RANK COMPLEXITY GAP FOR LOVÁSZ-SCHRIJVER AND SHERALI-ADAMS PROOF SYSTEMS

Stefan Dantchev and Barnaby Martin

Abstract. We prove a dichotomy theorem for the rank of propositional contradictions, uniformly generated from first-order sentences, in both the Lovász-Schrijver (LS) and Sherali-Adams (SA) refutation systems. More precisely, we first show that the propositional translations of first-order formulae that are universally false, i.e. fail in all finite and infinite models, have LS proofs whose rank is constant, independent of the size of the (finite) universe. In contrast to that, we prove that the propositional formulae that fail in all finite models, but hold in some infinite structure, require proofs whose SA rank grows polynomially with the size of the universe.

Until now, this kind of so-called "complexity gap" theorem has been known for tree-like Resolution and, in somehow restricted forms, for the Resolution and Nullstellensatz 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 refutation system (since the conference version of this paper, SA has been considered as a refutation system in several further papers). An interesting feature of the SA system is that it simulates LS, the Lovász-Schrijver refutation system without semidefinite cuts, in a rank-preserving fashion.

Keywords. Propositional proof complexity, Lift-and-project methods, Lovász-Schrijver proof system, Lower bounds, Complexity gap theorems

Subject classification. 68Q25: Analysis of algorithms and problem complexity.

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 of a set of clauses to a set of inequalities, when applied to propositional contradictions, gives rise to some very interesting propositional refutation systems based on different methods for solving Integer Linear Programming.

By tradition, the two most important ILP-based refutation systems are Cutting Planes, introduced as a general method for solving ILP in Gomory (1958), and as a refutation system in Cook *et al.* (1987), and Lovász-Schrijver (LS), introduced as a general method for solving ILP in Lovász & Schrijver (1991), and first considered as a refutation system in Pudlák (1999). A number of mixtures of ILP-based refutation systems and algebraic refutation systems are introduced and studied in Grigoriev *et al.* (2002).

Another method for solving ILP was proposed by Sherali and Adams in Sherali & Adams (1990), but was not explored as a propositional refutation system until the conference version of this paper, Dantchev (2007). Since then, SA has been considered explicitly as a refutation system in Dantchev *et al.* (2009). Furthermore, various results about integrality gaps have yielded lower bounds for SA as a refutation system, e.g. those given in Schoenebeck (2008). The SA relaxation is interesting in that it is a static and stronger version of LS, the Lovász-Schrijver relaxation without semidefinite cuts. More precisely, it is proved in Laurent (2003) that the rank k SA relaxation is tighter than the rank k LS relaxation. The fact that SA is stronger than LS as a refutation system follows from Rhodes (2008).

A number of lower bounds have been proven for ILP-based refutation 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 contradictions from Buresh-Oppenheim *et al.* (2006) as well as the LS rank lower bound for the Pigeonhole Principle from Grigoriev *et al.* (2002). No size lower bounds are known for LS (other than for its tree-like restriction – see Beame *et al.* (2007) and Pitassi & Segerlind (2009)), and it seems that the rank is the better complexity measure for LS in the same way that the degree is a good complexity measure for the algebraic refutation systems. More recently, and since the conference version of this paper, attention has focussed on integrality gaps for SA and its semidefinite variant, the Lasserre relaxation (see, for example, Charikar *et al.* (2009); Georgiou *et al.* (2009); Mathieu & Sinclair (2009); Schoenebeck (2008)). The Lasserre relaxations are tighter than each of LS, LS_+ and SA. Among a number of interesting results in Schoenebeck (2008), it is proved that the $\Omega(n)$th rank of Lasserre can not prove a random k-SAT formula unsatisfiable.

All results in Buresh-Oppenheim *et al.* (2006) and Grigoriev *et al.* (2002) are lower bounds for specific contradictions. The aim of this paper is to prove a very general LS rank lower bound that would apply to a large class of contradictions, namely those that can be expressed as FO sentences. Note that the Pigeonhole principle as well as the Least number principle (stating that a finite order has no minimal element) are such contradictions, 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 proved in Riis (2001):

THEOREM 1.1. *Given an 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 ε such that for every n, every tree-like Resolution refutation of $\mathcal{C}_{\psi,n}$ is of size at least $2^{\varepsilon n}$.

