for every $|C| \leq k$;
$$\sum_{i\in D} p_{\bigwedge_{j\in C} l_j \wedge l_i} \geq p_{\bigwedge_{j\in C} l_j} \tag{13}$$
for every $|C| \leq k$ and every small clause $\bigvee_{i\in D} l_i$ in $\mathcal{C}_{\psi.n}$, and
$$\sum_{x\in[n]} p_{\bigwedge_{j\in C} l_j \wedge S_i(\bar{c},x)} \geq p_{\bigwedge_{j\in C} l_j} \tag{14}$$

for every $|C| \leq k$, every Skolem relation S_i in $\mathcal{C}_{\psi.n}$, and every tuple $\bar{c}$.

Note that the equations (10) and (11) as well as the inequalities (12) do not depend on ψ. One should also bear in mind that in the LHSs of the inequalities (13) and (14), every term, whose index contains both a literal and its negation, simply vanishes.

We are now ready to state and prove the general SA rank lower bound lemma.

LEMMA 7. *For a given FO sentence ψ that has an infinite model, there is a constant a such that for every fixed k and every $n \geq 2^{k^a}$, the linear program $\mathcal{S}_{\psi,k,n}$ is consistent.*

PROOF. We consider a set $\mathfrak{M}$ of *labelled countable models* of $\mathcal{C}_{\psi,\omega}$. What this means is that we label the elements of a countable universe by the positive integers so that the m (distinct) constant symbols of ψ - $a_1, a_2, \ldots a_m$ are interpreted as the numbers $1, 2, \ldots m$, respectively. Note that if $\mathcal{M} \in \mathfrak{M}$ and $\mathcal{M}'$ is obtained from $\mathcal{M}$ by taking a permutation of (the labels) $\mathbb{N}_+$ that does not move $1, 2, \ldots$ and m, the models $\mathcal{M}$ and $\mathcal{M}'$ are distinct as labelled models even though they are isomorphic in the usual model-theoretic sense (as unlabelled models).

Given a labelled model $\mathcal{M}$ of $\mathcal{C}_{\psi,\omega}$ and a number $d \geq m$, we call the restriction of $\mathcal{M}$ to $[d]$ (the set of elements labelled by $1, 2, \ldots d$) an *initial segment* of $\mathcal{M}$ of size d, and denote it by $\mathcal{M}_d$ (note that $\mathcal{M}_d$ is not a model of $\mathcal{C}_{\psi,d}$ as ψ has no finite models). In other words, an initial segment of size d is any labelled (by $[d]$) finite structure of size d that contains the constants of ψ as the first m labels, and could be extended into a (countable) labelled model of $\mathcal{C}_{\psi,\omega}$.

Finally, we denote by $\mathfrak{M}_d\left(\bigwedge_{j \in C} l_j\right)$ the set of all initial segments of size d that are consistent with the conjunction $\bigwedge_{j \in C} l_j$, built upon R-variables and S-variables, whose arguments and witnesses are all in $[d]$. Clearly, the set $\mathfrak{M}_d\left(\bigwedge_{j \in C} l_j\right)$ is finite, and $\mathfrak{M}_d = \mathfrak{M}_d(\top)$ is the set of all labelled finite models of size d that could be extended to countable models of $\mathcal{C}_{\psi,\omega}$.

We can now define our valuation as follows. Given a fixed rank k, we set d to be $m + k \max\{p, q+1\}$, where p and q are the maximal arities of relation symbols in ψ and Skolem relations in $\mathcal{C}_{\psi,n}$, respectively. (Recall that we defined the arity of a Skolem relation to be the number of its arguments, thus excluding the witness, hence $q+1$ in the expression above.) In the end, set the values of the variables as

$$p_{\bigwedge_{j \in C} l_j} = \frac{\left|\mathfrak{M}_d\left(\bigwedge_{j \in C} l_j\right)\right|}{|\mathfrak{M}_d|}$$

for every $|C| \leq k$. Note that, formally speaking, the above definition is insufficient as, in general, some of the elements mentioned by the conjunct in the numerator may be outside $[d]$. With a slight abuse of notation, though, we fix the problem by taking a conjunct that is isomorphic to $\bigwedge_{j \in C} l_j$ and whose elements are all within $[d]$ and leave all the numbers from $[m]$ in place (recall that those represent the constants of ψ). Thus any two variables indexed by isomorphic conjunctions will have the same value.