Furthermore, 2 holds if and only if ψ has an infinite model.

Since Riis (2001), various complexity gap, or classification, theorems have appeared: for a restricted class of contradictions for Resolution in Dantchev & Riis (2003), for Nullstellensatz in Riis (2008), for Cutting Planes in Dantchev & Martin (2009) and for a parameterised variant of tree-like Resolution in Dantchev *et al.* (2007). In its strongest form, our result can be stated as follows.

THEOREM 1.2. *Given an 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 ε such that for every n, every SA refutation of $\mathcal{C}_{\psi,n}$ is of rank $\Omega(n^{\varepsilon})$.

Furthermore, 2 holds if and only if ψ has an infinite model.

In light of the rank-preserving simulation of LS by SA, this provides matching gap theorems for LS and SA.

The rest of the paper is organised as follows. In Section 2, we define the two refutation systems LS and SA, and explain the translation of an FO sentence into a family of finite propositional contradictions. The main part of the paper, Section 3, contains the proof of Theorem 1.2. It is divided into two – we first prove the "easy", constant LS rank, case in Section 3.1 and then move onto the "hard" non-constant lower bound for the SA rank in Section 3.2. We finally discuss some open questions.

2. Preliminaries

The Lovász-Schrijver (LS) refutation system. Lovász-Schrijver is a lift-and-project refutation 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 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

$$1 \geq p_v \geq 0 \text{ and } 1 \geq p_{\neg v} \geq 0.$$

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 inequality by a variable p_l, where l is a literal, in order to get a quadratic inequality. If the original inequality contained a term $p_{\neg l}$, the new quadratic term $p_l p_{\neg l}$ vanishes, i.e. does not appear in the result. If the original inequality contained a term p_l, the new quadratic term p_l^2 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 inequalities.

An LS derivation of an inequality from a set of inequalities (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 inequality that can be derived in a single step, using Rules 1, 2 or 3, from the inequalities that label the children of the node. The root of the tree is labelled by the inequality that is finally derived. The *rank* of an LS derivation is the maximal number of derivation steps, using Rule 1, i.e. multiplications of a linear inequality by a variable, over all branches (paths from the root to a leaf) of the derivation tree. The rank of an inequality with respect to a set of axioms is the minimal rank over all possible derivations of the inequality 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) refutation system. Sherali-Adams is a static refutation system, so we shall define SA refutations 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 $k+1$ variable-distinct literals l_i, $|I| \leq k+1$ (we shall write $|C| \leq k+1$ 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

$$1 \geq p_C \geq 0 \tag{2.1}$$

as well as the equations

$$p_{C \wedge v} + p_{C \wedge \neg v} = p_C \tag{2.2}$$

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

$$p_{\emptyset} = 1 \tag{2.3}$$

where $\emptyset$ is the empty conjunct (of size 0, i.e. $\emptyset = \top$). 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

$$\sum_{j \in J} p_{l_j \wedge C} \geq p_C \tag{2.4}$$

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 (alternatively, one may keep such variables and see that they must evaluate to zero due to (2.2)).

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 (2.2), (2.3) and inequalities (2.1), (2.4), 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 polynomial-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 in a rank-preserving fashion. The following proposition may be inferred from Laurent (2003), but we give its (easy) proof for completeness.

PROPOSITION 2.5. *An inequality that can be derived in LS rank k is consequent of the inequalities of SA rank k.*

PROOF. By induction on the rank k. The case $k = 0$ (the axioms) is trivial. Suppose that it is true for rank k. An inequality that can be derived in LS rank $k+1$ is a positive linear combination of inequalities of the form $\sum_{j\in J} \alpha_j p_{l_j} p_l \geq \beta p_l$, for real numbers α_j, β, where $\sum_{j\in J} p_{l_j} \alpha_j \geq \beta$ is derivable in LS rank k. But it is now clear, since $\sum_{j\in J} p_{l_j} \alpha_j \geq \beta$ was consequent on SA rank k, that $\sum_{j\in J} \alpha_j p_{l_j \wedge l} \geq \beta p_l$ ($\sum_{j\in J} \alpha_j p_{l_j} p_l \geq \beta p_l$) is consequent on SA rank $k+1$. □

Translation of FO sentences into propositional CNF formulae. We use the language of FO logic with equality but without function or constant symbols, i.e. we only allow relation symbols. The omission of constants is purely for technical simplicity (note that constants may be simulated by outermost added existential quantifiers). 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-variables* $R\left(x_{i_1}, x_{i_2}, \ldots x_{i_p}\right)$, 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 R-variables only.

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 $1 \leq i \leq k$, which give rise to *S-variables.* $S_i(x_1, x_2, \ldots x_i, y_i)$ 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.

$$\bigvee_{y_i=1}^{n} S_i(x_1, x_2, \ldots x_i, y_i) \tag{2.6}$$

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

$$\forall \overline{x}, \overline{y} \; \bigwedge_{i=1}^{k} S_i(x_1, \ldots x_i, y_i) \to \mathcal{F}(\overline{x}, \overline{y}). \tag{2.7}$$

We shall call clause (2.6) a "big" (or Skolem) clause, and a clause that results as in the translation of (2.7), 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 $S_i(x_1, x_2, \dots x_i, y_i)$ that are made explicit in $\mathcal{C}_{\psi,\omega}$. It is easy to see that $\mathcal{C}_{\psi,n}$ is satisfiable iff ψ has a model of size n.

Given a (propositional) variable of the form $R_i(c_1, c_2, \dots c_p)$ or $S_j(c_1, c_2, \dots c_p, x)$, we call $c_1, c_2, \dots c_p$ *arguments* of R_i or S_j, respectively. We call x the *witness* of S_j. We also call $c_1, c_2, \dots c_p$ 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 Lovász-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 refutation system for FO logic. We shall first introduce such a refutation system, which is in fact FO Resolution but presented in a tableau-style manner. We shall then show how to translate a refutation of an FO contradiction ψ into a constant-rank LS refutation of $\mathcal{C}_{\psi,n}$ ("constant" here and hereafter 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 witnessing some unary Skolem relation in ψ with the constant 1 and deriving further constants as Skolem witnesses of already derived constants as and when necessary. (Note that we tend to discount the empty model. It is, therefore, possible to have ψ with no finite models and no outermost existential quantifier. In this case we may instantiate a single constant at the outset to get us going.) Every internal node of $\mathcal{T}_\psi$ makes one of the following two kinds of queries:

1. A Boolean query of the form $R_i(c_1, c_2, \dots c_p)$ where R_i is a p-ary predicate

symbol from ψ, and $c_1, c_2, \dots c_p$ 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 $S_j(c_1, c_2, \dots c_q, x)$ where S_j is a q-ary Skolem relation witnessing a variable x for some already existing constants c_1, c_2, $\dots, c_q$. 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, $\{1, 2, \dots r\}$, 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 one of the small clauses of $\mathcal{C}_{\psi,\omega}$.

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 – in 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 refutation system for FO logic.

THEOREM 3.1. *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 an 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 an 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 for the infinite disjunction. Suppose now that ψ is an 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. $\square$

Example 3.2. We give an example of a decision tree $\mathcal{T}_\psi$ constructed as in the procedure just given. We consider the following sentence ψ which has no models:

$$\forall x \exists y\ R(x,y) \quad \wedge \quad \exists x \forall y\ \neg R(x,y).$$

As per our translation to propositional clauses, this is equivalent to the conjunction of the universal clauses

$$\begin{array}{ll} (i.) & \forall x \forall y\ \neg S_2(x,y) \vee R(x,y) \text{ and} \\ (ii.) & \forall x \forall y\ \neg S_1(x) \vee \neg R(x,y), \end{array}$$

together with the Skolem clauses

$$\begin{array}{c} \forall x \exists y\ S_2(x,y) \text{ and} \\ \exists x\ S_1(x). \end{array}$$