In other words, we set the values of the variables in a natural way - the variable $p_{\bigwedge_{j \in C} l_j}$ is meant to "represent" the conjunction $\bigwedge_{j \in C} l_j$, and it is a real variable in the interval $[0, 1]$, so it is natural that it is set to the fraction of initial segments that are consistent (satisfy) $\bigwedge_{j \in C} l_j$ or, if you prefer, the probability that an initial segment picked uniformly at random satisfies the conjunction. Note also that if $\bigwedge_{j \in C} l_j$ is inconsistent with $\mathcal{C}_{\psi,\omega}$, $p_{\bigwedge_{j \in C} l_j}$ is set to 0. What remains to be verified is that this valuation indeed satisfies the rank-k SA system for $\mathcal{C}_{\psi,n}$ for $n \geq 2^{k^a}$ for some suitable chosen constant a.

The (in)equalities (10) and (12) are trivially fulfilled as are the equations (11) - every initial segment consistent with $\bigwedge_{j \in C} l_j$ has either $l = \bot$ or $l = \top$ but not both. As for the small-clause inequalities (13) , it is enough to recall that every initial segment $\mathcal{M}_d$ is a substructure of a model of $\mathcal{C}_{\psi,\omega}$ and as such $\mathcal{M}_d$ satisfies every small clause $\bigvee_{i \in D} l_i$ whose elements are all within $[d]$, so we have

$$\mathfrak{M}_d\left(\bigwedge_{j \in C} l_j\right) = \bigcup_{i \in D} \mathfrak{M}_d\left(\bigwedge_{j \in C} l_j \wedge l_i\right)$$

for all conjuncts $\bigwedge_{j \in C} l_j$ and therefore

$$\sum_{i \in D}\left|\mathfrak{M}_d\left(\bigwedge_{j \in C} l_j \wedge l_i\right)\right| \geq \left|\mathfrak{M}_d\left(\bigwedge_{j \in C} l_j\right)\right|.$$

The only non-trivial case which, in fact, gives the lower bound, is the case of big clauses (14). Consider such an inequality of the form

$$\sum_{x \in [n]} p_{\bigwedge_{j \in C} l_j \wedge S_i(\bar{c}, x)} \geq p_{\bigwedge_{j \in C} l_j} \tag{15}$$

for some fixed conjunction $\bigwedge_{j \in C} l_j$. We need to consider two cases depending on the relationship between the witness x and the set E of all other elements mentioned by the conjunction $\bigwedge_{j \in C} l_j \wedge S_i(c_1, c_2, \ldots c_q, x)$.

1. The witness x cannot be outside the set E, i.e. $\bigwedge_{j \in C} l_j \wedge S_i(c_1, c_2, \ldots c_q, x)$ is not consistent with any model of $\mathcal{C}_{\psi,\omega}$ as long as $x \notin E$. In such a case we have

$$\mathfrak{M}_d\left(\bigwedge_{j \in C} l_j\right) = \bigcup_{x \in E} \mathfrak{M}_d\left(\bigwedge_{j \in C} l_j \wedge S_i(\bar{c}, x)\right),$$

which implies

$$\sum_{x \in E}\left|\mathfrak{M}_d\left(\bigwedge_{j \in C} l_j \wedge S_i(\bar{c}, x)\right)\right| \geq \left|\mathfrak{M}_d\left(\bigwedge_{j \in C} l_j\right)\right|,$$

i.e. the inequality (15) is trivially satisfied under our valuation.

2. The witness x can lay outside the set E, i.e. $\bigwedge_{j \in C} l_j \wedge S_i(c_1, c_2, \ldots c_q, x)$ is consistent with at least one model of $\mathcal{C}_{\psi,\omega}$ such that $x \notin E$. As we have already explained. all variables whose indices are isomorphic get the same value under our valuation, so we can break the LHS of the inequality (15) into two parts:

$$\sum_{x \in E} p_{\bigwedge_{j \in C} l_j \wedge S_i(\bar{c}, x)} + (n - |E|)\, p_{\bigwedge_{j \in C} l_j \wedge S_i(\bar{c}, x_0)} \geq p_{\bigwedge_{j \in C} l_j}$$