Figure 3.1 shows an FO decision tree for this system of clauses. The number following each # specifies the clause that has been contradicted. For example, the bottom right # comes from the knowledge $S_2(1,2)$ and $\neg R(1,2)$ – which contradicts the first universal clause. $\Diamond$

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

Lemma 3.3. *The inequality $\sum_{j=1}^{d} p_{l_j} - 1 \geq 0$ has an LS derivation of rank at most d from the inequalities*

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

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

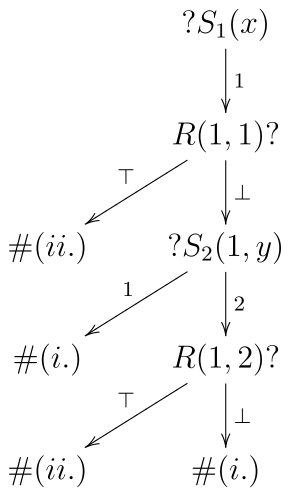

Figure 3.1: Decision tree for Example 3.2.

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 on i that the inequality

$$\beta \sum_{j=1}^{i} p_{l_j} + \sum_{j=i+1}^{d} \alpha_j p_{l_j} - \beta \geq 0 \tag{3.4}$$

has a rank i derivation. The basis case $i = 0$ is trivial. As for the inductive step, we multiply the inequality (3.4) by $p_{\neg l_{i+1}}$, and get

$$\beta \sum_{j=1}^{i} p_{l_j} p_{\neg l_{i+1}} + \sum_{j=i+2}^{d} \alpha_j p_{l_j} p_{\neg l_{i+1}} - \beta p_{\neg l_{i+1}} \geq 0. \tag{3.5}$$

For every $j \neq i+1$, we add the equation $p_{l_{i+1}} + p_{\neg l_{i+1}} - 1 = 0$ multiplied by $-p_{l_j}$ 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 (3.5) in order to transform any term of the form $p_{l_j} p_{\neg l_{i+1}}$ 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

(3.5). The final result then is

$$\beta \sum_{j=1}^{i+1} p_{l_j} + \sum_{j=i+2}^{d} \alpha_j p_{l_j} - \beta \geq 0, \tag{3.6}$$

which completes the inductive step. In the end we multiply the final inequality (3.5) 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. □

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

Lemma 3.7. *Whenever a node u in the tree $\mathcal{T}_\psi$ is labelled by a conjunction $\wedge_{j=1}^{d} l_j$, there is an LS derivation 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 an R-variable or an 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 3.3 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

$$n \sum_{i \in C} p_{\neg l_i} + \sum_{x \in [n]} p_{\neg S_j(\bar{c}, x)} - n \quad \geq \quad 0.$$

We now add the inequality above to the big clause $\sum_{x \in [n]} p_{S_j(\bar{c}, x)} - 1 \geq 0$ 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

$$n \sum_{i \in C} p_{\neg l_i} - 1 \quad \geq \quad 0.$$

Finally, an application of Lemma 3.3 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 3.7. 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 3.8. *Given an FO contradiction ψ, its standard translation into propositional CNF 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 that if an 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 a 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 polynomial, i.e. $\Omega(n^\varepsilon)$ for some constant ε, $0 < \alpha \leq 1$, that depends only on the FO sentence ψ.

We start by recalling the structure of the rank k SA system (linear program) for an FO sentence ψ, which we denote by $\mathcal{L}_{\psi,k,n}$. 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 $k+1$ literals l_j, each of which is made up either of an R-variable or an S-variable. $\mathcal{L}_{\psi,k,n}$ consists of the following equations and inequalities:

$$p_\emptyset = 1, \tag{3.9}$$

which takes care of the empty conjunct $\top$;

$$p_{\bigwedge_{j\in C} l_j \wedge l} + p_{\bigwedge_{j\in C} l_j \wedge \neg l} = p_{\bigwedge_{j\in C} l_j} \tag{3.10}$$

for every $|C| \leq k$ and every literal l whose variable is different from each of the variables of l_j;

$$1 \geq p_{\bigwedge_{j \in C} l_j} \geq 0 \tag{3.11}$$

for every $|C| \leq k+1$;

$$\sum_{i \in D} p_{\bigwedge_{j \in C} l_j \wedge l_i} \geq p_{\bigwedge_{j \in C} l_j} \tag{3.12}$$

for every $|C| \leq k$ and every small clause $\bigvee_{i \in D} l_i$ in $\mathcal{C}_{\psi,n}$, and

$$\sum_{x \in [n]} p_{\bigwedge_{j \in C} l_j \wedge S_i(\bar{c},x)} \geq p_{\bigwedge_{j \in C} l_j} \tag{3.13}$$