where x_0 is some number outside E, i.e. $x_0 \in [d] \setminus E$. Now recall that the values of p-variables do not depend on n, and neither does $|E|$. Therefore, if we choose n big enough (note that the other factor in the second term of the LHS is strictly positive as there is at least one model of $\mathcal{C}_{\psi,\omega}$ consistent with $\bigwedge_{j\in C} l_j \wedge S_i(\bar{c}, x_0)$), say

$$n \quad \geq \quad |E| + \frac{p_{\bigwedge_{j\in C} l_j}}{p_{\bigwedge_{j\in C} l_j \wedge S_i(\bar{c},x_0)}},$$

the inequality (15) will certainly be satisfied. The ratio of the two p-variables is trivially bounded by $|\mathfrak{M}_d|$. Given that $d = \Theta(k)$ (the hidden constant depends on ψ only) and there are finitely many relation symbols and (implicit) Skolem relations in ψ, each of them of finite arity, we conclude that $|\mathfrak{M}_d| \leq 2^{k^a}$ for some suitable chosen a that depends on the FO sentence ψ only.

□

We can finally state the SA rank lower bound theorem for FO sentences, which is a trivial consequence of Lemma 7.

THEOREM 8. *Given a First-Order sentence ψ that fails in the finite but has an infinite model, there is a constant a, $0 < a \leq 1$, such that every Sherali-Adams refutation of the propositional translation of ψ, $\mathcal{C}_{\psi,n}$, requires a rank at least $(\log n)^{\alpha}$, where n is the size of a finite model.*

4. CONCLUSION AND OPEN PROBLEMS

We have proven a rank gap for the proofs of certain uniformly generated propositional tautologies. The gap is between constant LS and poly-log SA. Our result leaves a number of open questions:

1. Can the gap be widened, i.e. can we replace poly-log SA by linear SA? Our guess is "yes", and it is supported by concrete examples, such as the Pigeon-Hole Principle and the Least Number Principle (see the lower bounds in [9]). As a matter of fact, Martin [7] has already improved the poly-log to poly, i.e. to n^{α} for some α, $0 < \alpha \leq 1$.

2. Can a gap be proven for LS_+, Lovász-Schrijver with semidefinite lifts? Note that such gap cannot be bigger that logarithmic since the Pigeon-Hole principle has a logarithmic rank LS_+ proof.

3. Can a gap be proven in the general case of Integer Linear Programs, not only for those that are translations of propositional contradictions? This may need some subtle definition of what it means for an infinite dimensional polytope to contain an integral point.

5. ACKNOWLEDGMENTS

I would like to thank Barnaby Martin for greatly helping me improve the presentation, both mathematically and linguistically. I am also grateful to the anonymous reviewers whose comments I found really helpful.

This work was partially supported by the EPSRC under grant EP/C526120/1.

6. REFERENCES

[1] J. Buresh-Oppenheim, N. Galesi, S. Hoory, A. Magen, and T. Pitassi. Rank bounds and integrality gaps for cutting planes procedures. In *44th IEEE Symposium on the Foundations of Computer Science*, pages 318–327, October 2003.

[2] W. Cook, R. Coullard, and G. Turan. On the complexity of cutting plane proofs. *Discrete Applied Mathematics*, 18:25–38, 1987.

[3] R. Gomory. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the AMS*, (64):275–278, 1958.

[4] D. Grigoriev, E. A. Hirsch, and D. V. Pasechnik. Complexity of semialgebraic proofs. *Moscow Mathematical Journal*, 2(4):647–679, 2002.

[5] M. Laurent. A comparison of the Sherali-Adams, lovász-Schrijver and Lasserre relaxations for $0-1$ programming. *Mathematics of Operations Research*, 28(3):470–496, 2003.

[6] L. Lovasz and A. Schrijver. Cones of matrices and set-functions and $0-1$ optimization. *SIAM Journal on Optimization*, 7(1):166–190, 1991.

[7] B. Martin. Personal communication.

[8] P. Pudlak. On the complexity of propositional calculus. In *Sets and Proofs*, pages 197–218. Cambridge University Press, 1999. Invited papers from Logic Colloquium'97.

[9] M. Rhodes. Rank lower bounds for the Sherali-Adams operator. To appear in *Computation and Logic in the Real World - Third Conference of Computability in Europe, CiE 2007*, LNCS, vol. 4497, 2007.

[10] S. Riis. A complexity gap for tree-resolution. *Computational Complexity*, 10:179–209, 2001.

[11] H. D. Sherali and W. P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3:411–430, 1990.