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

Note that the equations (3.9) and (3.10) as well as the inequalities (3.11) do not depend on ψ. One should also bear in mind that in the LHSs of the inequalities (3.12) and (3.13), 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 3.14. *For a given FO sentence ψ that has an infinite model, there are constants α, β such that for every fixed $k > \beta$ and every $n \geq k^{\alpha}$, the linear program $\mathcal{L}_{\psi,k,n}$ 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. Note that if $\mathcal{M} \in \mathfrak{M}$ and $\mathcal{M}'$ is obtained from $\mathcal{M}$ by taking a permutation of the positive integers (the labels), 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, 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 could be extended into a (countable) labelled model of $\mathcal{C}_{\psi,\omega}$. Call a labelled model $\mathcal{M}$ of $\mathcal{C}_{\psi,\omega}$ *amenable* if all S-variables with arguments in $[d]$ have witnesses in $[d+e]$. Let $\mathcal{M}_{d|e}$ be the restriction of an amenable model $\mathcal{M}$ of $\mathcal{C}_{\psi,\omega}$ to $[d+e]$ such that no S-variable witnesses are given when not all of their arguments are contained in $[d]$. Thus a model

$\mathcal{M}_{d|e}$, when extended to an amenable model $\mathcal{M}$, may actually have S-variable witnesses to $[d+e]$ (when not all arguments are in $[d]$) in $[d+e]$ or elsewhere – the point is that we will not record them if they are in $[d+e]$. We will never care about these witnesses and, since it will be confusing to have them, we explicitly ignore them. We denote by $\mathfrak{M}_{d|e}$, the set of $\mathcal{M}_{d|e}$ so derived from amenable models $\mathcal{M}$ of $\mathcal{C}_{\psi,\omega}$. In general, if e is not sufficiently large as a function of d, then $\mathfrak{M}_{d|e}$ may be empty. Finally, we denote by $\mathfrak{M}_{d|e}\left(\bigwedge_{j\in C} l_j\right)$ the class of models in $\mathfrak{M}_{d|e}$ that are consistent with the conjunction $\bigwedge_{j\in C} l_j$, built upon R-variables and S-variables, whose elements are from $[d+e]$ but whose S-variable arguments are all from $[d]$ (so we allow $\{d+1,\ldots,d+e\}$ to appear as arguments in R-variables).

Clearly, the set $\mathfrak{M}_{d|e}\left(\bigwedge_{j\in C} l_j\right)$ is finite, and $\mathfrak{M}_{d|e} = \mathfrak{M}_{d|e}(\top)$ is the set of all labelled finite models of size $d+e$, whose S-variable arguments in $[d]$ are witnessed in $[d+e]$, 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 $d := (k+1)\cdot\max\{p,q\}$, where p and q are the maximal arities of relation symbols in ψ and Skolem relations in $\mathcal{C}_{\psi,n}$, respectively, and $e := b\cdot d^q$, where b is the number of Skolem relations in $\mathcal{C}_{\psi,n}$. Note that e is sufficiently large to accommodate all potential S-variable witnesses to the set $[d]$, this is important as it will always force $\mathfrak{M}_{d|e}$ to be non-empty. Let $\bigwedge_{j\in C} l_j$ contain R-variables whose arguments are in $[d+e]$ and S-variables whose arguments are in $[d]$ and whose witnesses are in $[d+e]$. Set the values of our variables as

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

for every $|C| \le k+1$. If $\bigwedge_{j\in C} l_j$ contains R-variables whose arguments are outside of $[d+e]$ or S-variables whose arguments are outside of $[d]$ or whose witnesses are outside of $[d+e]$, then we fix the problem by taking a conjunct that is isomorphic to $\bigwedge_{j\in C} l_j$ and whose R-variable arguments are all within $[d+e]$ and whose S-variable arguments are within $[d]$ and whose witnesses are in $[d+e]$. 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, to 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 $n \geq k^\alpha$ for some suitable chosen constant α.

The inequalities (3.9) and (3.11) are trivially fulfilled as are the equations (3.10) – 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 (3.12) , it is enough to recall that every initial segment $\mathcal{M}_{d|e}$ is a substructure of a model of $\mathcal{C}_{\psi,\omega}$ and as such $\mathcal{M}_{d|e}$ satisfies every small clause $\bigvee_{i\in D} l_i$ whose R-variable arguments are in $[d+e]$ and whose S-variable arguments are within $[d]$ and whose witnesses are in $[d+e]$, so we have

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

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

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

The only remaining case which, in fact, gives the lower bound, is the case of big clauses (3.13). Consider such an inequality of the form

$$\sum_{x\in[n]} p_{\bigwedge_{j\in C} l_j \wedge S_i(\bar{c},x)} \geq p_{\bigwedge_{j\in C} l_j} \tag{3.15}$$

for some fixed conjunction $\bigwedge_{j\in C} l_j$, whose R-variable arguments are in $[d+e]$ and whose S-variable arguments are within $[d]$ and whose witnesses are in $[d+e]$. By the definition of our valuation

$$p_{\bigwedge_{j\in C} l_j \wedge S_i(\bar{c},1)} + \ldots + p_{\bigwedge_{j\in C} l_j \wedge S_i(\bar{c},d+e)} = p_{\bigwedge_{j\in C} l_j},$$

because we insist that the witnesses in our initial segments exist. Let $z := \max\{p,q,b\}$ and let $k \geq (z+1)z^z + 1 =: \beta$; the result proceeds in the following fashion.

$$\begin{aligned} n = d+e &:= (k+1)\max\{p,q\} + b^q((k+1)\max\{p,q\})^q \\ &\leq (k+1)^{q+3} && \text{as } k \geq \max\{p,q,b\}^q \\ &\leq k^{q+4} && \text{as } k \geq (z+1)z^z + 1 \end{aligned}$$

Finally, set $\alpha := q+4$ and the result follows. □

We can finally state the SA rank lower bound theorem for FO sentences, which is a trivial consequence of Lemma 3.14.

THEOREM 3.16. *Given a first-order sentence ψ that fails in the finite but has an infinite model, there is a constant ε, $0 < \varepsilon \leq \frac{1}{4}$, such that every Sherali-Adams refutation of the propositional translation of ψ, $\mathcal{C}_{\psi,n}$, requires a rank at least $\Omega(n)^{\varepsilon}$, where n is the size of a finite model.*

4. Conclusion and Open problems

Our result leaves a number of open questions:

1. Can the gap be widened, i.e. can we replace polynomial SA by linear SA? Our guess is "yes", and it is supported by concrete examples, such as the Pigeonhole principle and the Least number principle – see the lower bounds in Dantchev *et al.* (2009).

2. Can a gap be proven in a much more general and abstract setting, e.g. for the proof systems whose lines have "small" communication complexity, defined in Beame *et al.* (2010)? These systems include not only LS and SA but also LS_+ and stronger generalisations, such as the Lassere proof system. This could be rather difficult as the gap would be in a different place for different systems as the Pigeonhole principle has a rank two LS_+ proof.

Acknowledgements

A preliminary version of this paper – Dantchev (2007) – was presented at the 39th Annual ACM Symposium on Theory of Computing, San Diego, California. We are grateful to two anonymous referees for interesting observations as well as detailed corrections.

References

PAUL BEAME, TRINH HUYNH & TONIANN PITASSI (2010). Hardness amplification in proof complexity. In *STOC*, LEONARD J. SCHULMAN, editor, 87–96. ACM. ISBN 978-1-4503-0050-6.

PAUL BEAME, TONIANN PITASSI & NATHAN SEGERLIND (2007). Lower Bounds for Lovász–Schrijver Systems and Beyond Follow from Multiparty Communication Complexity. *SIAM J. Comput.* **37**(3), 845–869.

JOSHUA BURESH-OPPENHEIM, NICOLA GALESI, SHLOMO HOORY, AVNER MAGEN & TONIANN PITASSI (2006). Rank Bounds and Integrality Gaps for Cutting Planes Procedures. *Theory of Computing* **2**(1), 65–90. URL `http://www.theoryofcomputing.org/articles/v002a004`.

Moses Charikar, Konstantin Makarychev & Yury Makarychev (2009). Integrality gaps for Sherali-Adams relaxations. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009*, 283–292.

W. Cook, R. Coullard & G. Turan (1987). On the complexity of cutting plane proofs. *Discrete Applied Mathematics* **18**, 25–38.

S. Dantchev (2007). Rank complexity gap for Lovász-Schrijver and Sherali-Adams proof systems. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing, San Diego, California, USA, June 11-13, 2007*, 311–317. ACM. ISBN 978-1-59593-631-8.

S. Dantchev & S. Riis (2003). On Relativisation and Complexity gap for Resolution-based proof systems. In *The 17th Annual Conference of the EACSL, Computer Science Logic*, volume 2803 of *LNCS*, 142–154. Springer.

Stefan S. Dantchev & Barnaby Martin (2009). Cutting Planes and the Parameter Cutwidth. In *CiE*, 134–143.

Stefan S. Dantchev, Barnaby Martin & Mark Rhodes (2009). Tight rank lower bounds for the Sherali-Adams proof system. *Theor. Comput. Sci.* **410**(21-23), 2054–2063.

Stefan S. Dantchev, Barnaby Martin & Stefan Szeider (2007). Parameterized Proof Complexity. In *FOCS*, 150–160. Accepted in Computational Complexity.

Konstantinos Georgiou, Avner Magen & Madhur Tulsiani (2009). Optimal Sherali-Adams Gaps from Pairwise Independence. In *APPROX '09 / RANDOM '09: Proceedings of the 12th International Workshop and 13th International Workshop on Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, 125–139. Springer-Verlag, Berlin, Heidelberg. ISBN 978-3-642-03684-2.

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

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

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

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

CLAIRE MATHIEU & ALISTAIR SINCLAIR (2009). Sherali-adams relaxations of the matching polytope. In *STOC '09: Proceedings of the 41st annual ACM symposium on Theory of computing*, 293–302. ACM, New York, NY, USA. ISBN 978-1-60558-506-2.

TONIANN PITASSI & NATHAN SEGERLIND (2009). Exponential lower bounds and integrality gaps for tree-like Lovász-Schrijver procedures. In *SODA*, 355–364.

P. PUDLÁK (1999). On the complexity of propositional calculus. In *Sets and Proofs*, 197–218. Cambridge University Press. Invited papers from Logic Colloquium'97.

MARK RHODES (2008). Resolution Width and Cutting Plane Rank Are Incomparable. In *MFCS*, 575–587.

S. RIIS (2001). A Complexity gap for tree-resolution. *Computational Complexity* **10**, 179–209.

SØREN RIIS (2008). On the Asymptotic Nullstellensatz and Polynomial Calculus Proof Complexity. In *LICS*, 272–283.

GRANT SCHOENEBECK (2008). Linear Level Lasserre Lower Bounds for Certain k-CSPs. In *FOCS '08: Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, 593–602. IEEE Computer Society, Washington, DC, USA. ISBN 978-0-7695-3436-7.

H. D. SHERALI & W. P. ADAMS (1990). A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics* **3**, 411–430.

Manuscript received 16th April 2010

STEFAN DANTCHEV
School of Engineering and Computing Sciences,
University of Durham, DH1 3LE, UK.
s.s.dantchev@durham.ac.uk

BARNABY MARTIN
School of Engineering and Computing Sciences,
University of Durham, DH1 3LE, UK.
barnabymartin@gmail.com